TALIS Video Study National Report

Research report

November 2020

Jenni Ingram and Ariel Lindorff with Elaine McCann, Anna Riggall, and Naomi Sani
University of Oxford and Education Development Trust

Contents

List of figures

List of tables

Acknowledgements

Executive Summary

Introduction

The Teaching and Learning International Survey (TALIS) Video Study, run by the Organisation for Economic Co-operation and Development (OECD), provides new information on the teaching of mathematics in secondary schools across the eight participating countries/economies: Biobío, Metropolitana and Valparaíso (Chile), Colombia, England (UK), Germany, Kumagaya, Shizuoka and Toda (Japan), Madrid (Spain), Mexico and Shanghai (China). The TALIS Video Study complements the existing TALIS and PISA studies, providing additional evidence on classroom processes by drawing on direct measures of classroom teaching and instruction. By looking directly into the classroom through video-recorded observation and lesson artefact collection, the TALIS Video Study addresses some of the limitations of using teacher self-reported data. Furthermore, the TALIS Video Study provides new and rich information about classroom processes and practices and contributes to current understanding of how they are related to student learning and other outcomes. The Department for Education (DfE) commissioned Education Development Trust (EDT) and the University of Oxford to conduct the TALIS Video Study in England, where data collection was conducted between October 2017 and October 2018.

The OECD is releasing data from the TALIS Video Study as part of two international reports, a policy report, OECD (2020) *Global Teaching Insights: A Video Study of Teaching*[1] and a technical report. This report, which focuses upon England, is published simultaneously with the OECD's Policy Report which it complements by (i) providing a more focused and detailed analysis of the results in England and (ii) analysing differences within England across teachers and classes. It covers the teaching practices of mathematics teachers who participated in the study in England, along with teachers' and students' perceptions of these practices and the relationship between these practices and student outcomes. These outcomes include students' attainment, their personal interest in mathematics and their general self-efficacy (belief in their ability to succeed) in mathematics.

The results refer to the teaching practices of participating teachers and relate to a particular topic in mathematics, thus they should not necessarily be taken as an indication of mathematics teaching practices more widely or at the present time. It should also be noted that the analysis in each chapter uncovers associations (correlations) but cannot establish causal relationships among the different measures collected.

---

[1] Also referred to as the TALIS Video Study

Methodology

The TALIS Video Study focused on the teaching and learning of a single mathematics topic, quadratic equations, across all the participating countries and economies. It used measurement tools specifically designed for the study for the analysis of videos of mathematics teaching and the lesson artefacts[2] accompanying that teaching. The study design was longitudinal, capturing the outcome measures before and after the sequence of lessons that included quadratic equations. The procedures for data collection and coding of videos and artefacts were standardised across the participating countries and economies.

The data collected for TALIS Video Study included video recordings of two lessons from the unit of work[3] that included quadratic equations alongside the classroom artefacts or teaching materials from these videoed lessons and the following lessons. Teachers also completed two questionnaires, one before the start of the unit and one after the end of the unit. The students in these teachers' classes also completed questionnaires before and after the teaching of the unit, and also completed achievement tests before and after the teaching of the unit. Chapter 1 of this report describes these data collection methods in more depth.

The TALIS Video Study used a stratified, two-stage probability sampling design. Initially 100 schools were randomly selected from the school rosters used for TALIS 2018, with a further 200 replacement[4] schools identified. Once a school agreed to participate, three mathematics teachers were randomly selected from within the school and approached for participation, one at a time in the order they were selected. Once a teacher agreed to participate, consent was sought from the teacher and their students. Only classes where at least 15 students or 50% of the class consented were included in the study.

All differences and associations reported are statistically significant at the 5% level. Further details of the methodology and the analysis can be found in the England Technical Report[5].

This report focuses on England. Chapter 2 describes the profile of the teachers and students in England who participated in the TALIS Video Study. 85 teachers and 2,024 students from 78 schools participated. The proportion of mathematics teachers in the TALIS Video Study for England who were female was 58% and on average teachers had ten years of teaching experience. Around a quarter of teachers reported holding at least a Master's-level qualification. The average class size in the TALIS Video Study in

---

[2] Lesson artefacts included lesson plans, handouts and worksheets, textbook pages, visual materials such as the projected slides shown, and/or any homework set where they were available.
[3] The unit of work was the series of lessons that the teacher identified as including the topic of quadratic equations.
[4] Countries tried to recruit initially from the 100 "main sample" schools but were given two replacements (triplet system) for each in the event they declined to participate or could not be contacted.
[5] McCann, Riggall, Ingram, and Lindorff (2020)

England was 27.4 students. The majority of students participating were in Year 10, though the study also included students in Years 8, 9, and 11 depending on when the topic of quadratic equations was taught in the participating school.

Key Findings

Classroom management

Classes were generally well managed, characterised by organised and efficient routines, and frequent monitoring of the entire classroom, and disruptions were handled quickly and effectively. Most of the lesson time observed was spent on mathematics learning.

Teachers frequently worked with the whole class and students were also given several opportunities to work individually. Teachers rarely made use of small group work or pair work.

When students and teachers were asked about their perceptions, they generally reported few disciplinary issues, high levels of teacher awareness of the classroom, and efficient handling of disruptions. However, teachers perceived fewer disruptions than their students.

Social and emotional support

Teachers and classrooms were generally socially and emotionally supportive. Teachers and students frequently and consistently demonstrated respect for one another through manners, language, and tone of voice. Teachers regularly provided encouragement to their students and shared moments of warmth in their lessons.

Almost all teachers requested that students shared their thinking processes or rationales, such as by explaining or describing procedures taken or their reasoning behind procedures taken. Students in all classes sought guidance and shared their work publicly. Almost every teacher supported students to persist through their mathematical errors or struggles for a moderate length of time at least once in the videoed lessons.

The vast majority of teachers and students reported experiencing positive relationships within their mathematics classrooms. The majority of students also reported that their mathematics teacher helped them with their learning, gave them extra help when they needed it, and made them feel confident to do well in the topic. Teachers generally were more positive in their perception of the social-emotional environment than students.

Classroom discourse

Teachers and students were involved in classroom discourse. Lessons included questions that requested students to recall, report an answer, or define terms, and

questions that requested students to summarise, explain, classify, or apply processes. There were also some lessons that included questions that requested students to analyse, synthesise, justify, or conjecture. On average lessons included explanations of why ideas or procedures are the way they are, though in many lessons these focused on brief or superficial features of the mathematics. In the lesson artefacts, which included worksheets, homework tasks, and teacher presentations, students were not often asked for explanations of how or why mathematical procedures or relationships work.

Around three-quarters of teachers and students reported that there were frequently opportunities for students to explain their ideas. Around half of the teachers and half of the students reported that there were also frequent opportunities to critique arguments made by other students or that students had opportunities to engage in discussion amongst themselves.

Classes with a higher average score on the pre-test had a positive association with the average video rating overall within the discourse domain, as well as higher ratings on each of the nature of discourse, questioning, and explanations components than classes with a lower average score on the pre-test. This means that there were more detailed student contributions and explanations, and students were asked more questions that require them to analyse, synthesise, justify, or conjecture in classes with higher average prior-attainment compared to classes with lower average prior-attainment.

The quality of subject matter

Lessons were generally clear though explicit connections and explicit patterns and generalisations were not commonly used. Almost all teachers included some type of explicit goal in both their lessons in the study, and around two thirds (68%) of the lessons had explicit goals that focused on student learning. The use of real-world connections was rare in both the lesson videos and artefacts. A quarter of teachers included no real-world connections in any of the artefacts from their four lessons. Less than half of the lessons included a connection to other mathematical topics. Equations were used frequently, as would be expected for lessons focusing on quadratic equations. Graphs were used by many teachers, and slightly fewer teachers used tables, drawings or diagrams. Physical objects or models were very rarely used. Three-quarters of the teachers (89%) asked students to make a connection between these mathematical representations in at least one lesson.

Teachers and students reported that lessons were adapted to the class's needs and knowledge and that teachers changed their way of explaining when a student had difficulties understanding a topic or task. This adaption of the lessons did not always involve giving different work to students with different attainment levels, with 31% of teachers and 51% of students reporting this did not happen. They also reported that

lessons included a summary of recently learned content, that goals were set, and that teachers explained how new and old topics were related.

Student cognitive engagement with subject matter

Students were rarely asked to engage with cognitively demanding subject matter. The opportunities to use multiple-solution strategies or to engage in the rationale for processes and procedures were limited. All teachers asked their students to engage in analyses, creation, or evaluation work that was cognitively rich and required thoughtfulness at some point in their two observed lessons.

In the lesson artefacts, students were rarely asked to use multiple mathematical methods. Opportunities to repetitively use a specific skill or procedure were frequent in both the lesson videos and the accompanying artefacts. Teachers mostly used technology for communication purposes, rather than for conceptual understanding of mathematics, and students rarely used technology at all.

Most students reported that their teacher gave them tasks that required them to think critically and tasks that required them to apply what they had learned to new contexts. Most teachers and students also reported that teachers frequently explained why a mathematical procedure works, illustrated why a mathematical procedure works using concrete examples or graphics, and compared different ways of solving problems. Students reported that lessons were more cognitively demanding than their teachers did.

Assessment of and responses to student understanding

Teachers generally asked questions, gave prompts or used tasks that elicited student responses, responded to students' thinking using feedback loops that focused on why the students' thinking was correct or incorrect or why ideas or procedures are the way they are, used students' contributions, and provided clues or hints to support student understanding when students made errors or struggled mathematically.

The lesson artefacts rarely provided students with opportunities for self-evaluation that involved explicit reflection on their understanding.

There were positive associations between the proportion of students within classes who were female, the proportion of students who did not speak English at home, the average class pre-test score and the average level of parental education, and the overall average domain rating for assessment of and responses to student understanding.

Opportunity to learn (OTL)

On average teachers spent 7.6 hours on the topic of quadratic equations.The most common subtopic in the TALIS Video Study lessons was handling algebraic expressions,

which included working with brackets and algebraic terms. The most common solution method was solving quadratic equations by factorising. Fewer than a fifth of lessons included applying quadratic equations to real-life contexts.

Students' prior attainment was associated with the extent of OTL reported by students, with students in classes with higher average pre-test scores indicating a greater extent of OTL compared to students in classes with lower average pre-test scores. Students' perceived OTL was positively associated with students' attainment on the post-test, their personal interest in mathematics, and their self-efficacy in mathematics.

Relationships between student characteristics, teaching practices and student outcomes

Students with higher scores on the pre-test tended to score higher on the post-test, and vice versa, as would be expected. Pre-test scores explained half of the variation in post-test scores showing the strength of this relationship. All other student characteristics, such as language spoken at home or socio-economic status, by contrast, explained very little of the variation in post-test scores. After pre-test scores had been accounted for, no student characteristics had any significant relationships with the post-test scores except for home possessions. Students with higher scores on the home possessions scale tended to have higher scores on the post-test. Given the short time between pre-test and post-test, the pre-test score likely accounts for relationships that exist between student demographic characteristics and attainment in general that has been shown in other research[6]. When classes were grouped into four groups (quartiles) by their average scores on the pre-test, students taught in classes with lower pre-test performance on average tended to have lower scores on the post-test even after taking into consideration their individual pre-test attainment. For students with higher individual pre-test scores, the average level of pre-test attainment in their class made more of a difference to their post-test attainment; students in the lowest two groups of average class pre-test attainment were likely to have lower post-test scores than their peers with the same pre-test attainment in higher-attaining classes.

Students with higher self-efficacy with their previous teacher tended to have higher self-efficacy with their current teacher. General self-efficacy with students' previous teacher explained 22% of the variation in students' general self-efficacy with their current teacher. Students with higher personal interest in mathematics with their previous teacher tended to have higher personal interest in mathematics with their current teacher. Personal interest in mathematics with students' previous teacher explained 12% of the variation in personal interest in mathematics with their current teacher.

The teaching practice domains were combined in order to examine relationships between student outcomes and teaching practices. The discourse domain and the assessment of

---

[6] Strand (2014), Sammons et al. (2014)

and responses to student understanding were grouped as these reflected teaching practices that related to teaching in a range of curriculum areas. The quality of subject matter and student cognitive engagement domains were combined as these reflect teaching practices that are often specific to mathematics classrooms. There were no relationships between classroom management, social-emotional support, discourse and assessment, or mathematics instruction and student attainment on the post-test, general self-efficacy with the current teacher, or personal interest in mathematics with the current teacher before accounting for student and school characteristics.

Class average pre-test attainment made more of a difference to students in classes with teachers who had lower ratings for classroom management. Students in classes with lower average general self-efficacy with their previous teachers had lower general self-efficacy with their current teacher than students in other groups if the teacher's classroom management ratings were low, and higher general self-efficacy with their current teacher than students in other groups if the teacher's classroom management ratings were high. For students in classes with the lowest average personal interest in mathematics with the previous teacher, the teacher's observed social-emotional support made less of a difference to personal interest with the current teacher. For students in classes with higher personal interest in mathematics with the previous teacher, the higher a teacher's observed social-emotional support ratings, the higher students' personal interest in mathematics with the current teacher. There was also an overall significant positive relationship between social-emotional support and personal interest in mathematics with the current teacher after including this interaction.

1. Introduction

This report presents the results of the Teaching and Learning International Survey (TALIS) Video Study of mathematics teaching in England. This chapter gives the background to the TALIS Video Study and its implementation in England. It also outlines the structure of the report.

What is the TALIS Video Study?

The Teaching and Learning International Survey (TALIS) Video Study, run by the Organisation for Economic Co-operation and Development (OECD), provides new information on the teaching of mathematics in secondary schools across the eight participating countries/economies. The study was conducted in England between October 2017 and October 2018. This was the first TALIS Video Study, complementing the Teaching and Learning International Survey (TALIS) of teachers that was conducted in 2013 and 2018 in England. The Department for Education (DfE) commissioned Education Development Trust (EDT) and the University of Oxford to conduct the TALIS Video Study in England.

Findings from the TALIS Video Study are compared to the findings from TALIS 2018 and PISA 2018 to determine the similarity of the TALIS Video Study sample to the diverse population of teachers and students in England. Indeed, this study complements the existing TALIS and PISA studies by providing additional evidence on classroom processes by drawing on direct measures of classroom teaching and instruction. Many of the questions from the questionnaires in the TALIS Video Study are also part of the teacher questionnaire in TALIS 2018, or are part of the student questionnaire in PISA 2018. However, as Chapter 2 outlines, there are differences in the samples for TALIS 2018, PISA 2018 and the TALIS Video Study which limit the comparisons that can be made.

Background to the TALIS Video Study

The data collected through the TALIS Video Study provides information on mathematics teaching practices in England and a number of other countries. By looking directly into the classroom through videoed observation and lesson artefact collection, the TALIS Video Study addresses some of the limitations of using teacher self-reported data. Furthermore, the TALIS Video Study provides new and rich information about classroom processes and practices, and contributes to current understanding of how they are related to student learning and other outcomes.

The goals of the TALIS Video Study were:

- to understand which aspects of teaching are related to student learning and student non-cognitive outcomes

- to observe and document how teachers from participating countries or economies teach

- to explore the interrelationships of various teaching practices and the relationships among contextual aspects, such as teacher characteristics, teacher education and student characteristics, and classroom teaching and learning.

Participating countries

Eight countries or jurisdictions participated in TALIS Video Study: Biobío, Metropolitana and Valparaíso (Chile), Colombia, England (UK), Germany, Kumagaya, Shizuoka and Toda (Japan), Madrid (Spain), Mexico, and Shanghai (China). The findings from all these countries or jurisdictions are reported in the OECD Policy Report[7], whilst this report focuses on England.

What did the TALIS Video Study measure?

The TALIS Video Study collected data using teacher and student questionnaires and assessments, as well as videos of mathematics lessons with the accompanying artefacts for those lessons and the subsequent lesson. The study focuses on the secondary phase mathematics topic of quadratic equations. The lesson videos and artefacts were rated across six domains, which were aspects of teaching that previous research has shown to support students' learning, including: *Classroom Management*, *Social-Emotional Support*, *Discourse*, *Quality of Subject Matter*, *Engagement in Cognitively Demanding Subject Matter*, and *Assessment of and Responses to Student Understanding*.

In England, the topic of quadratic equations is usually taught across a three-year period, between Years 9 and 11. Many of the subtopics considered within the focal topic of quadratic equations for the TALIS Video Study are not in the National Curriculum for the majority of students in England. However, solving quadratic equations by factorising and identifying roots of quadratic functions graphically, as well as moving between numerical, algebraic, graphical, and diagrammatic representations of quadratic functions, are included. Other methods for solving quadratic equations, such as completing the square or by using the quadratic formula, as well as solving quadratic equations by factorising where the coefficient of the quadratic term is different from 1 (e.g. $18x^2 - 3x = 6$), are only included in the Higher Tier examinations and are often only taught to more highly attaining students[8]. In 2018 and 2019, over half of the students taking GCSE

---

[7] OECD (2020a)
[8] Department for Education (2014)

Mathematics were entered for the Foundation Tier[9]. Further details of the measures used in the TALIS Video Study can be found in the England Technical Report[10].

The questionnaires

The TALIS Video Study required all participating teachers and students to complete two questionnaires. The first of these questionnaires was completed before the unit of teaching within which quadratic equations was taught. In total, 1,923 students and 85 teachers completed the first questionnaire. The second questionnaire was completed after the end of the unit of teaching. In total, 1,883 students and 84 teachers completed the second questionnaire.

The teacher questionnaires asked teachers about their background and education, their beliefs, their motivation and their perception of the school environment. They were also asked about the class participating in the study and about their teaching during the unit on quadratic equations. This included questions about their goals for the lessons, the mathematical content covered, their teaching practices and their judgement of their effectiveness in teaching the unit. The teachers were also asked whether the videoed lessons represented their typical teaching. Additional questions were asked of teachers in England that focused on their experiences of Continuing Professional Development (CPD). Teachers also completed a Teacher Log of the topics they taught during the unit as part of the questionnaires.

The student questionnaires asked students about their background, their attitudes and feelings towards mathematics, and the learning and teaching of mathematics. In particular, they included several items asking students about their personal interest in mathematics and their self-efficacy in relation to mathematics. The first questionnaire focused on students' attitudes and feelings towards mathematics with their previous mathematics teacher, whilst the second questionnaire focused on students' attitudes and feelings with their current mathematics teacher.

Teachers were also asked to log the number, length, and content of subtopics (e.g. real-life applications, solving quadratic equations by "completing the square") for the lessons throughout the unit on quadratic equations.

The assessments

Students took a pre-test focused on their general mathematics knowledge two weeks before the start of the unit on quadratic equations. They then took a post-test within two weeks of conclusion of the unit. The post-test had a narrower focus than the pre-test in order to provide more precise measures of students' knowledge and understanding of

---

[9] Ofqual (2019)
[10] McCann, Riggall, Ingram, and Lindorff (2020)

quadratic equations. In England, 1,892 students completed the pre-test and 1,873 students completed the post-test.

The videos

Two lessons from the unit on quadratic equations were videoed; one occurred during the first half of the unit, and the second occurred later in the unit. This enabled the study to consider differences in the purpose of each lesson, as lessons that introduce a new topic may differ from those where students are developing their fluency.

Videos were rated separately for two different types of observation codes: components and indicators. These codes depended upon whether and how the teaching practices could be observed and evaluated by a rater and are detailed in the subsequent chapters for each domain and the England Technical Report[11]. Videos were rated by two different trained raters for each type of video rating. Components were higher inference codes that were used to rate the lesson videos. Video component ratings were rated every 16 minutes. A holistic code for each domain was also given by raters which was applied to the components within each domain. All components and holistic ratings were rated on a four-point scale. Indicators were lower inference codes that either categorised or rated interactions. Indicators were rated every eight minutes. The ratings for video components and indicators focus on the quality, nature, or presence of a specific practice and did not consider whether some practices were better than others.

The artefacts

In addition to the videos of lessons, artefacts from those lessons, and the lessons that followed, were also collected. These included lesson plans, handouts and worksheets, textbook pages, and visual materials – such as the projected slides shown and/or any homework set – where they were available. Student work completed during the lessons was not collected. Artefacts were collected from each of the videoed lessons as well as from each subsequent lesson, resulting in four artefact sets for each teacher. In addition, the next formal examination that included quadratic equations was also collected to document the formal expectations for student understanding.

All artefacts were rated by two different trained raters across the four domains that related specifically to mathematics teaching: *discourse*, *quality of subject matter*, *student engagement in cognitively demanding subject matter*, and *assessment of and responses to student understanding*. Further details for the artefact components are detailed in the subsequent chapters for each domain where they are used and in the England Technical Report[12].

---

[11] McCann, Riggall, Ingram, and Lindorff (2020)
[12] Ibid.

Overview of the report

The rest of this report is divided into seven main chapters which present the results from the TALIS Video Study. Each chapter is organised around a series of questions which form the headings for each section. A summary at the start of each chapter provides some key findings.

Chapter 2 considers the profile of the teachers, students, classes and schools that participated in the TALIS Video Study, including their background characteristics.

Chapter 3 describes the *classroom management* of the lessons included in the TALIS Video Study, including teachers' and students' perspectives. This chapter begins with a description of the ratings of the lesson videos for the *classroom management* domain. It then examines teachers' and students' perceptions of classroom management, before examining the experiences that classes with different characteristics have of classroom management.

Chapter 4 describes the *social and emotional support* environment of the classrooms included in the TALIS Video Study, including teachers' and students' perspectives. The chapter begins with a description of the lesson video ratings for the *social and emotional support* domain. It then examines teachers' and students' perspectives of the *social and emotional support* during the topic of quadratic equations. The chapter ends by examining the differences between teachers with different background characteristics and between classes with different characteristics on the average video ratings for the domain.

Chapters 5, 6, 7 and 8 focus on the mathematical aspects of teaching measured in the domains within the study, with Chapter 5 focusing on *discourse*, Chapter 6 focusing on the *quality of subject matter*, that is the quality of mathematics in the lesson, Chapter 7 focusing on *student engagement in cognitively demanding subject matter*, and Chapter 8 focusing on teachers' *assessment of and responses to student understanding*. Each chapter begins with a description of the overall video ratings for the domain. This is followed by an examination of teachers' and students' perspectives of the lessons within the topic of quadratic lessons in relation to the domain.

Chapter 9 describes the opportunities that students had to learn different aspects of quadratic equations. This chapter explores students' exposure to different solution methods, and their opportunities to work with mathematics in different ways. The chapter begins by describing the *opportunities to learn* as reported by teachers. Next the *opportunities to learn* as shown in the artefacts are described. The chapter ends by examining the relationship between *opportunities to learn* and student and class characteristics.

Chapter 10 explores the relationships between student characteristics and students' achievement and other non-cognitive outcomes (personal interest in mathematics and

self-efficacy in mathematics) as well as between different teaching practices and those outcomes.

More detailed analyses of international results can be found in the OECD report on the TALIS Video Study, which also includes results for England[13]. The OECD will also publish full details of their analysis in the [OECD Technical Report](#)[14].

Interpreting the findings

There is inherent variation in human populations which can never be summarised with absolute accuracy. All research and data collection that uses samples, such as the sampling of schools, teachers and students in the TALIS Video Study, is affected by *sampling error*. The results reported here are a best estimation of the total population of mathematics teachers teaching the topic of quadratic equations. Another random sample of schools and teachers would likely provide slightly different results. Statistical methods are used to measure how good the estimations are.

Measurement error relates to the results obtained by each individual teacher and student. It takes into account variations in their scores or ratings which are not directly due to the underlying attribute or behaviour being measured, but which are influenced by other factors related to individuals or to the nature of the instruments used in data collection.

As a consequence of these two areas of uncertainty, interpretations of differences between two sets of results are often meaningless. Repeating the measurements could affect both the size and the direction of any differences. For this reason, this report focuses on *statistically significant* differences in average scores or ratings. Statistically significant differences are unlikely to be a result of random fluctuations due to sampling or measurement error. Where significant differences are reported, the significance level is taken to be 5%. Similarly, when interpreting results of regression models, this report focuses on those coefficients that are statistically significant at the same level (5%). Where statistically significant differences between teachers or classes are found, these may be a result of a number of factors. For some of these factors data was not collected in the TALIS Video Study. Therefore, the analysis is only able to explain the reasons for differences between teachers to a limited extent.

There are occasions where some figures or proportions represented within a table do not sum to the total due to rounding. This is because of a combination of rounding and missing values. Throughout the report the arithmetic mean average has been used unless stated otherwise. The distribution of teacher average ratings on many of the measures in TALIS Video Study are illustrated using density plots to show the general shape of the distributions. The more peaked the curve is in a density plot, the more

---

[13] OECD (2020a)
[14] OECD (2020b)

classrooms in England have average scores concentrated around a few values, that is the more densely populated that range is under the curve. Note that figures such as density plots used in this report may be sensitive to kernel smoothing, so readers should focus on the overall shape of these figures rather than individual features that represent small numbers of classrooms.

Disclaimer

The TALIS Video Study is an OECD project. The development of the Study's instrumentation and data analyses and drafting of international reports were contracted by the OECD to RAND, ETS[15] and DIPF[16]. The authors of this work are solely responsible for its content. The opinions expressed and arguments employed in this work do not necessarily represent the official views of the OECD or its member countries.

---

[15] Educational Testing Service
[16] Leibniz Institute for Research and Information in Education

2. The characteristics of teachers, students, and schools

Key findings

- In the TALIS Video Study, 85 teachers from 78 schools and a total of 2,024 students participated in England.

- Fifty-eight per cent of the mathematics teachers were female. This was a smaller proportion than in TALIS 2018 (64%) which surveyed teachers from all curriculum areas in lower secondary schools.

- The average teacher in the TALIS Video Study had 10 years of teaching experience. This was lower than the average teaching experience in TALIS 2018 for England, and lower than the other participating countries in the TALIS Video Study.

- The qualifications of the teachers participating in the TALIS Video Study were similar to those surveyed in TALIS 2018, with around a quarter of secondary teachers in England reporting they held at least a Master's-level qualification.

- Sixty-nine per cent of the TALIS Video Study mathematics teachers had taken mathematics courses equivalent to those needed for a degree in their mathematics education or training.

- The vast majority (85%) of teachers reported that they were aware of and understood the 2016 Standard for Teachers' Professional Development, and 91% felt that the professional development they had done had had a positive impact on their teaching.

- The average class size in the TALIS Video Study in England was 27.4 students, which is larger than the average class size of 24.5 reported in TALIS 2018.

- The majority of students participating were in Year 10, though the TALIS Video Study also included students in Year 8, 9 and 11.

- All participating schools were state funded, and 74% of schools were in urban areas.

Introduction

This chapter describes the characteristics of the teachers, students, classes, and schools that participated in the TALIS Video Study.

In England, 85 teachers from 78 schools and a total of 2,024 students participated. This chapter begins by describing the characteristics of the participating teachers and comparing them to the lower secondary teachers from England who participated in the TALIS 2018 study. The chapter then explores the characteristics of the participating students and compares them to those of the students from England who participated in PISA 2018. Finally, the characteristics of the participating classes and schools are explored.

The overall TALIS Video Study approach to recruitment sought to achieve a random sample of mathematics teachers in each participating country, however it is difficult to ascertain whether the England sample is representative, since it was necessary to approach more teachers than originally envisaged to gain agreement for video recording of lessons. In addition, to the authors' knowledge there is no national source of data on the characteristics of secondary teachers of mathematics for particular age groups in England and so comparisons with a national pattern are precluded. Comparisons with the sample of secondary teachers from all curriculum areas that participated in TALIS 2018 are made which reveal both similarities and differences between the samples.

The characteristics of teachers

The TALIS 2018[17] study showed that, in most nations, more women than men are teachers. This was also found in the TALIS Video Study with 58% of participating teachers being female, compared to 64% of all lower secondary teachers in England and an OECD average of 68% female teachers in the TALIS 2018 study. The TALIS Video Study only includes mathematics teachers so the difference in proportions of female teachers from TALIS 2018 may be a result of the focus on just one group of teachers rather than all lower secondary teachers as in TALIS 2018.

Similar to the England TALIS 2018 sample, the England sample in the TALIS Video Study included more relatively inexperienced teachers compared to other countries. The difference was even more pronounced in the TALIS Video Study where teachers had on average 10.00 years' experience teaching and 9.89 years teaching mathematics. This compares to an average of 13 years' teaching experience in England in TALIS 2018 which was also around 4 years lower than the OECD average. Figure 2.1 illustrates the mathematics teaching experience of the participating teachers. The proportion of teachers in their early career (five or fewer years' teaching experience) was 35%.

---

[17] Jerrim and Sims (2019)

Figure 2.1: Years of mathematics teaching experience for participating teachers
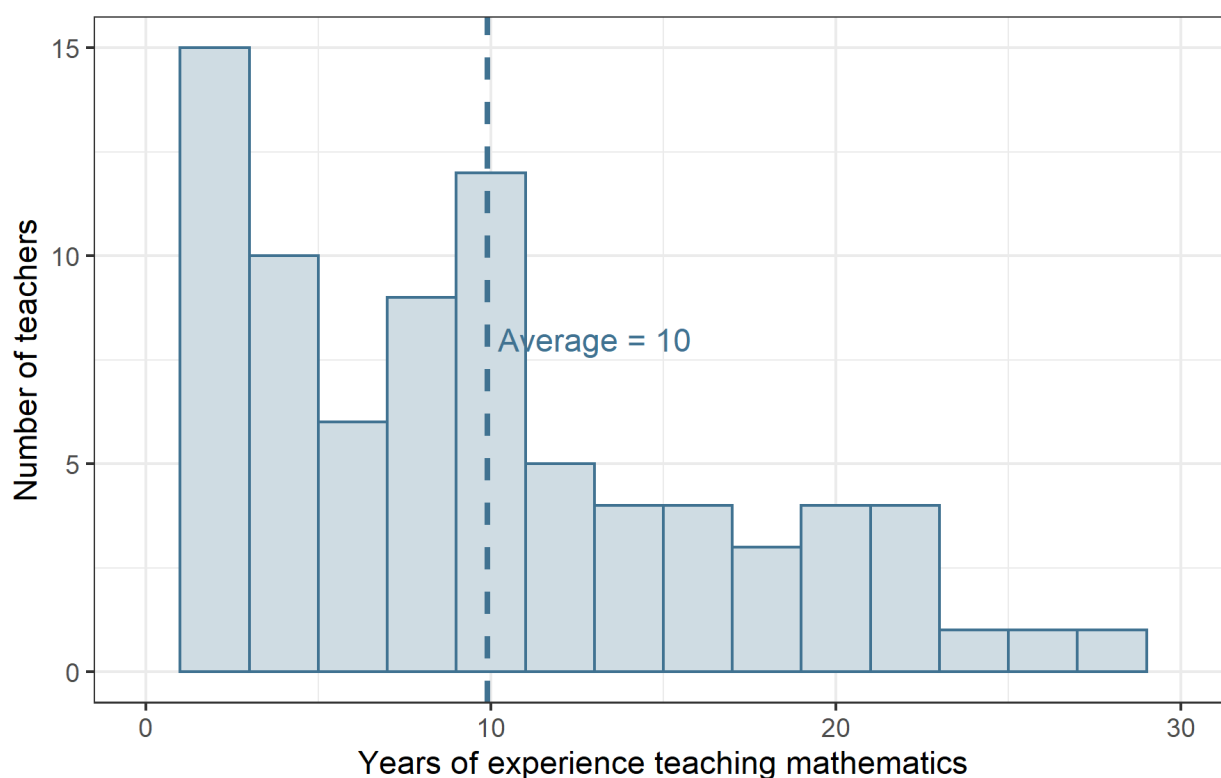
Source: TALIS Video Study teacher data file

The level of qualification held by teachers in the TALIS Video Study was very similar to the level held by the teachers in the England sample for TALIS 2018: around a quarter of secondary teachers in England reported holding at least a Master's-level qualification. This is below the OECD average (44%)[18]. The majority of teachers (84%) in the TALIS Video Study received their teaching qualification through a standard teacher education route with an Initial Teacher Training (ITT) provider.

Teachers were asked a series of questions about their own mathematics education: 82% had completed at least a Bachelor's Degree in a mathematics-related subject; 13% had completed a Master's Degree or Doctorate in a mathematics-related subject; and 69% of participating mathematics teachers had taken mathematics courses equivalent to those needed for a degree in their mathematics education or training.

---

[18] ISCED (International Standard Classification of Education) is an international classification of educational qualifications. This provides a framework to facilitate comparisons of educational qualifications across countries. However, previous research has suggested that not all qualifications fit easily into the ISCED classification schema (Schneider, 2008).

Table 2.1: Comparison of TALIS 2018 and TALIS Video Study teacher demographics (England)

Characteristics	TALIS 2018	TALIS Video Study
Female	64%	58%
Teaching experience (years)	13	10
Highest Degree: Undergraduate	73%	73%
Highest Degree: Master's	24%	24%
Highest Degree: Doctorate	2%	1%

Source: TALIS Video Study teacher data file and TALIS 2018 dataset

Teachers were also asked about their self-efficacy in relation to teaching the participating class. As illustrated in Figure 2.2, teachers' self-efficacy was generally fairly high, with teachers stating that they were confident in their abilities to support student learning in a variety of ways.

Figure 2.2: Teachers' self-efficacy
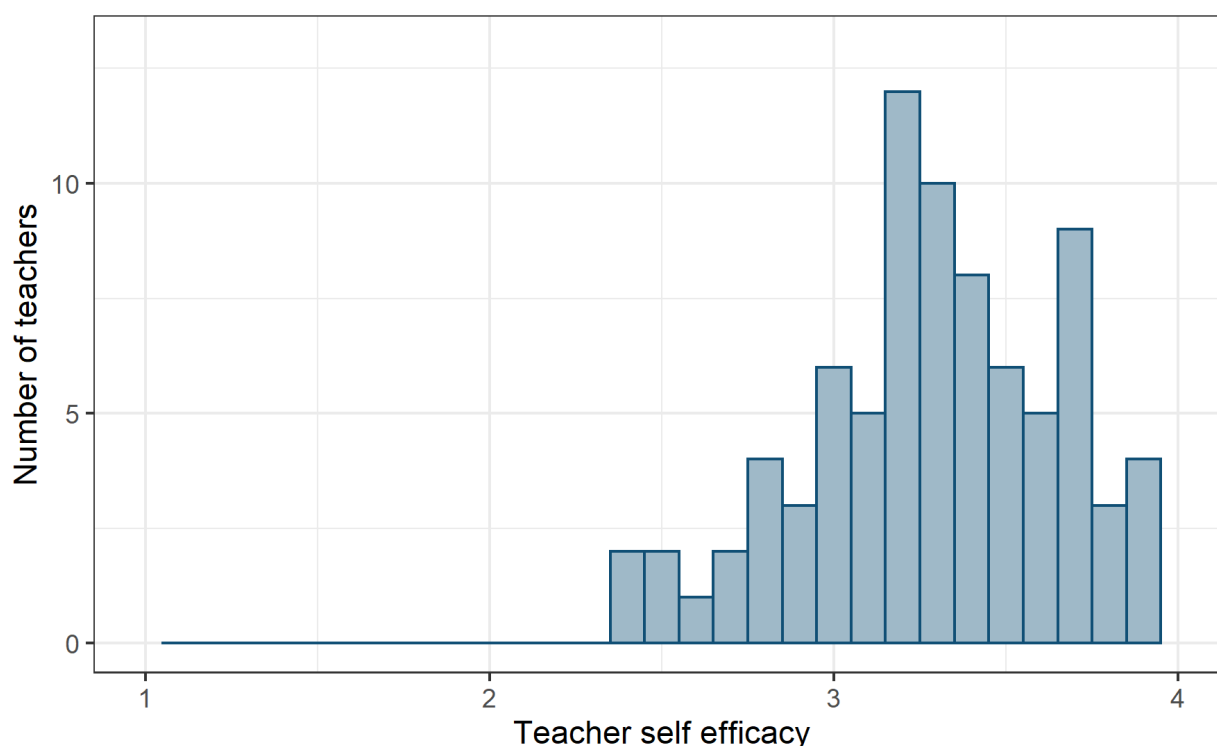
Note: Items measuring self-efficacy used a four-point scale
Source: TALIS Video Study teacher data file

What were teachers' beliefs and attitudes about teaching mathematics?

The pre-questionnaire also asked teachers about their beliefs and attitudes about teaching mathematics. Teachers in England largely felt responsible for their students' learning and progress, and in particular for the effectiveness of their own teaching. There was more of a mixed response about their feelings of responsibility for students liking or valuing mathematics.

Almost all the teachers (92%) agreed or strongly agreed that students should be allowed to think of solutions to practical problems themselves before being shown how they are solved, as shown in Figure 2.3. Around three-quarters of teachers (76%) agreed or strongly agreed that their role as a teacher was to facilitate students' own inquiry and that students learn best by finding solutions to problems on their own (72%). The vast majority of teachers (80%) also stated that thinking and reasoning processes were more important than specific curriculum content.

Figure 2.3: Teachers' beliefs about teaching mathematics
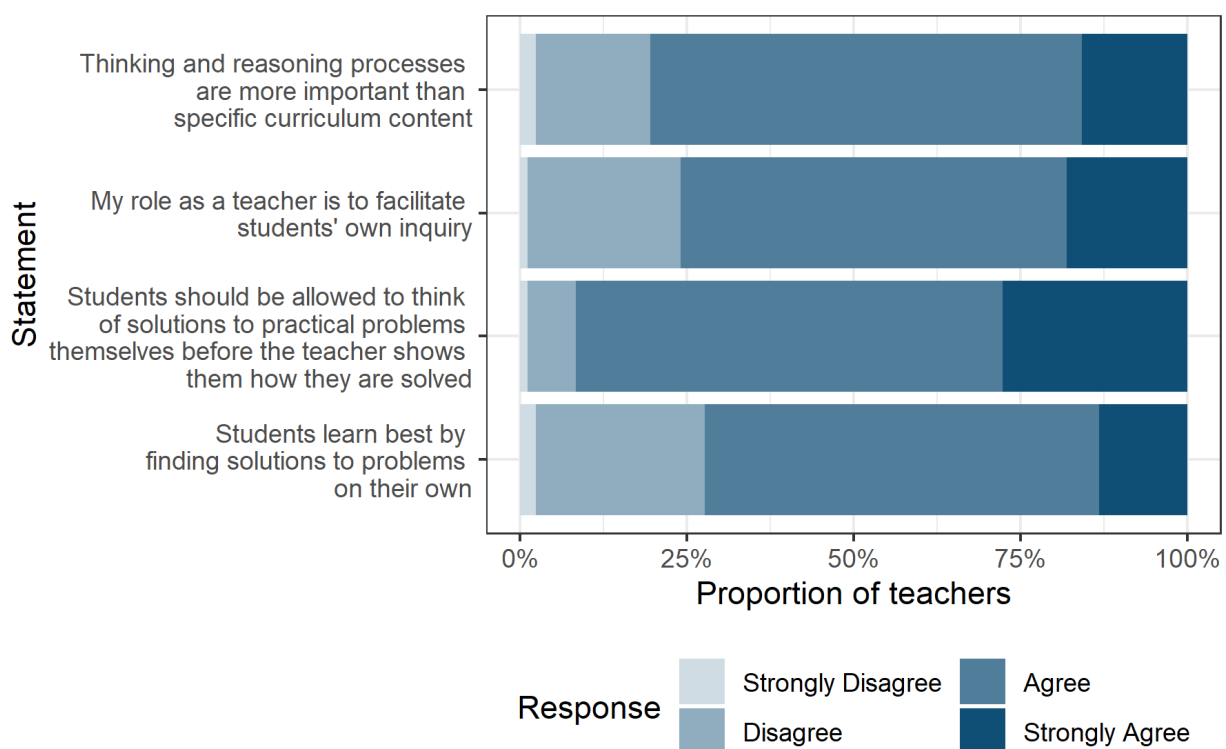
Source: TALIS Video Study teacher data file

What were teachers' experiences of observation and videoing of their teaching?

Teachers were asked how often their teaching had been videoed during their teaching career. The majority reported that they had never been videoed, as can be seen in Table 2.2, with only three teachers reporting that they had been videoed more than ten times in their career. When asked how often their teaching had been observed, 87% reported that their lessons were observed occasionally during the school year, with 29% stating that this happened frequently.

Table 2.2: Teachers' experiences of their lessons being videoed

Frequency of being videoed	Number of teachers	Proportion of teachers
Never	43	51%
1-3 times	27	32%
4-10 times	12	14%
More than 10 times	3	4%

Further information on responses to these items can be found in the England Technical Report

Source: TALIS Video Study teacher data file

How satisfied were mathematics teachers with their job as a teacher?

The teachers in the TALIS Video Study were very positive about their role as a teacher, as shown in Table 2.3. Almost all the teachers agreed or strongly agreed that they enjoyed working at their current school and that they would recommend this school as a good place to work. The vast majority also agreed or strongly agreed that if they could decide again, they would still choose to work as a teacher, and only two teachers regretted their decision to become a teacher. Around a third of teachers agreed or strongly agreed that the teaching profession was valued in society.

Table 2.3: Job satisfaction of mathematics teachers

Statement	Number of teachers	Proportion of teachers
The advantages of being a teacher clearly outweigh the disadvantages	71	85%
If I could decide again, I would still choose to work as a teacher	73	86%
I would like to change to another school if that were possible	13	15%
I regret that I decided to become a teacher	2	2%
I enjoy working at this school	79	93%
I wonder whether it would have been better to choose another profession	37	44%
I would recommend this school as a good place to work	78	92%
I think that the teaching profession is valued in society	30	35%
I am satisfied with my performance in this school	77	91%
All in all, I am satisfied with my job	74	87%

Further information on responses to these items can be found in the England Technical Report
Proportions and numbers represent teachers who agreed or strongly agreed with each statement
Source: TALIS Video Study teacher data file

Teachers were also positive about the professional development they received and the guidance they had received on the Standard for Teachers' Professional Development published in 2016[19]. Most teachers (86%) reported that they were aware of the Standard and understood what they meant. The majority of teachers were also using the Standard to inform their development decisions. Almost all teachers (91%) felt that their professional development was having a positive impact on their teaching practice. On average, teachers reported taking part in 2.2 hours of training and other activities related to continuing professional development (CPD) in their most recent working week, and 76% agreed or strongly agreed that they were satisfied with the training and CPD activities that they had undertaken in the last 12 months.

Teachers were also asked what influenced their decision to pursue a career in teaching. The most common reason chosen was a desire to "make a difference in young people" but three-quarters were also driven by "a desire to encourage and inspire children and

---

[19] Department for Education (2016)

young people to pursue mathematics". Around a quarter of the participating teachers had changed their career to pursue teaching.

The characteristics of students

This section describes the students' background characteristics and how they compare with the students who participated in PISA 2018. PISA was a representative study of 15-year-old students, while the TALIS Video Study focused on year groups where quadratic equations are taught. Differences in student age and year group are a consequence of the design of the TALIS Video Study. Differences in gender, first- or second-generation immigrant background, and parental education are likely to be a consequence of the sampling, including a smaller sample size, which may mean that the sample for the TALIS Video Study is not similar to the diverse population of students in England that PISA 2018 is representative of.

What were the participating students' backgrounds?

The participating students were in Years 8, 9, 10, and 11, with the vast majority of students in Year 10, as can be seen in Table 2.4. These year groups are equivalent to international grades 7, 8, 9, and 10 respectively. This differs from PISA 2018 where the target population was 15-year-olds.

Table 2.4: Distribution of the year groups of students

Year group	Number of students	Proportion of total participating students	Number of classes
8	59	3%	2
9	369	18%	14
10	1,442	71%	62
11	154	8%	7

Source: TALIS Video Study student data file

The average age of the participating students was 15 years, ranging from 12 to 16 years old, as can be seen in Figure 2.4.

Figure 2.4: Ages of the students participating
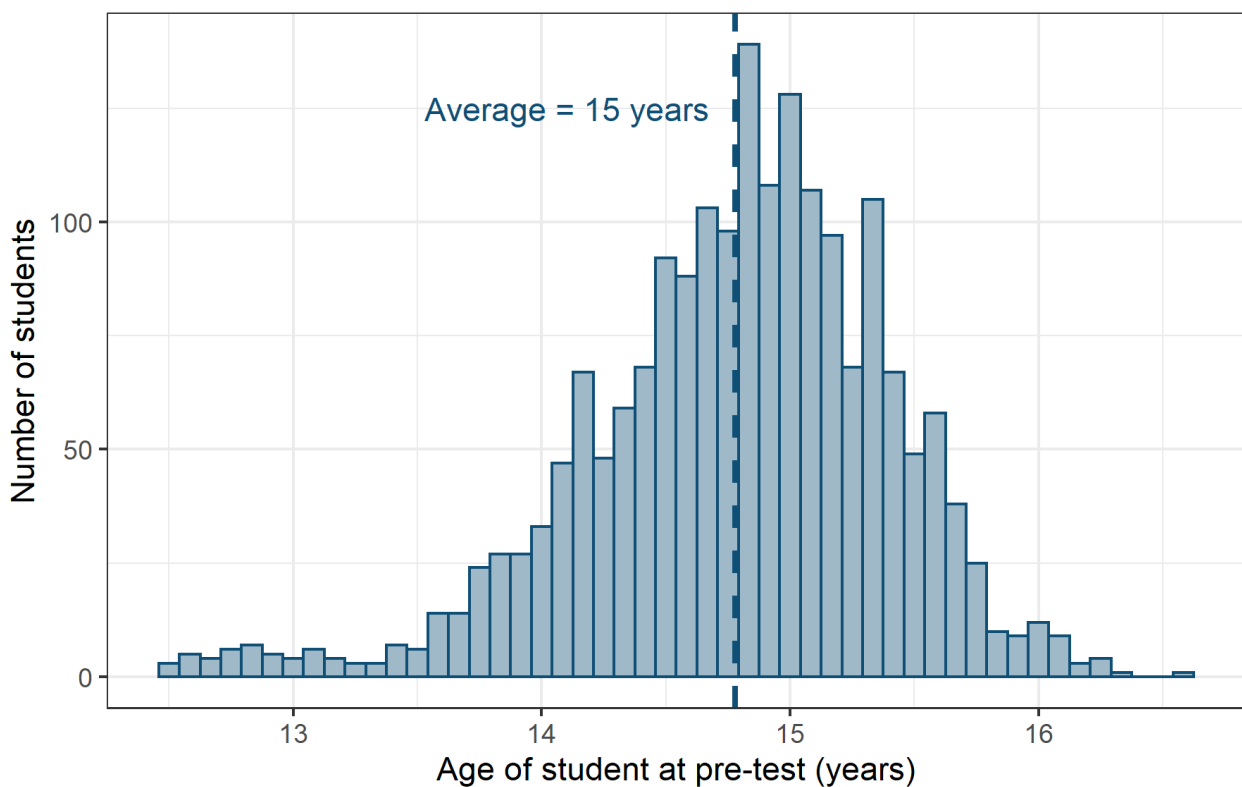
Source: TALIS Video Study student data file

In England, the age of students participating in the TALIS Video Study was both younger (on average) and more variable than the age of students who participated in PISA 2018, due to the focus on the teaching of a particular mathematical topic rather than a particular age of students. PISA included students who were aged between 15 years, 3 months and 16 years, 2 months at the time of assessment.

The TALIS Video Study student questionnaire also collected information about students' gender and whether students were first- or second-generation immigrants. Table 2.5 shows how the characteristics of the participating students in the TALIS Video Study differed from those in PISA 2018.

Table 2.5: Student characteristics

Characteristic	TALIS Video Study	PISA 2018
Proportion of female students	54%	51%
Average age	15	16
Proportion of first- or second-generation immigrant students	16%	22%
Proportion of students who speak English at home	92%	88%

Source: TALIS Video Study student data file and PISA 2018 dataset

Socio-economic status in the TALIS Video Study was based on students' responses to questions about their parents'/carers' backgrounds and education, as well as possessions in their homes. This differs slightly from the questions used in PISA 2018 which also included questions about parents' and carers' occupations.

Parents'/carers' education is based on the maximum International Standard Classification of Education (ISCED) level for each parent/carer, converted to years of education based on the education qualification system in England. In the TALIS Video Study almost half of the students had at least one parent/carer who had attended university, as can be seen in Table 2.6.

Table 2.6: Parents'/Carers' education level

Years of parental/carer education	Level of parental/carer education	Number of students	Proportion of students
<10	Before GCSE	26	1%
11	GCSE	272	13%
12	A-Level	267	13%
14	Non-University Tertiary Education	295	15%
16	University	952	47%
Data not available		212	10%

Source: TALIS Video Study student data file

What were students' experiences of mathematics teaching?

The TALIS Video Study included a focus on non-cognitive outcomes in addition to student attainment. These non-cognitive outcomes were students' self-reported personal interest in mathematics and confidence, or self-efficacy, in mathematics. These non-cognitive outcomes were measured through questions about their experiences with their previous teacher, with their current teacher, and during the topic of quadratic equations that was the focus of the study. The questions focusing on students' interests or self-efficacy with their previous teacher were used as pre-measures of students' non-cognitive outcomes (see Chapter 10 and the England Technical Report)[20].

Table 2.7 shows that whilst many students were interested in mathematics with their previous teacher (56%), fewer thought what they were talking about in their mathematics lessons was interesting, and less than one in five were curious about their next mathematics lesson. The average personal interest score with their previous mathematics teacher was 2.31 on a scale of 1 to 4, where 4 represents a high level of personal interest.

Table 2.7: Students' reported interest in mathematics with their previous teacher

Statement	Proportion of students
I was interested in mathematics	56%
I often thought that what we were talking about in my mathematics class was interesting	45%
After mathematics class I was often already curious about the next mathematics class	18%
I wanted to deal more intensively with some topics discussed in my mathematics class	39%

Further information on responses to these items can be found in the England Technical Report
Proportions represent the students who agreed or strongly agreed with each statement
Source: TALIS Video Study student data file

More students reported being interested in mathematics with their current mathematics teacher than with their previous mathematics teacher, but they found the focal topic of quadratic equations less interesting than mathematics in general, and only a minority were curious about their next mathematics lesson, as can be seen in Table 2.8. The average personal interest score for the sequence of lessons that included quadratic equations was 2.

---

[20] McCann, Riggall, Ingram, and Lindorff (2020)

Table 2.8: Students' reported interest in mathematics with their current teacher

Statement	Proportion of students (in general)	Proportion of students (during topic)
I am interested in mathematics	76%	59%
I often think that what we are talking about in my mathematics class is interesting	68%	54%
After mathematics class I am often already curious about the next mathematics class	33%	24%
I would like to deal more intensively with some topics discussed in my mathematics class	58%	36%

Further information on responses to these items can be found in the England Technical Report

Proportions represent the students who agreed or strongly agreed with each statement

Source: TALIS Video Study student data file

Many students reported not being confident in mathematics when they were with their previous mathematics teacher, with an average general self-efficacy score of 2.2 on a scale of 1 to 4. Table 2.9 shows that whilst many students had not felt confident in mathematics with their previous teacher, more than half expected to do well in mathematics.

Table 2.9: Students' reported confidence in mathematics with their previous teacher

Statement	Proportion of Students
I believed I would receive an excellent mark in mathematics	28%
I was confident I could understand the most difficult material on mathematics	32%
I was confident I could do an excellent job on the mathematics assignments and tests	32%
I expected to do well in mathematics	53%
I was confident I could master the mathematics skills being taught	40%

Further information on responses to these items can be found in the England Technical Report
Proportions represent the students who said these statements were very true or extremely true of them
Source: TALIS Video Study student data file

More students reported being more confident in mathematics with their current mathematics teacher than with their previous teacher, but the opposite was found for the focal topic of quadratic equations, as shown in Table 2.10.

Table 2.10: Students' reported confidence in mathematics with their current teacher

Statement	Proportion of students (in general)	Proportion of students (during topic)
I believe I will receive an excellent mark in mathematics	46%	22%
I am confident I can understand the most difficult material on mathematics	51%	24%
I am confident I can do an excellent job on the mathematics assignments and tests	44%	28%
I expect to do well in mathematics	63%	42%
I am confident I can master the mathematics skills being taught	53%	39%

Further information on responses to these items can be found in the England Technical Report
Proportions represent the students who said these statements were very true or extremely true of them
Source: TALIS Video Study student data file

The average general self-efficacy in mathematics score for the sequence of lessons that included quadratic equations was 2.2.

Students also completed two attainment tests. The first test was completed before the teaching of the unit of quadratic equations began, and the second test was completed at the end of the teaching unit. The pre-test focused on students' general mathematics knowledge and pre-existing knowledge of quadratic equations and included 30 questions. The post-test measured students' knowledge of quadratic equations following instruction and included 24 questions.

The average mark on the pre-test was 22 out of 30, and the average mark on the post-test was 11 out of 24. These marks were then used to calculate a score that was standardised across the countries taking part. Across all countries that took part in the TALIS Video Study the average score was 200 points with a standard deviation of 25 points.

For England, the average score on the pre-test was 202, with a standard deviation of 18 points. Figure 2.5 shows the distribution of scores for England on the pre-test. The median score for England on the pre-test was 199.

Figure 2.5: Student attainment on the pre-test
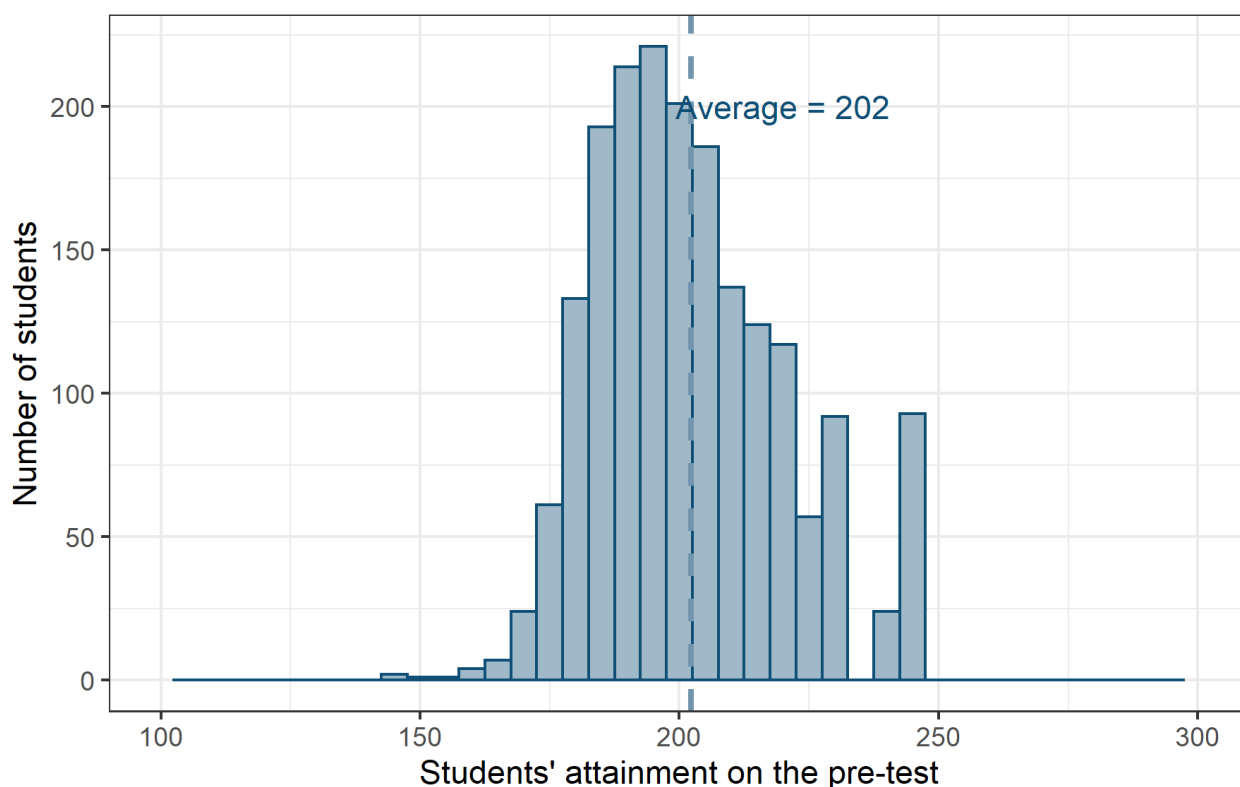
Source:

The average score on the post-test for England was 195 with a standard deviation of 14 points. This is lower than the average of 200 across the TALIS Video Study countries at post-test. Figure 2.6 shows the distribution of scores for England on the post-test. The median score for England on the post-test was 193. This difference in scores on the post-test could be influenced by the difference in *opportunities to learn* (see Chapters 6 and 7) the mathematics included in the post-test experienced by students in England compared to students in the other participating countries. As outlined in the introduction, the focal topic of quadratic equations involves many subtopics that are not in the National Curriculum for the majority of students in England. Solving quadratic equations by factorising and identifying roots of quadratic functions graphically, as well as moving between numerical, algebraic, graphical, and diagrammatic representations of quadratic functions, are included in the National Curriculum for all students. Other methods for solving quadratic equations, such as completing the square or by using the quadratic formula, as well as solving quadratic equations by factorising where the coefficient of the quadratic term is different from 1, are only included in the Higher Tier examinations and are often only taught to more highly attaining students[21]. In 2018 and 2019 over half of the students taking GCSE Mathematics were entered for the foundation tier[22].

---

[21] Department for Education (2014)
[22] Ofqual (2019)

Figure 2.6: Student attainment on the post-test
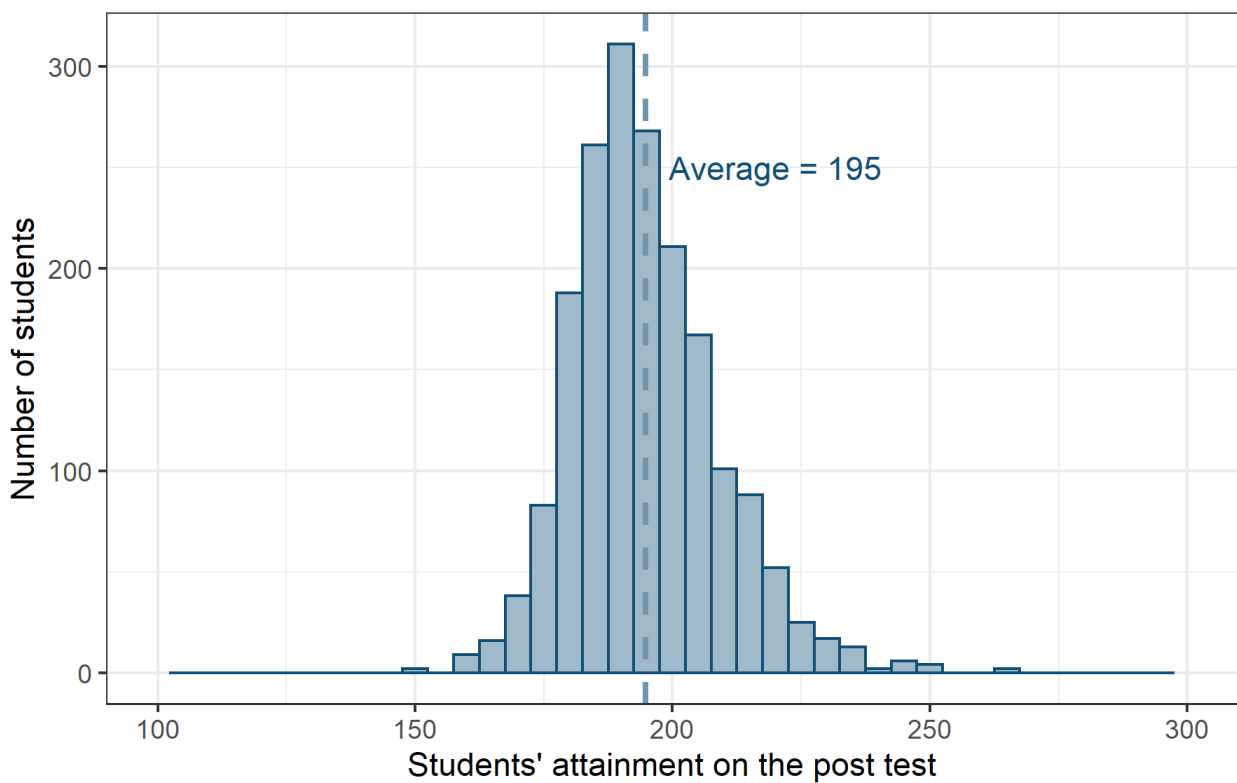
The characteristics of schools and classes

What were the sizes of participating classes?

The average class size in the TALIS Video Study was 27.4 students. The smallest class size was 15 students and the largest was 34 students. In TALIS 2018, the average comparable class size in England was 24.5 students, which was slightly above the OECD average of 23.8 students. Table 2.11 illustrates how there was variation in class sizes for different year groups.

Table 2.11: Class sizes across different year groups

Year group	Number of classes	Average class size	Smallest class size	Largest class size
8	2	31.5	31	32
9	14	30.0	21	33
10	62	26.7	15	34
11	7	27.6	25	34

Source: TALIS Video Study student and teacher data files

What were the characteristics of the participating classes?

In England, it is a common practice in secondary schools to group students by prior attainment (often termed setting), particularly in Key Stage 4, and for lower attaining groups to have smaller class sizes to support students' learning. Figure 2.7 illustrates the variation in class sizes for different levels of attainment. In Year 10, the classes with the lowest attainment on the pre-test had an average class size of 22.6 students and the classes with the highest attainment on the pre-test had an average class size of 29.2 students. This difference may be a result of the different class sizes that arise from grouping students by prior attainment.

Figure 2.7: Relationship between class size and attainment on the pre-test
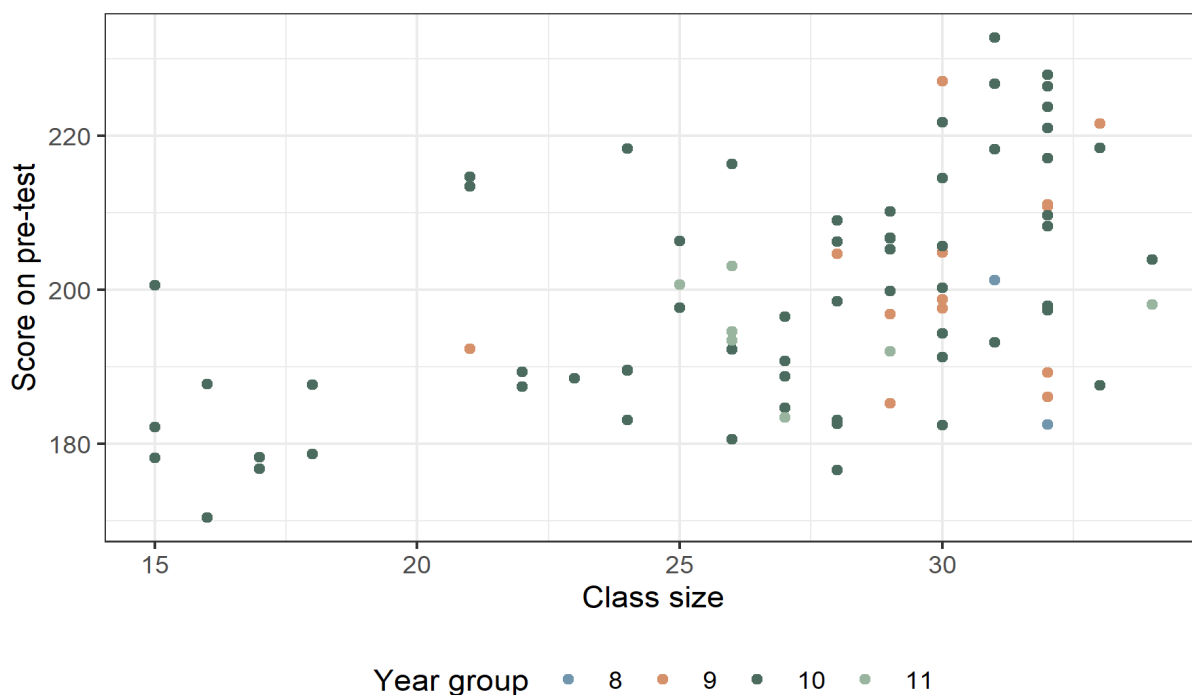
Source: TALIS Video Study student data file

The class with the lowest average attainment and the class with the highest average attainment were in Year 10. Several students in both Years 9 and 10 achieved the highest possible score on the pre-test. The topic of quadratic equations is taught at different times of the year and in different year groups. Whilst the topic of quadratic equations is in both the foundation and the higher tier GCSE examination specifications, the topic is often taught towards the end of the foundation tier course, whereas it is taught far earlier for those taking the higher tier examination. Furthermore, only some aspects of the topic of quadratic equations included in this study are included in the National Curriculum for students taking the foundation tier examinations[23].

In the England sample for TALIS Video Study only one class in Year 11 had an average attainment above the median suggesting that the sample does not include the full range of prior attainment in mathematics within that year group.

All the schools that participated were state-funded schools, as no independent schools[24] consented to participating. Of the participating state-funded schools, 23% were faith schools and 12% were selective schools.

Nearly three-quarters (74%) of the schools were in urban areas. Table 2.12 shows the type of schools that took part.

Table 2.12: Participating school types:

School type	Proportion of schools
Academy (converter)	76%
Academy (sponsor-led)	10%
Community	9%
Foundation	3%
Voluntary aided	4%

Note: One school is categorised as both a Foundation school and an Academy Sponsored school

Source: Secondary School Performance Tables (2018)

The schools were recruited from across England, as can be seen in Table 2.13.

---

[23] Department for Education (2014)
[24] Only five were included in the sample provided

Table 2.13: Regions participating schools came from

School region	Proportion of schools
East Midlands	12%
East of England	4%
London	5%
North East	8%
North West	24%
South East	12%
South West	15%
West Midlands	14%
Yorkshire and the Humber	4%

Source: TALIS Video Study school data file

Summary

The TALIS Video Study in England included 85 teachers from 78 schools, with a total of 2,024 students participating. The mathematics teachers in the TALIS Video Study had an average of 10 years of teaching experience, and 58% were female. The majority of students in the TALIS Video Study in England were in Year 10. The average class size in the TALIS Video Study in England was 27.4 students. Schools included were all state funded, and 74% of participating schools were in urban areas. Teachers had mostly positive feelings about their job. However, only 35% of teachers agreed or strongly agreed that the teaching profession is valued in society. Teachers largely felt responsible for their students' learning and progress, and in particular for the effectiveness of their own teaching. When it came to professional development, 86% of teachers reported that they were aware of and understood the 2016 Standard for Teachers' Professional Development. Furthermore, 91% of teachers felt that the professional development they had completed had had a positive impact on their teaching.

3. How were mathematics classrooms managed?

Key findings

- Classes were generally well managed: teachers generally had organised and efficient routines, they monitored what was happening in the entire classroom and they dealt with disruptions quickly and effectively.

- Most teachers said that their students were aware of what was allowed and what was not allowed and that they knew why certain rules were important.

- Most teachers reported that they were immediately aware of students doing something else, and that they were aware of what was happening in the classroom, even if they were busy with an individual student.

- Only a small minority (8%) of teachers agreed or strongly agreed that they lost a lot of time due to students interrupting lessons during the teaching of quadratic equations.

- Most students also reported that their mathematics teacher managed to stop disruptions quickly.

- On average, lower attaining classes were observed to experience more disruptions and fewer organised and efficient routines. The classes with the lowest average ratings for the classroom management domain were also some of the lowest attaining classes on the pre-test.

Introduction to the classroom management domain

In this chapter the three main components of teaching practices within the domain of *classroom management* are considered: classroom *routines*; teacher *monitoring*; and *disruptions*. Each were measured on a scale of 1 to 4, where a higher rating represented higher quality and/or higher presence of particular behaviours. Videos were also given a holistic rating across the three main components on the same scale. In addition, analysis of the videos included indicators of the proportion of *time on task* that students spent working, and the activity structures seen, such as whole class, small group, pair, or individual work.

Routines considers the organisation and efficiency of the routines for common managerial tasks within the classroom. These tasks include handing out books, taking the register, or moving students from pair work to group work. A rating of 4 means that all routines within the lesson were organised and did not waste any time.

Monitoring involves the teacher monitoring what is happening in the entire classroom. This can include the teacher maintaining physical proximity to the students, scanning the entire classroom, facing the students, calling upon a range of students and noticing student progress. The focus is on the monitoring of the whole class and not just a small group of students. A rating of 4 means that the teacher frequently monitored the entire classroom and did so consistently.

Disruptions are instances when teachers', students' or external actors' behaviours draw significant attention away from the subject matter or classroom activities. This includes fire alarms, technology failures, students arriving late, and student misbehaviour. This component focuses on how quickly and effectively the teacher deals with any disruption, as well as the number of disruptions that occur, and considers the instructional time lost due to these disruptions. A rating of 4 could mean that the teacher handled any disruptions quickly and effectively so that whilst there may have been interruptions, no instructional time was lost. A rating of 4 could also indicate that there were no disruptions during the lesson.

Time on task was also measured within this domain. This measures the extent to which there is loss of lesson time due to activities or situations that are not directly focused on mathematical learning, such as classroom routines, off-topic discussions, or transitions between tasks or stages of the lesson. Mathematical learning includes the full range of activities in which the ideal student should be engaged, including listening to the teacher explaining, doing group work, or working individually on a problem. This indicator was measured on a scale of 1 to 4 where a rating of 4 represents very little time being lost during the lesson.

In this domain the different *activity structures* and the frequency of these activity structures were also measured. This indicator measures the extent to which whole group work, small group work (three or more students), pair work, or individual work are used at different stages of the lesson.

How did teachers monitor and manage their classes?

In England, the lessons in the TALIS Video Study were generally rated as well-managed with an average rating of 3.74. Teachers typically had organised and efficient routines, they monitored what was happening in the entire classroom and they dealt with disruptions quickly and effectively.

There was also little variation in classroom management as illustrated in Figure 3.1. Overall, the lowest rating for classroom management was 3.11 and the highest was 4.00, with some teachers achieving the maximum rating across all three components. There was more variability within *monitoring* but for all three components more than three-quarters of teachers had an average rating above 3.5.

Figure 3.1: Distribution of teacher average ratings for the classroom management domain
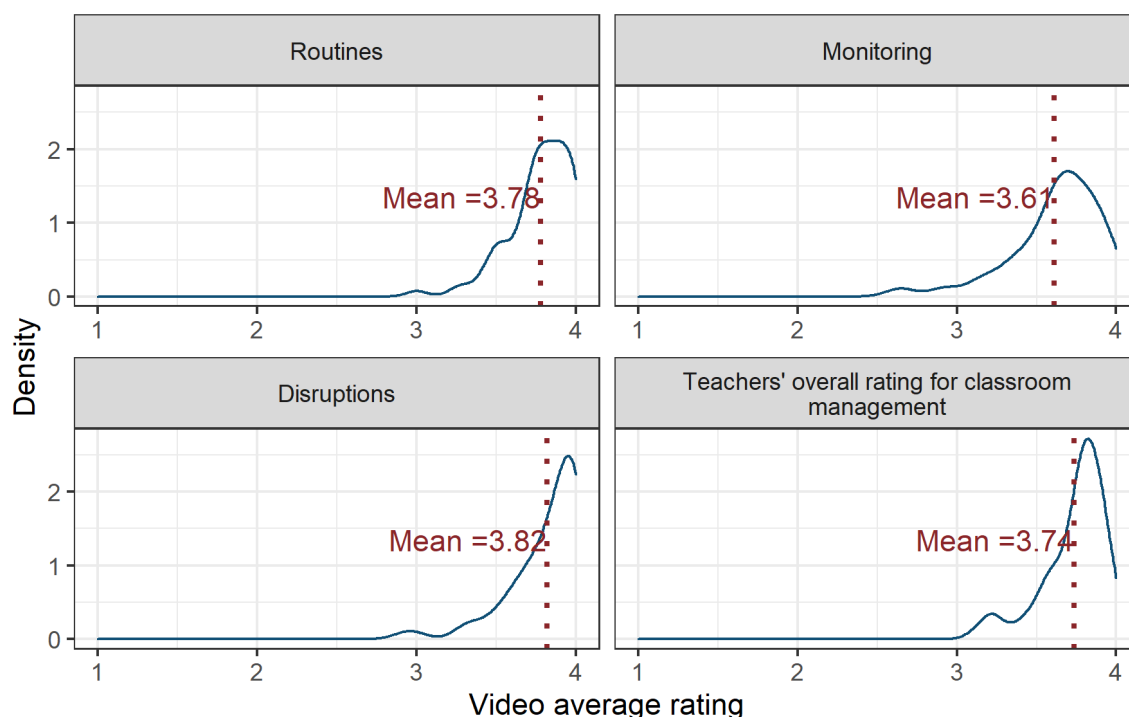
Most lessons lost very little time to activities, tasks, or dialogue that was not focused on mathematical learning. The average rating for *time on task* was 3.86 with all teachers having an average rating over 3.

Teachers frequently worked with the whole class and students were also given several opportunities to work individually. Teachers rarely made use of small group work or pair work. Each 8-minute segment of video was rated as to whether and for how much of the segment each activity structure was used, and teachers could use more than one type of activity structure during a segment. Only 21% of teachers used small group work at all in their lessons, whereas 60% used pair work at least once.

What were teachers' and students' perceptions of classroom management?

At the end of the teaching unit, both teachers and students were asked a series of questions about classroom management in general and during the teaching of the unit. The proportions of teachers or students agreeing or strongly agreeing with each statement are summarised in Table 3.1.

Table 3.1: Similarities and differences in teachers' and students' perceptions of classroom management.

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
In this/our teacher's class, students/we were aware of what was allowed and what was not allowed	98% (98%)	94% (92%)
I managed/Our teacher manages to stop disruptions quickly	98% (96%)	85% (84%)
In this/our teacher's class, students knew/we know why certain rules were important	96% (96%)	92% (90%)
I reacted/Our teacher reacts to disruptions in such a way that the students stopped disturbing learning	93% (94%)	75% (72%)
I was immediately aware of students doing something else	93% (88%)	68% (67%)
I was/Our teacher is aware of what was happening in the classroom, even if I was/he or she is busy with an individual student	93% (96%)	77% (75%)
When the lessons began, I/our mathematics teacher had to wait quite a long time for these students to quieten down	21% (13%)	31% (26%)
I lost/we lose quite a lot of time because of students interrupting the lessons	21% (8%)	32% (24%)
In this class/our teacher's class, transitions from one phase of the lesson to the other (e.g., from class discussions to individual work) took/take a lot of time	15% (8%)	25% (24%)
There was much disruptive noise in this classroom	13% (10%)	29% (23%)

Further information on responses to these items can be found in the England Technical Report
Proportions represent the teachers or students who agreed or strongly agreed with each statement
Source: TALIS Video Study teacher and student data files

Most teachers said that their students were aware of what was allowed and what was not allowed and that they knew why certain rules were important. The students also mostly agreed or strongly agreed with these statements. The majority of both teachers and students reported that classrooms were generally well managed, though there were differences between teachers' and students' perceptions of classroom management.

Teachers were more likely to agree or strongly agree that disruptions were handled quickly and in a way that did not disturb learning than their students. Students were also more likely to agree or strongly agree that there were more disruptions and time lost due to these disruptions, classroom noise or transitions within lessons than their teachers.

The proportion of teachers in the TALIS Video Study who agreed or strongly agreed that they lost a lot of time due to students interrupting lessons was 21%, but during the topic of quadratic equations this dropped to 8%. These are both lower than the proportion of teachers in TALIS 2018 where 28% of lower-secondary teachers in England agreed or strongly agreed. The proportion of teachers who agreed or strongly agreed that there was a lot of disruptive noise in the classroom was 13% in general and 10% for the topic of quadratic equations, which were also lower than the proportion of teachers in TALIS 2018 where 23% of lower-secondary teachers in England agreed or strongly agreed.

Did teachers with different backgrounds manage classrooms differently?

There was very little variation in the average ratings in the *classroom management* domain as the vast majority of teachers were rated highly on all three components – *routines*, *monitoring*, and *disruptions*. There were no significant relationships between the average video rating on the overall domain or any of the components within this domain and teachers' self-efficacy. There were also no significant differences between teachers with different qualifications (teacher training and Masters'-level education), or between male and female teachers in the overall *classroom management* domain, or on the average rating for *routines*, *monitoring*, and *disruptions*. In particular, there were no significant differences between teachers with different levels of experience as mathematics teachers, as shown in Table 3.2. This is in contrast to the results from TALIS 2018 which found a significant difference in perceived behaviour management in classrooms between teachers who had been working for five years or fewer as a teacher compared to teachers who had worked for more than 20 years. This could potentially be explained by the smaller sample of teachers within the TALIS Video Study compared to TALIS 2018, and the sample for TALIS Video Study focusing exclusively on mathematics teachers.

Table 3.2: Relationship between classroom management and teachers' experience

Teaching experience	Number of teachers	Routines	Monitoring	Disruptions	Overall
5 years or less	29	3.74	3.63	3.78	3.72
6 to 10 years	21	3.80	3.66	3.89	3.79
11 to 15 years	15	3.76	3.46	3.75	3.66
15 to 20 years	8	3.84	3.62	3.82	3.76
More than 20 years	10	3.84	3.63	3.85	3.77

Source: TALIS Video Study teacher data file

Did different classes have different experiences of classroom management?

In general, most classes had similar experiences of classroom management. There were no significant associations between students' experiences of classroom management according to different class proportions of female students, students with low socio-economic status (as measured by level of parental education and number of home possessions), or proportion of students who did not speak English at home. There were also no significant differences in experiences of classroom management between classes with different levels of student personal interest in mathematics, nor those with different levels of self-efficacy in mathematics based on their responses to the pre-questionnaire.

Figure 3.2 illustrates that there were significant associations between the class average score on the pre-test and the average ratings for the *classroom management* domain overall, as well as the *routines* and *disruptions* components in particular. The higher the class average score on the pre-test, the higher the average teacher rating for the domain and these components. There was no significant association between class average pre-test scores and the *monitoring* component.

Figure 3.2: Teacher average classroom management ratings versus class average pre-test scores
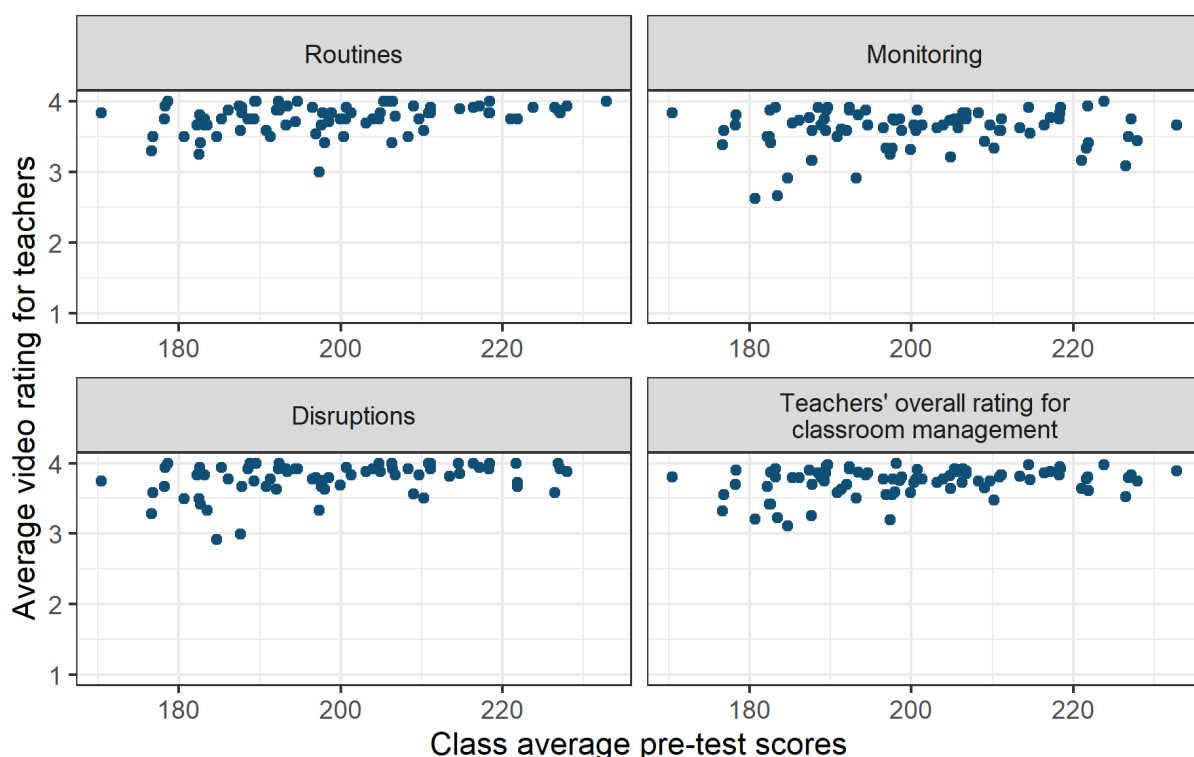
There was more variability in students' experiences of classroom management if they were lower attaining when compared to higher-attaining students, and this difference is statistically significant for the three components (*routines*, *monitoring*, and *disruptions*), as well as on the overall domain average rating. Some classes with lower average scores on the pre-test experienced classrooms with some of the lowest average ratings on each of the *classroom management* components but other classes with similar average scores on the pre-test experienced classrooms with some of the highest average ratings on these components.

Classes were grouped according to their prior attainment, with the top quarter of classes being identified as higher attaining and the bottom quarter of classes being identified as lower attaining. Table 3.3 shows the average ratings and the standard deviations of each of the higher- and lower-attaining quartiles. On each of the components, the lower-attaining classes experienced more variation in classroom management than the higher-attaining classes.

Table 3.3: Variation in classroom management experiences of classes of different levels of prior attainmnet

Component	Higher-attaining classes' average rating (s.d.)	Lower-attaining classes' average rating (s.d.)
Routines	3.88 (0.11)	3.69 (0.21)
Monitoring	3.62 (0.25)	3.53 (0.40)
Disruptions	3.90 (0.15)	3.67 (0.32)
Overall domain rating	3.80 (0.13)	3.63 (0.27)

Source: TALIS Video Study teacher data file

Summary

Classrooms in the TALIS Video Study in England were well managed, with an average rating of 3.74. Lesson videos were rated on *routines*, *monitoring*, and *disruptions*, and for all three components more than three-quarters of teachers were rated above 3.5 on average on the 4-point scales. Teachers generally had organised and efficient routines, monitored what was happening in the entire classroom, and dealt with disruptions quickly and effectively. Teachers and students both perceived mathematics classrooms to be managed well. There were no significant differences in classroom management average ratings between teachers with different characteristics. Higher-attaining classes tended to experience more organised and efficient routines and fewer disruptions to their lessons. Whilst many of the lower-attaining classes experienced high average ratings within the *classroom management* domain, a few of the lower-attaining classes experienced some of the lowest average ratings within this domain. There was more variability in students' experiences of classroom management if they were lower attaining when compared to higher-attaining students, and this difference was statistically significant.

4. What did social and emotional support look like?

Key findings

- Teachers and classrooms were generally socially and emotionally supportive.

- Teachers and students frequently and consistently demonstrated respect for one another through manners, language, and tone of voice.

- Students in all classes sought guidance and shared their work publicly.

- Teachers consistently provided encouragement to their students and shared moments of warmth in their lessons.

- Teachers were nearly unanimous when it came to their attitudes about creating positive relationships with their students, agreeing that they got along well with their students, showed an interest in their students' wellbeing, and really listened to their students.

- Most students reported that their teacher gave extra help when needed and continued teaching until students understood.

- Most students had positive perceptions of their relationship with their teacher, feeling heard, cared about, and treated fairly.

- While 84% of students had a sense of belonging in their mathematics class, 13% of students felt awkward and out of place.

- Classes with higher average self-efficacy scores, personal interest scores, or pre-test scores tended to have higher average ratings on the respect component.

Introduction to the social and emotional support domain

In this chapter, the three main components of teaching practices within the domain of *social and emotional support* are considered: *respect*, *encouragement and warmth*, and *risk-taking*. Each was measured on a scale of 1 to 4, where a higher rating represented higher quality or higher prevalence of particular behaviours. Videos were also given a holistic rating across the three main components. In addition, analysis of the videos included indicators of student *persistence* and the *public sharing* of mathematics.

Respect measures the extent to which teachers and students demonstrate respect for one another by using respectful language, listening to one another, using appropriate names, using a respectful tone of voice and using other traditional markers of manners. This component also takes into account disrespectful interactions between the teacher

and students, or between students. These include threats, mean or degrading comments, physical aggression such as pushing someone or slamming down materials, and comments after which a student or the teacher demonstrates shame. A rating of 4 means that the teacher and students frequently and consistently demonstrated respect for one another, and that there were no negative interactions between any student and the teacher or between students.

Encouragement and warmth measures the extent to which the teacher and/or the students provide encouragement to students throughout their work. This includes the teacher reassuring students when errors are made, making positive comments, and complimenting students' work. This component also includes moments of shared warmth such as smiling, laughing, joking, and playfulness. A rating of 4 on this component means that the teacher and/or the students frequently provided encouragement to students throughout their work and that there were frequent moments of shared warmth.

The final component considered in this chapter is *risk-taking*. This measures students seeking guidance or voluntarily taking risk by publicly sharing their private work. A rating of 4 on this component means that students either frequently sought guidance and/or they frequently voluntarily shared their private work publicly.

The *persistence* indicator describes the extent to which students persist through errors or mathematical struggles with the teacher's support. This indicator was also measured on a scale of 1 to 4. To achieve a rating of 4, students demonstrated that they were aware they had made an error or were engaged in a mathematical struggle and they persisted through this at significant length and/or in significant depth with the teacher's support.

Requests for public sharing measures the extent to which the teacher requests that students share their private mathematical thinking publicly. This indicator was measured on a scale of 1 to 3, where a rating of 3 represents that the teacher requested students to share their private mathematical thinking publicly and this shared work has more than limited detail.

Were mathematics classrooms socially and emotionally supportive environments?

In the TALIS Video Study classes, classrooms were socially and emotionally supportive environments with an overall average rating of 3.14. In particular, teachers and students demonstrated respect for one another in the majority of classrooms, with 88% of teachers being rated over 3 on average. The average rating was lower for the component of *encouragement and warmth*, which also had somewhat more variation (measured by the standard deviation) than the other components; however, the average was still above the midpoint of the scale. The teacher average ratings for *risk-taking* have been excluded from the overall rating for the domain by the ISC, as a result of the variation in practices between countries, so this component is reported separately.

Figure 4.1 illustrates the average ratings on each of the components within the *social and emotional support* domain. The vast majority (93%) of teachers consistently provided more than occasional encouragement to their students throughout their work and shared moments of warmth in their lessons. Students in all classrooms sought guidance and shared their private work publicly.

Figure 4.1: Distribution of teacher average scores for the social-emotional domain
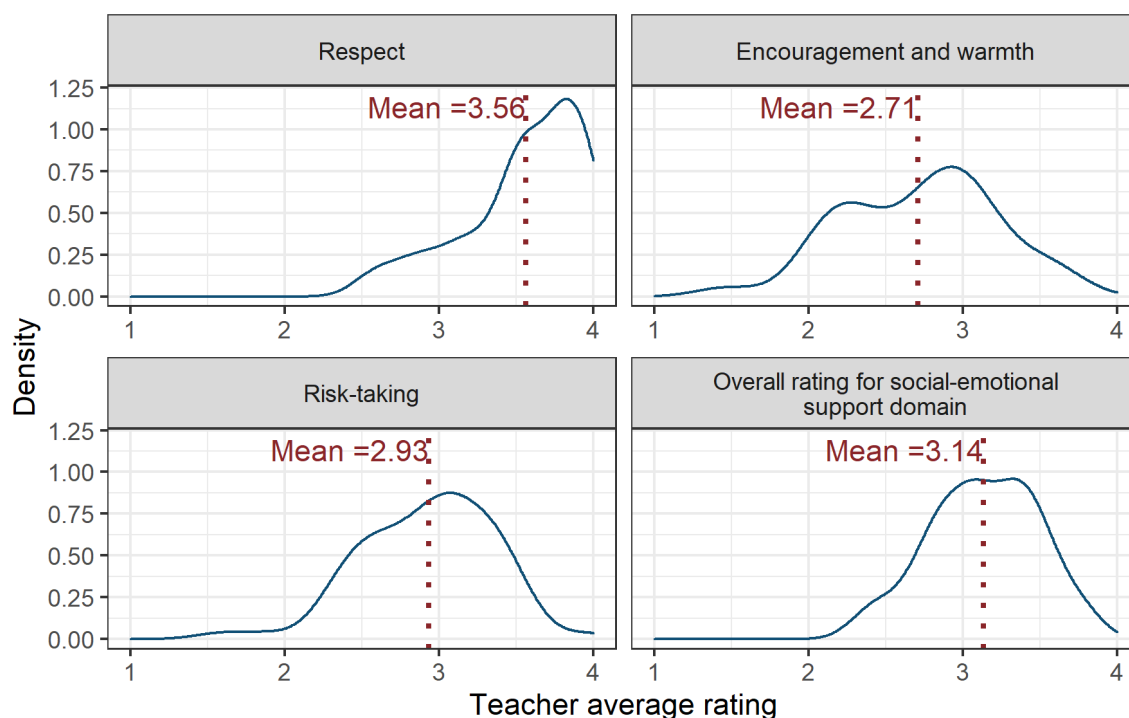
Indicators were measured for every 8-minute segment of a lesson. There was evidence of students persisting through errors or mathematical struggles with the teacher's support in 63% of lesson segments, though in the majority of these cases the teachers only addressed the mathematical errors or struggles briefly or superficially, or ignored them. Almost every teacher supported students in persisting through their mathematical errors or mathematical struggle for a moderate length of time at least once.

In all lessons, teachers requested some students share their private mathematical thinking publicly at some point and almost all teachers (94%) requested that students share their private mathematical thinking publicly in more than limited detail, that is, students shared something that revealed their thinking processes or rationales, for example by explaining or describing procedures taken or their reasoning behind procedures taken.

What were teachers' and students' perceptions of social and emotional support in lessons?

In the questionnaires at the end of the teaching unit, teachers and students were asked a series of questions about the social and emotional support provided during the teaching of that unit.

Table 4.1: Teachers' and students' perceptions of social and emotional support

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
I/Our mathematics teacher gave extra help when these students/we needed it	100% (100%)	92% (89%)
I/Our mathematics teacher helped these students/us with their/our learning	100% (100%)	96% (94%)
I aim to make these students feel confident in their ability to do well in the course/ Our mathematics teacher made me feel confident in my ability to do well in the course	100% (100%)	82% (80%)
I aimed to make these students feel confident in their ability to learn the material/ Our mathematics teacher made me feel confident in my ability to learn the material	100% (100%)	83% (80%)
I/Our mathematics teacher appreciated it when different solutions came up for discussion	100% (98%)	90% (86%)
I listened to these students' views on how to do things/Our mathematics teacher listened to my view on how to do things	98% (95%)	81% (74%)
I aimed to show these students that I understood them/I felt that our mathematics teacher understood me	95% (96%)	76% (75%)
I/Our mathematics teacher continued teaching until these students/we understood	89% (88%)	87% (85%)
I/Our mathematics teacher encouraged these students/me to find the best way to proceed by myself	86% (75%)	78% (73%)

Further information on responses to these items can be found in the England Technical Report
Proportions are of teachers or students who agreed or strongly agreed
Source: TALIS Video Study teacher and student data files

Teachers and students agreed that the mathematics classroom was a supportive environment. As can be seen in Table 4.1 teachers were more likely to agree with the statements than their students. Students were also more likely to agree when describing their teacher in general rather than during the teaching of quadratic equations.

Teachers and students also reported good relationships with each other. Almost all teachers reported that they got along well with their students and listened to what their students had to say. Similarly, almost all teachers also stated that they treated their students fairly and that they made their students feel that they really cared about them. All teachers reported that they showed interest in their students' well-being, as shown in Table 4.2. Students' perceptions of their relationships with their teacher were not as positive as their teacher's. The largest difference was where almost all teachers agreed or strongly agreed that they made their students feel that they really cared about them whereas as this statement was only agreed or strongly agreed with by around two-thirds of the students.

Table 4.2: Teachers' and students' perceptions of their relationships with each other

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
I got along well with these students/my mathematics teacher	100% (98%)	90% (89%)
I showed interest in these students' well-being/ My mathematics teacher was interested in my well-being	100% (100%)	81% (77%)
I/My mathematics teacher treated these students/me fairly	100% (99%)	93% (91%)
I/My mathematics teacher really listened to what these students/I had to say	96% (98%)	81% (79%)
I made students feel I really cared about them/ My mathematics teacher made me feel he/she really cared about me	96% (99%)	70% (68%)

Further information on responses to these items can be found in the England Technical Report
Proportions are of teachers or students who agreed or strongly agreed
Source: TALIS Video Study teacher and student data files

Students were also asked a series of statements about their sense of belonging in their mathematics class (see Table 4.3). The vast majority of students reported that they felt

like they belonged in their mathematics class. However, 13%(12%) of students reported feeling awkward and out of place and a similar proportion reported that they felt like an outsider in their mathematics class.

Table 4.3: Students' sense of belonging in their mathematics lessons

Statement	Proportion of students in general (during topic)
I felt like I belonged in my mathematics class	84% (82%)
I felt awkward and out of place in my mathematics class	13% (12%)
I felt like an outsider (or left out of things) in my mathematics class	12% (12%)
I felt lonely in my mathematics class	8% (9%)

Further information on responses to these items can be found in the England Technical Report
Proportions are of teachers or students who agreed or strongly agreed
Source: TALIS Video Study student data file

Did teachers with different backgrounds support students differently?

No statistically significant differences within the domain of *social and emotional support* were identified between teachers of different genders, with different qualifications, with different levels of self-efficacy, or with different levels of teaching experience.

There was a significant difference for teachers who had mathematics courses equivalent to degree level within their own mathematics education or training in the *encouragement and warmth* component, but this difference was not significant once other teacher background characteristics were taken into account. There was a significant difference in the average overall domain rating between teachers who had a degree or an equivalent level of training in their mathematics education and those who did not, which remained significant after controlling for other teacher characteristics. Teachers who had taken courses equivalent to degree level mathematics had significantly higher average ratings for the overall *social and emotional support* domain than teachers who had not.

Did different classes have different experiences of social and emotional support?

There were no significant differences within the domain of *social and emotional support* between classes with different proportions of male and female students, or for classes

with different proportions of students who did not speak English at home. There were also no significant differences within this domain dependent upon the socio-economic status of students as measured by their home possessions or by the level of their parents' or carers' education.

Further, there were no differences for the overall average rating for the *social and emotional support* domain between classes with different average levels of personal interest in mathematics, pre-test scores, or self-efficacy in mathematics as reported on the pre-questionnaire. This was also true for the *risk-taking* component.

Figure 4.2: Average teacher ratings for the social-emotional domain versus class average personal interest in mathematics with previous teacher scores
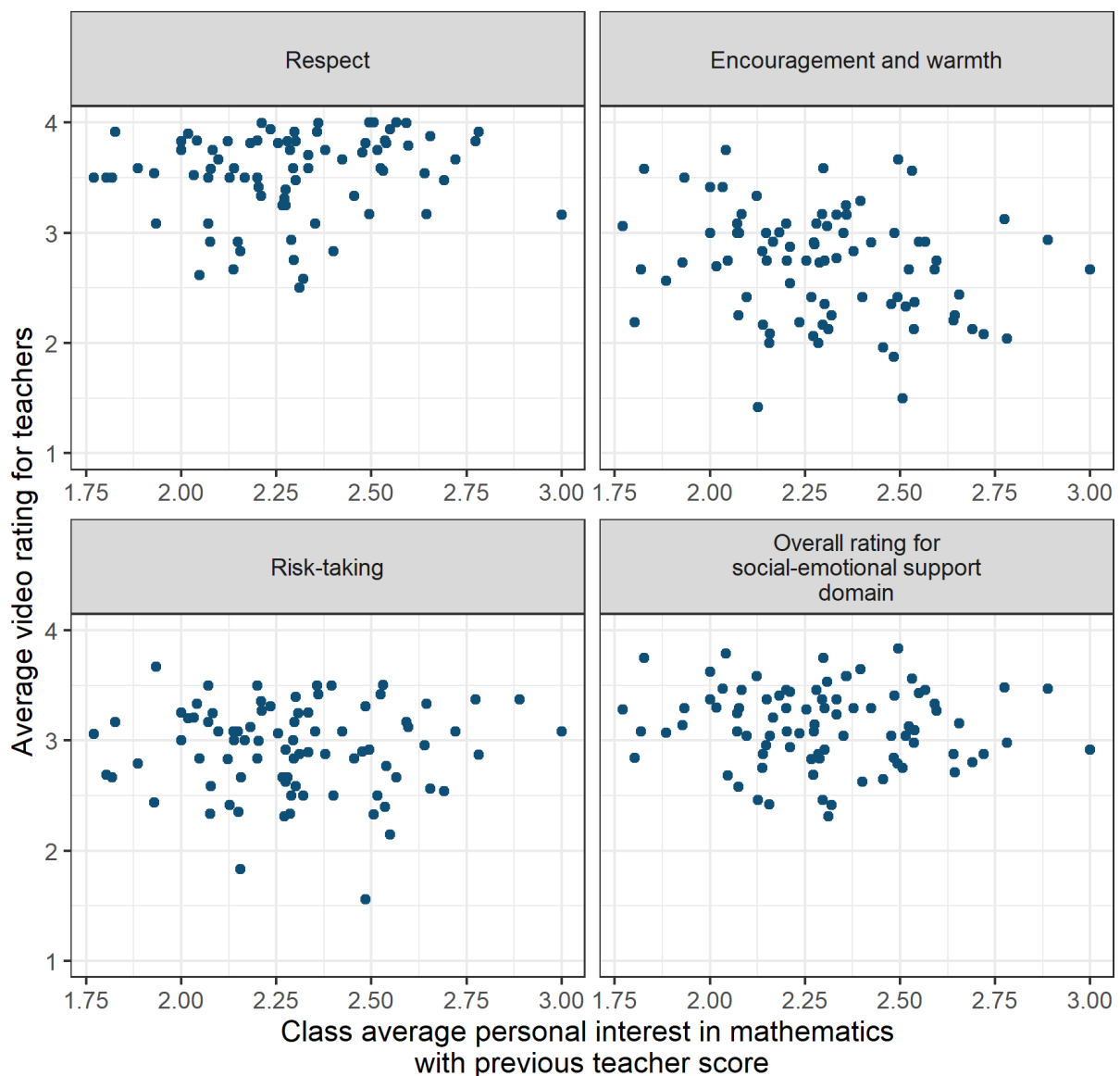
Source: TALIS Video Study teacher data file

There were significant associations between class average self-efficacy score, class average score for personal interest in mathematics, and class average score on the pre-test and the average ratings on the *respect* component. Classes with higher average general self-efficacy scores or personal interest in mathematics scores with their previous teacher, or higher average pre-test scores, tended to have higher average ratings for *respect*.

For the *encouragement and warmth* component classes with higher average scores for personal interest in mathematics with their previous teacher had significantly lower average ratings for this component compared to classes with a lower average personal interest in mathematics. However, this relationship was no longer significant after controlling for other class characteristics. Classes with higher proportions of students with a first- or second-generation immigrant background also had lower average ratings for the *encouragement and warmth* component compared to classes with a lower proportion. This significant difference remained when other class characteristics were controlled for.

Summary

Teachers and classrooms in the TALIS Video Study in England were socially and emotionally supportive, with an average rating of 3.14. Teachers reported providing very high levels of *social and emotional support* in their classrooms. A high proportion of students felt supported in their learning by their mathematics teacher. The majority of students also had positive perceptions of their relationship with their teacher, feeling heard, cared about, and treated fairly. There were no significant differences in *social and emotional support* ratings between teachers with different characteristics except that teachers who had taken courses equivalent to degree level mathematics offered significantly more *social and emotional support* than teachers who had not. There were no significant differences on the overall *social and emotional support* ratings between classes with different characteristics.

5. What was the quality of discourse in mathematics classrooms?

Key findings

- Classrooms had a mix of teacher- and student-directed discourse, with students occasionally providing detailed contributions.

- Lessons included questions that requested students to recall, define, summarise, and apply rules, and students were sometimes asked to analyse, justify, or synthesise material.

- Students were not commonly asked for explanations of their mathematical thinking in the lesson artefacts; however, approximately three-quarters of both teachers and students reported that students frequently or always had opportunities to explain their ideas.

- There were no significant differences in the quality of or opportunity for discourse provided between teachers with different qualifications (teacher training and Master's-level education), self-efficacy ratings, or years of mathematics teaching experience.

- There were significant differences between classes with different levels of prior attainment, with high-attaining classes tending to have higher average ratings on the nature of discourse, questioning, and explanations than lower-attaining classes.

- Classes with a higher average socio-economic status, based on the number of home possessions and level of parental education, were also significantly associated with higher video ratings overall on the discourse domain compared to classes with a lower average socio-economic status.

- There was a positive association between classes' average score for general self-efficacy as measured on the student pre-questionnaire and the average rating on the questioning component, with classes with higher average self-efficacy being associated with higher average questioning ratings compared to classes with lower average self-efficacy scores.

Introduction to the discourse domain

Three main components of teaching practices were considered within the domain of discourse: the *nature of discourse*, *questioning*, and *explanations*. Each were measured on a rating of 1 to 4, where a higher rating represented higher quality or higher

prevalence of particular behaviours. Videos were also given a holistic rating across the three main components. In addition, analysis of the videos included indicators of the *discussion opportunities* within the lesson. The artefacts were also rated for whether students were asked to give explanations.

The *nature of discourse* refers to the opportunities students have to participate in classroom discourse and the extent to which students' discourse is characterised by detailed contributions. Discourse is defined as any communication in the classroom by the teacher and/or students and includes written communication. Teacher-directed discourse is communication in which the teacher has control over the pattern of questions and answers. This can include the teacher introducing ideas, procedures, or processes as well as conversations in which the teacher initiates a question, to which a student responds, before the teacher evaluates this response (commonly known as the IRE or IRF sequence of interaction[25]). In teacher-directed discourse, students do not substantially shape the direction or nature of the mathematical discourse. Detailed contributions are those that include detail about the mathematics being worked on, not just short answers that give an answer or define a term, for example. A rating of 4 represents discourse that is rarely teacher-directed and where students' discourse is frequently characterised by detailed contributions. A rating of 1 indicates that the discourse is teacher-directed and that students' discourse does not include any detailed contributions.

Questioning describes to what extent teachers ask questions that request students to engage in a range of types of cognitive reasoning. These questions can be asked in written or oral forms and include any questions asked on worksheets or in textbooks used in the lesson. The rating within this component measures the extent to which the teacher asks questions that request students to analyse, synthesise, justify, or conjecture, as well as questions that request students to recall, report answers, summarise, explain, classify, apply rules processes or formulae, or define terms. A rating of 4 represents an emphasis on questions that request students to analyse, synthesise, justify, or conjecture. A rating of 1 represents discourse where the questions generally request students to recall, report an answer, provide yes or no answers, or define terms.

Explanations describes the nature and extent of teacher and student explanations. These explanations can be written or verbal and are defined as descriptions of why ideas or procedures are the way they are. Explanations are statements that clarify, rationalise, or justify. Only explanations that focus on why ideas or processes are the way they are were included, but these can be quite brief. For example, a teacher might ask "why is it called a constant?" and a student answer of "it doesn't change" would be considered an explanation that focuses on why. A rating of 4 indicates that explanations focus on lengthy and/or deeper features of the mathematics, whereas a rating of 1 indicates that

---

[25] IRE stands for Initiation, Response and Evaluation and is described in more detail in Mehan (1979). IRF stands for Initiation, Response and Feedback and is described by Sinclair and Coulthard (1975).

there are no explanations of why ideas or procedures are the way they are, either by the teacher or the students.

The lesson videos were also rated for whether there were *discussion opportunities* present or not in each 16-minute segment. Discussions were defined as extended conversations between and among the teacher and many students, where students do much of the talking. Though the teacher guides the discussion towards a learning goal, discussions are predominantly based on student ideas and characterised by student-to-student interaction.

Lesson artefacts were rated on only one component, *asking for explanations*, within the *discourse* domain. This measures the extent to which students are asked to explain or justify their thinking about mathematical procedures and concepts. Students may be asked to explain their mathematical processes or ideas, to describe mathematical relationships, or to develop arguments that justify why processes work, why expressions are equivalent, why mathematical relationships are true, or why quantitative reasoning is justified or valid.

What was the nature of discourse in mathematics classrooms?

In England, the lessons observed in the TALIS Video Study had a variety of teacher and student discourse with an average rating of 2.44 for the domain. Typically lessons included discourse that was sometimes teacher-directed but that could be also occasionally characterised by detailed student contributions. Lessons usually included questions that requested students to recall, report an answer or define terms, and questions that requested students to summarise, explain, classify, or apply rules, processes or formulae. There were also some lessons that included questions that requested students to analyse, synthesise, justify, or conjecture. On average, lessons included explanations of why ideas or procedures are the way they are, though in many lessons these focused on brief or superficial features of the mathematics.

There was also little variation between teachers in *discourse* ratings as illustrated in Figure 5.1. The majority of teachers had an average rating of between 2 and 3 for each component and the domain overall, with over three-quarters of the teachers having an average rating of between 2 and 3 for *questioning* and the domain overall.

Figure 5.1: Distribution of average teacher ratings for the discourse domain
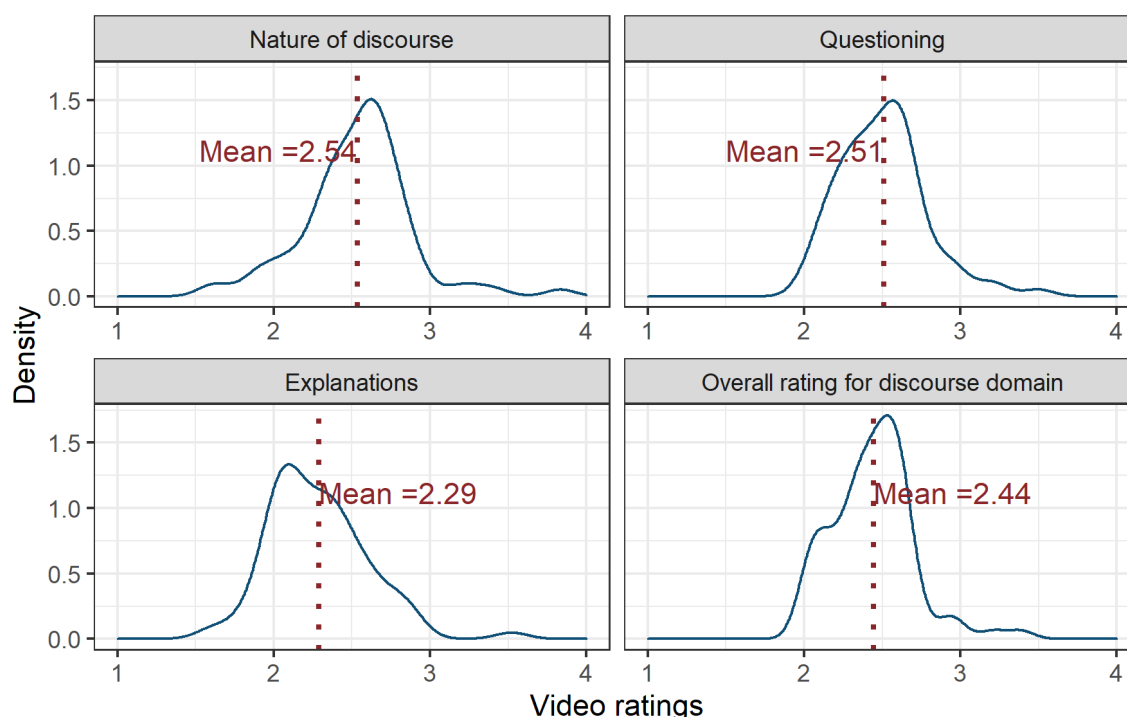
The lessons were also rated for whether there were opportunities for extended conversations between the teacher and the students, or between students, where students do much of the talking. Only a minority (17%) of teachers used these *discussion opportunities* such that a segment of instruction engaged students in discussions that were clear and focused on the learning objectives in their lessons.

The average rating of the artefact component *asking for explanations* within the *discourse* domain was 1.4, suggesting that students were not commonly asked for explanations of their mathematical thinking in the lesson artefacts. A rating of 1 indicates that students may have been asked to recall facts or definitions, or to follow algorithms, but they were not asked for explanations of how or why mathematical procedures or relationships work. In the majority of the lessons the artefacts did not include questions and tasks that asked students for explanations.

The majority (72%) of teachers did include some degree of *asking for explanations* within the four lessons that artefacts were collected for, with at least one lesson scoring a rating of 2. This signifies that students were asked to show their work, describe the sequence of steps they used, or explain why expressions are equivalent. To achieve a rating of 3, however, lesson artefacts had to include the students being asked to explain *how* and to explain *why* a particular mathematical procedure or relationship works within the same lesson. Around 60% of teachers asked students for some form of explanation in the

lesson artefacts for at least one of their lessons, but only a small minority of teachers (n = 6) asked students for explanations of both how and why in at least one of their lessons.

The average rating of *asking for explanations* did not change significantly over the trajectory of the four lessons for which artefacts were submitted, meaning that teachers did not ask students for a higher level of explanations of their thinking as the lesson sequence on quadratic equations progressed. Additionally, whilst the lesson artefacts on their own may not have explicitly shown evidence of students being asked to explain why or how, teachers may have provided oral instructions that clarified their expectations as to how much students should explain/justify their answers or annotate their work.

What were teachers' and students' perceptions of the opportunities for classroom discourse during the lessons?

In the questionnaire at the end of the teaching unit the teachers were asked a series of questions about how frequently they gave students opportunities to engage in discourse during the unit containing quadratic equations. Students were asked the same questions but in relation to their perceptions of their mathematics teachers' use of these opportunities. Table 5.1 shows the proportion of teachers and students who reported that these tasks and activities happened frequently or always during the topic of quadratic equations. Around three-quarters of teachers and students reported that there were frequently or always opportunities for students to explain their ideas. In general, over half of the teachers and students reported that there were also frequent opportunities to critique arguments made by other students or that students had opportunities to engage in discussion among themselves. Both teachers and students reported fewer opportunities to explain ideas during the topic of quadratic equations than usual in their mathematics lessons.

Table 5.1: Teachers' and students' perceptions of opportunities for discourse

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
I/My mathematics teacher gave these students/me opportunities to explain their/my ideas	93% (73%)	81% (76%)
I/My mathematics teacher encouraged these students/me to question and critique arguments made by other students	65% (42%)	52% (50%)
I/My mathematics teacher required these students/us to engage in discussions among themselves/ourselves	68% (55%)	53% (55%)

Further information on responses to these items can be found in the England Technical Report

Proportions represent the teachers or students who answered frequently or always

Source: TALIS Video Study teacher and student data files

Did teachers with different backgrounds provide different opportunities for discourse in their lessons?

There were no statistically significant differences in the teacher video ratings overall for the *discourse* domain between teachers with different qualifications, numbers of years spent teaching mathematics, or self-efficacy scores. There were no teacher background characteristics where there was a significant difference in ratings on the *nature of discourse*, *questioning*, or *explanations* components.

Did classes with different characteristics experience different opportunities for discourse in their lessons?

There were no significant associations between the video ratings within the *discourse* domain and classes which had different proportions of female students, or different class average scores for personal interest in mathematics with their previous teacher. There was a significant association between classes with higher average socio-economic status based on the number of home possessions and the level of parental education which had higher average ratings for the *discourse* domain overall than classes with lower average socio-economic status, but there were no significant associations for any of the individual components that made up the domain.

Classes with higher proportions of students with a first- or second-generation immigrant background had higher average ratings for the *questioning* component compared to classes with lower proportions of students with a first- or second-generation immigrant background. Additionally, classes with higher proportions of students who spoke English at home had lower average ratings for the *questioning* component compared to classes with lower proportions of students who spoke English at home.

Classes with higher average scores on the pre-test had a significantly positive association with the average video rating overall within the *discourse* domain, and higher ratings on each of the *nature of discourse*, *questioning*, and *explanations* components, than classes with lower average scores on the pre-test. There was a positive association between classes' average score for general self-efficacy as measured on the student pre-questionnaire and the average rating on the *questioning* component, such that classes with higher average self-efficacy were associated with higher average ratings compared to those with lower average self-efficacy, but there were no significant relationships with the other components or the overall rating.

Summary

Classes experienced a mixture of teacher- and student-directed discourse, with students occasionally providing detailed contributions. Lessons included questions that requested students to recall, define, summarise, and apply rules, with students sometimes being asked to analyse, justify, or synthesise material. While students were not commonly asked for explanations of their mathematical thinking in the lesson artefacts, the lesson videos revealed that explanations of concepts or procedures were present but were typically brief or focused on mainly superficial features of the mathematics. There were no significant differences in the opportunities for discourse provided by teachers with different backgrounds. Classes with higher average attainment tended to experience higher average ratings on the *nature of discourse*, *questioning* and *explanations* compared to classes with lower average attainment on the pre-test. Classes with a higher average socio-economic status tended to have higher average ratings for the *discourse* domain overall than classes with a lower average socio-economic status.

6. What was the quality of subject matter in mathematics classrooms?

Key findings

- Teachers had a high level of clarity in their lessons, with students easily following the logical progression of content.

- The majority of teachers provided few explicit opportunities for students to look for patterns or to use repeated reasoning to understand mathematical relationships.

- The majority of both students and teachers reported that teachers provided summaries of recently learned content, goals at the beginning of instruction, and explanations of how new and old topics were related.

- Multiple representations of the same mathematical idea were commonly present in lessons, with teachers making explicit connections between representations.

- Students were rarely exposed to connections between mathematics content and real-world contexts.

- The majority of teachers only occasionally planned variations of activities to reinforce concepts or extend to more advanced content; however, teachers and students overwhelmingly agreed that teachers' instruction did adapt based on students' learning needs.

- Teachers reporting higher levels of self-efficacy were associated with higher average ratings for the clarity of their lessons.

- Classes with higher average levels of students' socio-economic status, in terms of the number of home possessions and level of parental education, experienced lessons with more explicit connections to other topics.

- Class average attainment on the pre-test and class average general self-efficacy scores from the pre-questionnaire were both positively associated with the teacher average rating for clarity.

Introduction to the domain

The three main components of teaching practices that were considered within the domain of the *quality of subject matter* include *explicit connections*, *explicit patterns and generalisations*, and the *clarity* of teaching. Each were measured on a rating of 1 to 4, where a higher rating represented higher quality or higher prevalence of particular

behaviours. Videos were also given a holistic rating across the three main components. In addition, analysis of the videos included indicators of the sharing of learning goals, the accuracy of the mathematics, connections to real-world contexts and other mathematical topics, summaries of the mathematics being learnt, the different representations used, and the organisation of the procedural instruction. The artefacts were rated within this domain for *explicit learning goals*, *addressing diverse student needs*, *connecting mathematical representations*, *explicit patterns and generalisations*, and *real-world connections*.

Video components

The *explicit connections* component measures the extent to which the teacher or students make explicit instructional connections between any two aspects of the subject matter. These aspects include ideas, procedures, perspectives, representations, or equations. Explicit instructional connections between quadratic equations and mathematical topics outside of quadratic equations or in the real world were counted if they concerned ideas, equations, representations, perspectives or procedures in those topics and real-world settings. The detail about the connection could be provided for any aspect of the mathematics, regardless of the importance to the mathematics, or the surface or deep nature of the mathematics. A rating of 4 indicates that there were at least two instructional connections between ideas, procedures, perspectives, representations, or equations and that these connections were explicit and clear, and at least one was elaborated. A rating of 1 indicates that either there were no instructional connections or that the connections that were present were implicit.

The *explicit patterns and generalisations* component measured whether the teacher or students explicitly looked for patterns in their work together, and whether they also generalised from the specific work students were doing to a foundational concept and/or definitions underlying the specific work. A pattern was defined as an ordered set of mathematical objects, such as numbers, equations, graphs, problems, or a recurring sequence. The definition of a generalisation was taken from Kaput (1999, p.136)[26] and involves "deliberately extending the range of reasoning or communication beyond the case or cases considered, explicitly identifying and exposing commonality across cases, or lifting the reasoning or communication to a level where the focus is no longer on the cases or situations themselves, but rather on the patterns, procedures, structures, and the relations across and among them (which, in turn, become new, higher level objects of reasoning or communication)". There needed to be at least two examples referred to or investigated from which a generalisation or pattern is developed, and students must have been explicitly asked to look for the pattern. A rating of 4 indicates that the teacher or the students looked for patterns in the mathematical work, and these identified patterns focused on one or more deeper features of the mathematics, or that explicit

---

[26] Kaput, J. J. (1999, p.136).

generalisations were developed from the mathematics under consideration and focused on foundational concepts, ideas and/or definitions. All explicit connections needed to be clear and correct. A rating of 1 indicates that neither the teacher nor the students looked for patterns in the mathematical work or they did not generalise from the mathematical work.

Clarity refers to the extent to which the mathematical content around the learning goal of the lesson is presented clearly and students appear to follow along with the content of the lesson. One way in which raters determined if the mathematics was clearly presented was to observe students' questioning patterns. When multiple students asked the same question about how two things are related, or what the next step in the process would be when the teacher had recently stated how they were related or what the steps were, these were taken to be an indication of a lack of clarity. A rating of 4 indicates that the mathematical concepts, tasks, student response patterns, or discussions in the lesson were clear and that there were no instances in which students demonstrated that they did not understand a logical element of the lesson. A rating of 1 indicates that the mathematical concepts, tasks, student response patterns, or discussions in the lesson were generally murky. There were also multiple instances in which students demonstrated they did not understand the same logical element(s) of the lesson, that is, there was a pattern to students' behaviours around clarity.

Video indicators

The lesson videos were also rated for *explicit learning goals*, *accuracy*, *real-world connections*, *connecting mathematical topics*, *organisation of procedural instruction* and *mathematical summary*. Each of these indicators is rated on a scale of 1 to 3. For *explicit learning goals* a rating of 3 indicates that the teacher explicitly stated or wrote the learning goal(s), whereas a rating of 2 indicates that the teacher explicitly stated or wrote the activities or topic(s) but without explicitly stating the learning goal(s). *Real-world connections* measures the extent to which what was being learned was connected or applied to something outside of school, a real-life problem, or a student's life experiences. A rating of 3 on this indicator means that there was more than one moderate connection or at least one strong connection between the mathematical content being learned and real-life problems or students' life experiences. *Connecting mathematical topics* measures the extent to which the topic being learned was connected to other mathematical topics (not including quadratic functions). A rating of 3 means there was more than one moderate connection or one strong connection between the mathematical topic being learned and other mathematical topics. *Organisation of procedural instruction* measures the degree of organisation, detail and correctness in the presentation of content when describing procedures or the steps of a procedure, where procedures are defined as instructions for completing a mathematical algorithm or task. A rating of 3 means the procedures presented were correct, well organised and detailed. *Mathematical summary* measures the extent to which the teacher or students provided a

summary of the mathematics under consideration in that particular lesson, that is, a review of what had or should have been learned in that lesson. A rating of 3 means there was at least one explicit summary of the mathematics being learned. This summary reviewed a significant amount of mathematical work the class had done and the summary was clear.

The videos were also rated for the *type of representations* used, including graphs, tables, drawings or diagrams, equations and expressions, or objects. These were rated as present or not present.

Artefact components

The artefacts were rated across six components: *accuracy of materials*, *explicit learning goals*, *addressing diverse student needs*, *connecting mathematical representations*, *explicit patterns and generalisations*, and *real-world connections*. *Explicit learning goals* measures the extent to which the teacher has explicit mathematical learning goals for the students. Like the indicator measure, a rating of 3 means that the lesson materials described specific learning objectives for students that were framed with specific reference to student learning, mastery or understanding. *Connecting mathematical representations* measures the extent to which students are asked to make connections between multiple representations of the same mathematical idea. A rating of 3 means the students were asked to make the connection, whereas a rating of 2 means the teacher or the lesson materials made the connection. *Explicit patterns and generalisations* measures the extent to which students are asked to use patterns or repeated reasoning to understand quantitative relationships, make conjectures, make predictions, or derive general methods or rules. A rating of 3 means the students were asked to use the patterns or make the generalisations, whereas a rating of 2 means the teacher or lesson materials used the patterns or made the generalisation. *Real-world connections* measures the extent to which the mathematics being learned is connected to or applied to real-world contexts outside of mathematics (including but not limited to students' experiences), and the features of the context are incorporated into the activity (that is, the connection is relevant to solving the problem). A rating of 3 means that the students had to figure out how to connect or apply mathematics to a real-world context. A rating of 2 means that some of the mathematics in the lesson was set in a real-world context, but the context was unnecessary to carry out the activity, or the model that connected mathematics to a real-world situation was given to students. *Addressing diverse student needs* measures the extent to which the artefacts reveal planned variation in lesson content based on students' learning needs. A rating of 3 means that the lesson materials included both reinforcement and advancement options for students who could benefit from them. A rating of 2 means that either the materials included reinforcement or they included advancement options. *Accuracy of materials* indicates whether the materials include a major mathematical error. "Major errors" includes problems that were worked

out incorrectly during the lesson, terms that were defined incorrectly, or the equating of two non-equivalent mathematical terms.

What was the quality of subject matter in mathematics classrooms?

In England the lessons in the TALIS Video Study were generally clear with an average rating of 3.37 for the *clarity* component. The use of *explicit connections* and *explicit patterns and generalisations* were not as common, as illustrated in Figure 6.1. There was little variation in the *clarity* ratings across teachers. Over three-quarters of teachers had an average rating between 3 and 4 for *clarity*, with some teachers having the maximum rating across both lessons. Similarly, there was little variation in the amount teachers included *explicit patterns and generalisations* within their lessons with over three-quarters of teachers with an average rating between 1 and 2.

Figure 6.1: The distribution of ratings for the quality of subject matter domain
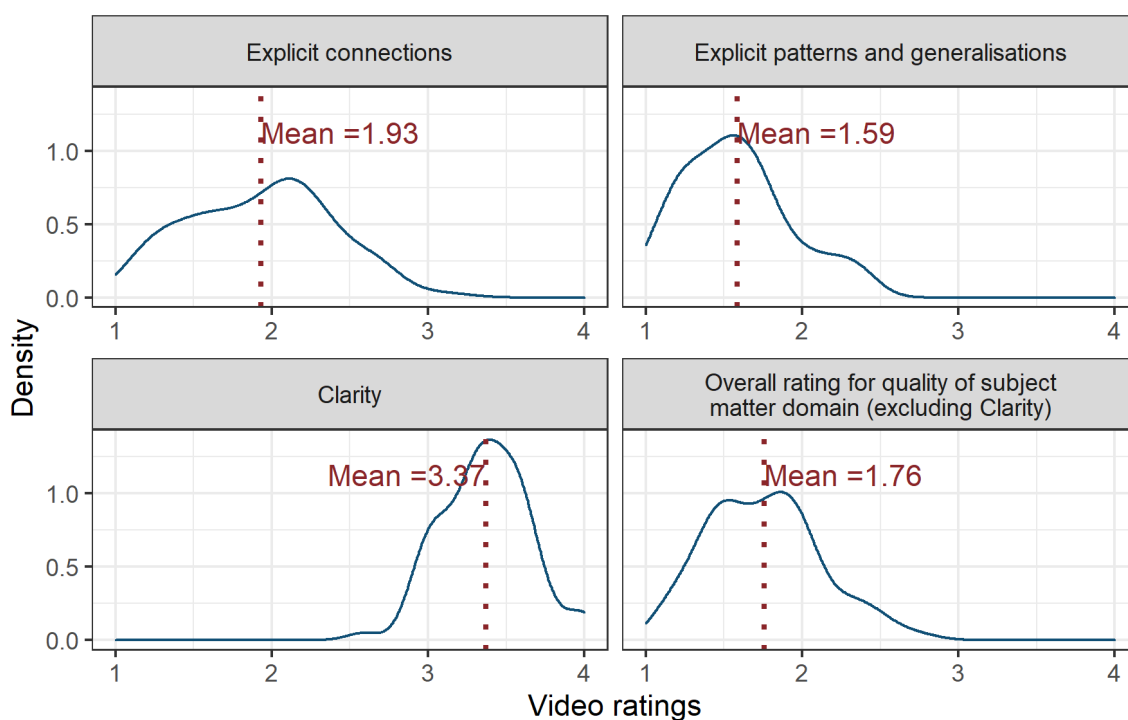
The use of explicit learning goals

Almost all teachers included explicit goals in both of their lessons that were videoed in the study, and around two-thirds (68%) of the lessons had explicit learning goals. In addition, the lesson artefacts were analysed for the extent to which the teacher had

explicit mathematical learning goals for students. Lesson artefacts received an overall average rating of 2.3 on *explicit learning goals*. This suggests that teachers often have a clear sense of the topics they intend to teach in a lesson, or how they will cover these topics. With the overall average above 2, teachers were sometimes using explicit student learning goals. The proportion of teachers that included explicit learning goals that focused on student learning in at least one lesson was 54%. Teacher averages on this component covered the full range of the 3-point scale; this wide distribution illustrates the large variety amongst teachers when it comes to explicitly including learning goals for students in their lesson artefacts. Note that the rating of lesson artefacts on *explicit learning goals* does not take into account whether the learning goals were communicated to students.

Connections and representations

Real-world connections were measured in both the videos and the artefacts. Over 70% of teachers did not use *real-world connections* or only used weak *real-world connections* during the lessons that were videoed. The teacher average rating for *real-world connections* was 1.04. The average rating of *real-world connections* on the artefacts was 1.3, suggesting that students were not often exposed to real-world contexts within the topic of quadratic equations. The majority of teachers had an average rating of less than 2 across the 4 sets of lesson artefacts.

A distinction was made between real-world contexts where the context was unnecessary to carry out the activity or the model that connected mathematics to a real-world context was given to the students (a rating of 2), and where students had to figure out how to connect or apply mathematics to a real-world context (a rating of 3). While a small proportion (19%) included a real-world context where the students had to figure out how to connect or apply mathematics to the real-world context in at least one lesson, a large minority of 26% of teachers had a maximum rating of 1 across the lessons, meaning that no *real-world connections* were provided at all in their four-lesson sequence.

The *connecting mathematical topics* indicator shows that very few connections were made in the videoed lessons to other mathematical topics, with an average teacher rating of 1.07. Less than half of the lessons included a moderate or strong connection to other mathematical topics.

Lessons were rated every 8 minutes for the *types of representations* used. The representations considered were graphs, tables, drawings or diagrams, equations, or physical objects. Equations were used in almost every segment of a lesson as would be expected for lessons focusing on quadratic equations. Around a fifth of lesson segments used graphs and slightly fewer used tables, drawings, or diagrams. Physical objects or models were very rarely used.

The average rating of *connecting mathematical representations* was 2.1, with around two-thirds (69%) of teachers receiving an average rating of 2 or higher. Multiple representations of the same idea were commonly present in lessons, and teachers made the connections between mathematical representations of the same idea explicit for students. Three-quarters of the teachers (89%) asked students to make a connection between mathematical representations themselves in at least one lesson. Only two teachers did not include any connections in their lesson artefacts for any of the lessons. Connecting one mathematical representation to another, such as being able to represent an equation graphically, can help deepen students' understanding of mathematical concepts.

The majority of teachers included a *mathematical summary* in at least one of their lessons, though these generally reviewed a small amount of the mathematical work the class had done or were somewhat vague. Over half of the lessons in the study did not include a summary, either by the teacher or the students, of the mathematics being learned.

The average teacher rating for *organisation of procedural instruction* was 1.91. However, since a rating of 1 could mean that there were no procedures in the segment, or that the procedures presented were somewhat correct, or that the procedures presented were not particularly organised and lacked detail, it is not clear whether there were no procedures in many of the lesson segments or whether the procedures that were present were not organised.

What was the quality of subject matter in the lesson materials?

The average rating of *explicit patterns and generalisations* for artefacts was 1.3 meaning that few teachers on average included *explicit patterns and generalisations* within the artefacts for their lessons. Over half of teachers did include *explicit patterns and generalisations* within at least one of the lesson artefact sets. Fewer than a fifth of teachers had at least one set of lesson artefacts that had a rating of 3 for *explicit patterns and generalisations*, and no teachers had a rating of 3 across all of their lessons.

The average rating of *addressing diverse student learning needs* was relatively low at 1.5. Overall there were few variations provided to support diverse student learning needs within the lesson artefact sets. Only a fifth of teachers (20%) had an average rating of 2 or more. However, the majority of teachers provided variation of some kind at some point within their artefacts for the four lessons, with a quarter (26%) of teachers providing both support and extension options within at least 1 lesson artefact set. This does not necessarily mean that many teachers did not vary their lesson content based on students' learning needs, only that the artefacts on their own do not show evidence of this variation.

What were teachers' and students' perceptions of how teachers adapted their teaching based on students' learning needs?

In the questionnaire at the end of the teaching unit the teachers were asked a series of questions about how teaching was adapted during the unit containing quadratic equations. Students were asked the same questions but in relation to their perceptions of their mathematics teachers' adaptations of their teaching. Table 6.1 shows the proportion of teachers and students who strongly agreed or agreed with each statement when asked about their perceptions of the lessons in general and during the topic of quadratic equations. All the teachers and most students agreed or strongly agreed that the teacher adapted the lessons to the class's needs and knowledge and almost all teachers and students agreed or strongly agreed that the teacher changed their way of explaining when a student had difficulties understanding a topic or task. This adaptation of the lessons to meet the class's needs and knowledge did not always involve giving different work to students with different attainment levels, with slightly more than two-thirds of teachers reporting that they gave different work to different students, and around half of the students reporting this.

Table 6.1: Teachers' and students' perceptions of how teachers adapted instruction based on student learning needs

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
I/Our mathematics teacher adapts the lessons to this/my class's needs and knowledge	100% (100%)	88% (86%)
I/Our mathematics teacher changes the way of explanation when a student has difficulties understanding a topic or task	100% (99%)	88% (83%)
I/Our mathematics teacher changes the structure of the lesson on a topic that most students find difficult to understand	95% (88%)	78% (65%)
I/Our mathematics teacher gives different work to students of different ability levels	71% (69%)	46% (49%)

Further information on responses to these items can be found in the England Technical Report
Proportions represent the teachers or students who agreed or strongly agreed with each statement
Source: TALIS Video Study teacher and student data files

This is also seen in the artefact analysis of *addressing diverse student learning needs*. Whilst the artefacts from the lessons showed little evidence that there was planned variation in lesson content based on students' learning needs, both teachers and students reported that the lessons themselves were adapted to the class's needs and knowledge of the topic, including teachers changing the way they explained when students had difficulties, changing the structure of the lessons and giving different work to students with different levels of attainment.

What were teachers' and students' perceptions of the clarity of teaching?

In the questionnaire at the end of the teaching unit the teachers were also asked a series of questions about the clarity of their teaching during the unit containing quadratic equations. Students were asked the same questions but in relation to their perceptions of their mathematics teachers' clarity of instruction. Table 6.2 shows the proportion of teachers and students who reported that particular activities related to clarity happened frequently during the topic of quadratic equations. For all activities associated with the clarity of the lessons the majority of teachers and students reported that they happened frequently. Almost all teachers reported that these activities happened at least occasionally, with at most three teachers reporting that they never did these activities during the series of lessons. In general students reported that the activities associated with clarity happened less frequently than what their teachers reported.

Table 6.2: Teachers' and students' perceptions of the clarity of lessons

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
I/My mathematics teacher presented a summary of recently learned content	70% (68%)	56% (56%)
I/My mathematics teacher set goals at the beginning of instruction	80% (79%)	64% (68%)
I/My mathematics teacher explained what I/they expected these students/us to learn	92% (81%)	82% (79%)
I/My mathematics teacher explained how new and old topics are related	86% (80%)	62% (62%)

Further information on responses to these items can be found in the England Technical Report
Proportions represent the teachers or students who agreed or strongly agreed with each

Did teachers with different backgrounds provide different experiences in the quality of subject matter in their lessons?

For this domain, the individual components were considered separately, and no overall average rating was used in the tests for significance due to the weak associations between the different components.

There were no significant differences within the *quality of subject matter* domain between teachers of different genders, with different qualifications, or in relation to the number of years' experience a teacher had in teaching mathematics. There was a positive association between teachers reporting higher general self-efficacy scores and their average rating on the *clarity* component.

Did classes with different characteristics experience a different quality of subject matter in their lessons?

There were no significant associations between classes with different proportions of female students within the *quality of subject matter* domain. There were also no significant associations between classes with different proportions of students who did not speak English at home or with different proportions of students with a first- or second-generation immigrant background on the average ratings within the *quality of subject matter* domain.

There was a significant positive difference between classes with higher average levels of students' socio-economic status, as measured by their parents' or carers' education level and number of home possessions, and the teacher average ratings for *explicit connections* but not with any of the other components within the domain, nor on the overall domain rating. Classes with higher average levels of students' socio-economic status experienced lessons where the teacher or students made explicit connections between any two aspects of the subject matter more frequently than classes with lower average levels of students' socio-economic status.

There were no significant associations between classes with higher average levels of students' personal interest in mathematics or general self-efficacy as measured on the pre-questionnaires and the *explicit connections*, or *explicit patterns or generalisations* component ratings. There was also no significant association between classes with higher average levels of students' personal interest in mathematics and the overall rating for the *quality of subject matter* domain. There was a significant positive association between the *clarity* component and classes with a higher average score for general self-efficacy in mathematics. This was also true for classes with a higher average score on

the pre-test who also experienced lessons with higher average video ratings for the *clarity* component compared to classes with a lower average score on the pre-test.

Summary

Classes in the TALIS Video Study in England experienced clear lessons but with limited opportunities to understand connections between mathematical content or form generalisations based on patterns. There were no associations between teachers' background characteristics and their average rating for the overall *quality of subject matter* domain. There was a significant positive association between classes with higher average levels of students' socio-economic status with the frequency and quality of *explicit connections* their teachers made. There was also a significant positive association between the *clarity* component and classes with a higher average general self-efficacy score. This was also found for classes with a higher average score on the pre-test, who also experienced lessons with higher average video ratings for the *clarity* component compared to classes with a lower average score on the pre-test.

7. How were students cognitively engaged?

Key findings

•   All teachers provided a wealth of opportunities for students to develop fluency with specific mathematical skills through practice or repetition.

•   Lessons did not often ask students to engage with cognitively demanding subject matter.

•   All teachers provided students with some opportunities to understand the rationale behind the procedures and processes with which they were working, and almost three-quarters of teachers and students reported that teachers frequently or always asked questions that helped students understand how or why a procedure works.

•   Most teachers rarely provided opportunities for students to use multiple-solution strategies or reasoning approaches, and the artefacts showed that teachers occasionally included opportunities for, but did not require the use of, multiple mathematical methods.

•   Around half of the teachers did not use technology beyond making communication more efficient.

•   Many more students than teachers believed that cognitively demanding tasks were frequently or always provided.

•   Classes with a higher average score on the pre-test tended to experience lessons with higher average video ratings for the *engagement in cognitively demanding subject matter*, *multiple approaches to or perspectives on reasoning*, and *understanding of subject matter procedures and processes* components, and overall, compared to classes with a lower average score on the pre-test.

Introduction to the domain

For the domain of *student cognitive engagement* the three components were *engagement in cognitively demanding subject matter*, *multiple approaches to or perspectives on reasoning*, and *understanding of subject matter procedures and processes*. Each were rated on a scale of 1 to 4, where a higher rating represented higher quality or higher prevalence of particular behaviours. Videos were also given a holistic rating across the three main components. In addition, analysis of the videos included indicators of *metacognition*, *opportunities for practice*, and the use of *technology*. Lesson artefacts were rated within this domain for *opportunities to practice a*

skill or procedure, *using multiple mathematical methods*, and *technology for understanding*.

The *engagement in cognitively demanding subject matter* measure described the extent to which students engaged in analyses, creation, or evaluation work that was cognitively rich and required thoughtfulness. Analysis was defined as the detailed examination or exploration of the features and relationships among mathematical procedures, processes, ideas, or topics. Creation was defined as formulating or inventing a way to solve a problem or devising a way to solve a new problem or type of problem. Evaluation was defined as determining the significance or conditions of a mathematical idea, topic, representation, or process. Work that engaged students' cognitive processes beyond recall, recitation, and the rote application of procedures was described as cognitively rich and requiring thoughtfulness. Such work frequently requires students to grapple with problems and ideas and the relationships among mathematical ideas, topics, representations, and processes. The focus was only on students' engagement in cognitively demanding subject matter. A rating of 4 within this component indicates that students frequently engaged in analyses, creation, or evaluation work that was cognitively rich and required thoughtfulness. A rating of 1 indicates that there was only a brief engagement with cognitively demanding subject matter or no engagement at all.

Multiple approaches to and perspectives on reasoning measured the extent to which students used multiple-solution strategies and/or reasoning approaches. It focused on multiple approaches that students used to solve problems, not multiple solutions that they came up with. It included teachers having various students solve one type of quadratic equation using different approaches, or having students approach a single problem using different approaches. A rating of 4 indicates that students generally used two procedures or reasoning approaches to solve the problem or type of problem, or they used more than two procedures or reasoning approaches to solve the problem or type of problem in some depth. A rating of 1 indicates that students generally used a single procedure or reasoning approach to solve the problem or type of problem, or that there was no evidence of how many approaches students were using.

Understanding of subject matter procedures or processes described the extent to which students engaged in opportunities to understand the rationale(s) for subject matter procedures and processes. This included students stating the goals or properties of procedures and processes, stating why a procedure or a solution was the way it was, or visually designating the elements or steps in a process or procedures. It included understanding that was visible in students' spoken words or written work. It captures whether students understand why or how a procedure works or what makes that procedure or process appropriate, which is different from students understanding what a procedure is. This component measured the frequency with which students attended to the rationale for the procedures and processes they engaged with. A rating of 4 indicates that they did this frequently whereas a rating of 1 indicates that either they did not

engage in the rationale or there was no evidence that they were attending to the rationale for the procedures and processes.

To what extent did teachers ask students to engage in cognitively demanding mathematics?

In England the lessons in the TALIS Video Study did not often ask students to engage with cognitively demanding subject matter, with an overall average rating of 1.86. Students did attend to the rationale for the procedures and processes they were working with in their lessons, but not very often. Students' opportunities to use two or more procedures or reasoning approaches to solve problems or types of problems, and their opportunities to engage in analyses, creation, or evaluation work that was cognitively rich and required thoughtfulness were more limited, as illustrated in Figure 7.1.

Figure 7.1: Distribution of teacher average ratings for the student cognitive engagement domain
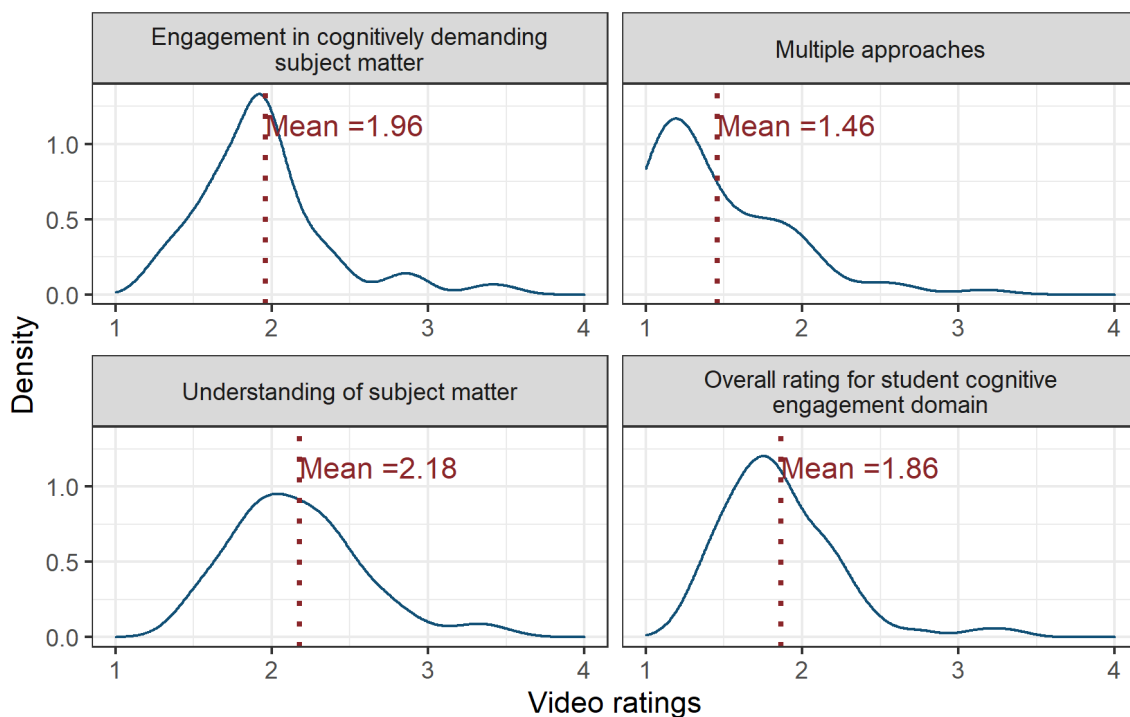
There was some variation between teachers in students' opportunities to engage in cognitively demanding subject matter, again illustrated in Figure 7.1. All teachers asked their students to engage in analyses, creation, or evaluation work that was cognitively rich and required thoughtfulness at some point. All teachers also gave their students opportunities to understand the rationale for the procedures and processes within the

lesson. However, some teachers never gave their students opportunities to use multiple-solution strategies or reasoning approaches in the videoed lessons.

Multiple methods often exist to solve the same type of mathematical problem, and an awareness or comparison of different methods can aid students in understanding the strengths of various approaches, as well as where and why they might use one method over another. Lesson artefacts were analysed regarding the types of opportunities provided for students to use and compare multiple mathematical methods on a single, or a set of related, mathematical tasks or activities.

The average rating for *using multiple mathematical methods* was 1.4, on a 3-point scale, revealing a lack of tasks and activities where students are required or have the option to use multiple mathematical methods. A rating of 1 signifies that students were given tasks with a single specified, or otherwise evident, approach. A rating of 2 indicates that students had the *opportunity* to use different methods to complete a task or activity, or a set of tasks or activities, whereas a rating of 3 indicates that the students were *required* to use multiple mathematical methods. On average, teachers only occasionally included opportunities for their students to use multiple mathematical methods in their lessons, though the majority of the teachers did include these opportunities at least once in the artefact sets for their lessons.

The videoed lessons were also rated for the opportunities students had to engage in the repetitious use of a specific skill or procedure. All teachers gave students opportunities to repetitively use a specific skill or procedure and, in many cases, students spent more than half of the 8-minute segment repetitively using the specific skill or procedure. The average rating for *repetitive-use opportunities* was 2.96 on a 3-point scale, where a rating of 1 indicates that students did not engage in the repetitive use of a specific skill or procedure. A rating of 2 represents that students had an opportunity to use a procedure or skill for less than half of the segment whereas a rating of 3 represents that they had this opportunity for half or more of the segment.

Artefacts from lessons were analysed for the extent to which students were provided with opportunities to develop fluency with specific mathematical skills or procedures through practice or repetition. All teachers gave their students *opportunities to practice a skill or procedure*, with an average rating of 2.8. Furthermore, many teachers provided a wealth of opportunities for students to practice in mathematics lessons, with no teachers having an average rating below 2. To achieve a rating of 3, artefacts had to demonstrate that students were provided with more than five opportunities to repeat and develop fluency with a specific mathematical skill or procedure during a lesson or the homework from that lesson. A rating of 2 indicates that up to five opportunities were provided and a rating of 1 indicates that no opportunities were given.

Almost all teachers provided students with more than five opportunities to repeat and develop fluency in at least one lesson artefact set, with just two teachers offering a maximum of five opportunities within all their lessons. The amount of practice students

received remained approximately the same across the four lessons for which artefacts were submitted.

The lesson videos were also rated on an indicator that measured whether teachers asked students to engage in *metacognition* by explicitly asking students to reflect on their own thinking. The average rating on this indicator was 1.06. The maximum rating of 3 was given when students were asked to engage in *metacognition* longer than briefly and/or in some depth and a rating of 1 was given when students were not asked to engage in *metacognition*. Only around one in ten lessons included an opportunity for students to engage in *metacognition* either briefly and/or in some depth.

The use of technology in mathematics lessons

The increased use of technology in classrooms has led to questions about the role it can play in the learning process and when it is being used to aid students' understanding versus simply being used. Both the videos and the lesson artefacts were rated for what technology was used, as well as for how it was used to support understanding. Almost all teachers used a smartboard during lessons with 85% of the 16-minute segments including their use, but the only other technology used with the whole class was non-graphical calculators (4% of segments) and visualisers or overhead projectors (3% of segments). Around 11% of lesson segments involved no class technology. Software was used to assist or support learning of the mathematical topic through simulations, instructional games, interactive graphing etc. in around 10% of lessons.

In around a quarter of lesson segments students used calculators, but otherwise technology was rarely used by students individually or in small groups. Only a few lesson segments involved students using graphical calculators, computers, or tablets. The videos were also rated on the extent to which students used *technology for understanding*. Around 77% of lessons included the use of technology, with an average rating of 2.10 indicating that teachers in England largely used technology for communication. *Technology for understanding* was rated on a 4-point scale where a rating of 4 means that technology was used exclusively for conceptual understanding or for a mixture of communication and conceptual purposes, where it was used at least once to support conceptual understanding. A rating of 3 means that technology was used once in a limited way to support conceptual understanding. A rating of 2 means that technology was only used for communication purposes, and a rating of 1 on this indicator indicates that technology that required electricity was not used.

Lesson artefacts were analysed for the extent to which students were given opportunities to use dynamic computerised learning tools to make and test conjectures, look for patterns, and develop understanding of mathematical concepts and relationships. A rating of 1 indicates that a lesson did not use technology, or it used technology only to make communication more efficient, such as students viewing projected slides. A rating of 2 indicates that the lesson used technology as a tool to make computation or graphing

more efficient, to reinforce teaching, for practice, assessment or feedback to the teacher, or for checking correctness (e.g. students being asked to use a calculator to check their answers). A rating of 3 indicates that the lesson used technology to explore mathematical relationships and extend students' understanding.

The average rating of *technology for understanding* from the lesson artefacts was 1.2. Approximately half (51%) of the teachers received an average rating of 1 on their lesson artefacts when it came to *technology for understanding*, meaning that technology was not used beyond visual presentations in any of their lesson artefacts. Less than 10% of teachers used technology to explore mathematical relationships and extend students' understanding at least once.

One possibly hidden source of error for this component is that lesson artefacts alone may not always make it apparent how and when certain technologies, such as calculators, are being used, particularly if their use is common practice in that classroom.

What were teachers' and students' perceptions of the cognitive engagement opportunities within the lessons?

In the questionnaire at the end of the teaching unit the teachers were asked a series of questions about how frequently they used questions, tasks, or prompts that gave students opportunities to engage in cognitively demanding subject matter during the unit containing quadratic equations. Students were asked the same questions but in relation to their perceptions of their mathematics teachers' use of these tasks and activities. Table 7.1 shows the proportion of teachers and students who reported that these tasks and activities happened frequently or always, in general and during the topic of quadratic equations. In all cases a larger proportion of students than teachers reported that these tasks and activities happened at least frequently. Whilst many teachers did not report that these tasks and activities happened frequently, the majority did report that they happened at least occasionally. Only 14% of teachers reported that they presented tasks for which there was no obvious solution, whilst 54% of students reported that their teacher did this during the lessons on the topic of quadratic equations. The vast majority of students also reported that their teacher gave them tasks that required them to think critically and tasks that required them to apply what they had learned to new contexts.

Table 7.1: Teachers' and students' perceptions of the cognitive demand of lessons

Statement	Proportion of teachers in general (during topic)	Proportion of students in general (during topic)
I/My mathematics teacher presented tasks for which there is no obvious solution	33% (14%)	49% (54%)
I/My mathematics teacher presented tasks that required these students/me to apply what they/I have learned to new contexts	78% (54%)	85% (82%)
I/My mathematics teacher gave tasks that required these students/me to think critically	55% (48%)	82% (81%)
I/My mathematics teacher asked these students/me to decide on their/my own procedures for solving complex tasks	48% (33%)	55% (57%)

Further information on responses to these items can be found in the England Technical Report
Proportions represent the teachers or students who said frequently or always for each statement
Source: TALIS Video Study teacher and student data files

These questions were measured on a 4-point scale, ranging from "never" to "always". Assigning each of these points a value from 1 to 4, individual teachers' and students' average perceptions of cognitively demanding tasks and activities can be calculated. Figure 7.2 shows the distribution of this average perception of cognitively demanding tasks and the difference between teachers' perceptions and students' perceptions. Students generally perceived lessons to be more cognitively demanding than their teachers. Several students (6%) stated that their teacher always presented each of the tasks detailed in the statements in Table 7.1 in their lessons, whereas no teacher reported this for all four statements.
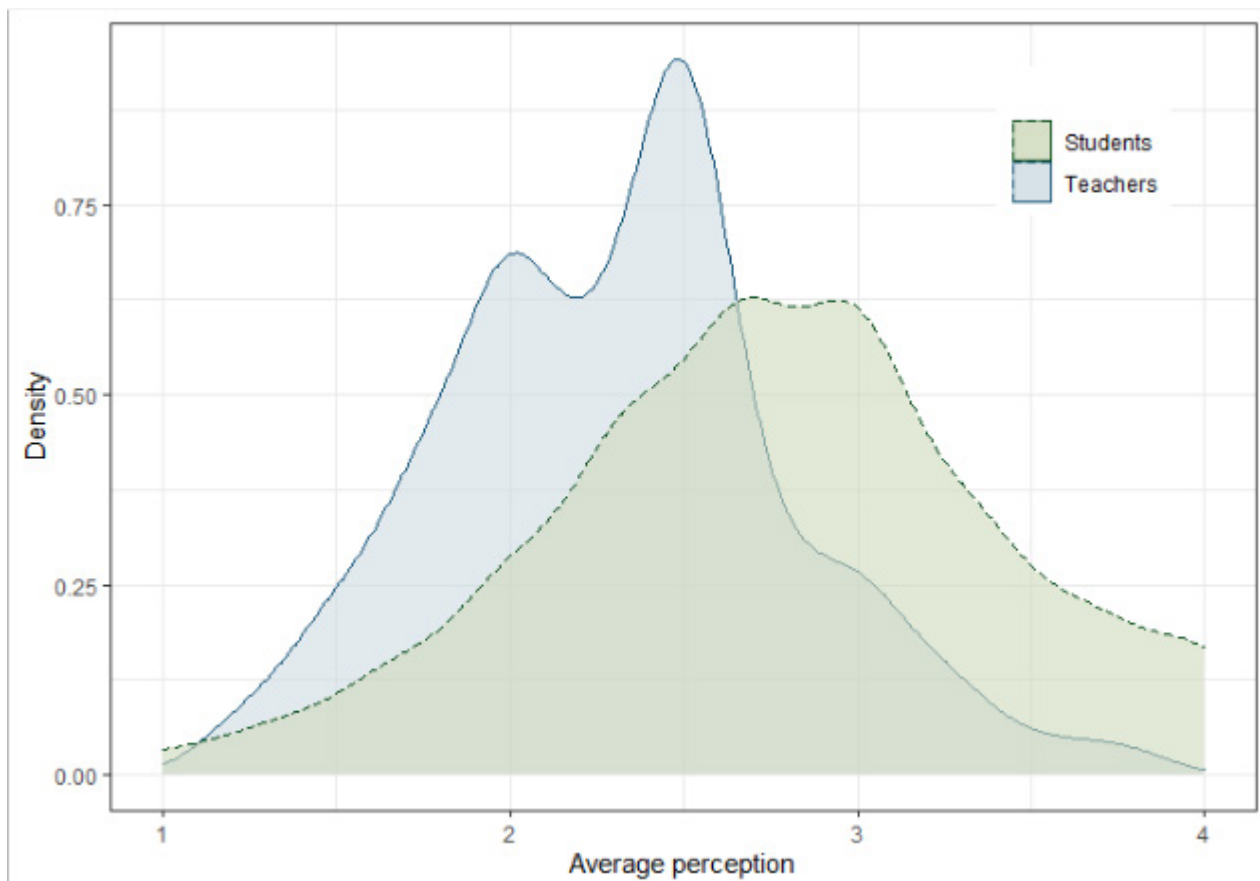
Figure 7.2: Distribution of teachers' and students' perceptions of the cognitive demand of lessons

Source: TALIS Video Study teacher and student data files

Teachers and students were also asked a series of questions around their perceptions of what the teacher did to support students' understanding of the subject matter during the topic of quadratic equations. As Table 7.2 shows, the proportion of teachers and students who reported that each of the teacher actions occurred frequently or always were similar. Almost three-quarters of teachers and students reported that the teachers frequently or always asked questions that helped students understand why a procedure works. A majority of both teachers and students also reported that teachers frequently or always explained why a mathematical procedure works, illustrated why a mathematical procedure works using concrete examples or graphics, and compared different ways of solving problems.

Table 7.2: Teachers' and students' perceptions of teachers' support of students' understanding during the topic

Statement	Proportion of teachers during the topic	Proportion of students during the topic
Explained why a mathematical procedure works	68%	68%
Illustrated why a mathematical procedure works using concrete examples or graphics	65%	62%
Asked questions that helped these students understand why a procedure works	75%	71%
Compared different ways of solving problems	62%	76%

Further information on responses to these items can be found in the England Technical Report
Proportions represent the teachers or students who said frequently or always to each statement
Source: TALIS Video Study teacher and student data files

Did teachers with different backgrounds provide different opportunities for students to engage in cognitively demanding activities in their lessons?

There were no significant associations in the *student cognitive engagement* domain related to teachers' gender, teaching qualifications, levels of self-efficacy or with the number of years' experience a teacher had in teaching mathematics. This was the case both for the overall domain rating and for the ratings on the individual components within the domain.

Did classes with different characteristics experience different opportunities for engaging in cognitively demanding activities in their lessons?

There were no significant associations between classes with different proportions of female students and teacher video ratings within the *student cognitive engagement* domain. There were also no significant associations between classes with different proportions of students who did not speak English at home or with different proportions of students with a first- or second-generation immigrant background. There were no significant associations within the domain average ratings and students' socio-economic status as measured by their number of home possessions, though there was a positive

association between the average level of parental education and the teacher average rating for *multiple approaches* and with the overall average domain rating.

There were no significant associations between classes with different average levels of students' personal interest in mathematics or general self-efficacy in mathematics as measured on the pre-questionnaires and the overall domain rating or the *multiple approaches to or perspectives on reasoning* or *understanding of subject matter procedures and processes*. There was a significant association between classes' average score for personal interest in mathematics on the pre-questionnaire and teachers' average rating on the *engagement in cognitively demanding subject matter* component, with classes with a higher average personal interest score experiencing higher average ratings for *engagement in cognitively demanding subject matter*.

There were significant associations between classes with a higher average score on the pre-test and all the average teacher ratings on each of the components and the domain overall. Classes with a higher average score on the pre-test experienced lessons with higher average video ratings for the *engagement in cognitively demanding subject matter*, *multiple approaches to or perspectives on reasoning*, and *understanding of subject matter procedures and processes* components, and overall, compared to classes with a lower average score on the pre-test.

Summary

Students participating in the TALIS Video Study in England were provided with some opportunities for cognitive engagement, with an overall rating of 1.86 on a 4-point scale. Students were not often asked to engage with cognitively demanding subject matter through analyses, creation, or evaluation work. All teachers provided students with some opportunities to understand the rationale behind the procedures and processes with which they were working. However, some teachers never provided students with the opportunity to use multiple-solution strategies or reasoning approaches during the videoed lessons. On average, teachers occasionally included opportunities for multiple mathematical methods, with some teachers not providing these opportunities at all. Teachers provided a wealth of opportunities for students to develop fluency with specific mathematical skills through practice or repetition. Students perceived lessons to be more cognitively demanding than teachers. Teachers with different backgrounds provided students with similar levels of opportunities for cognitive engagement. Classes with a higher average score on the pre-test experienced lessons with higher average video ratings within the *student engagement in cognitively demanding subject matter* domain compared to classes with a lower average score on the pre-test.

8. How did teachers assess students' thinking and respond to it?

Key findings

- Lessons included a range of assessments of, and responses to, student understanding.

- Teachers asked questions or gave prompts that elicited detailed student responses and sometimes used feedback loops to understand students' thinking.

- Most teachers frequently adapted their teaching to student understanding and provided cues or hints when students struggled.

- Students were very rarely provided with opportunities for self-evaluation.

- There was significant negative association between the overall average domain ratings for assessment of and responses to student understanding and the proportion of students within a class who spoke English at home.

- Classes with higher average attainment on the pre-test tended to experience classrooms with more elicitations of detailed student thinking.

Introduction to the domain

The three main components of teaching practices that were considered within the domain of *assessment of and responses to student understanding* were *eliciting student thinking*, *teacher feedback*, and *aligning instruction to present student thinking*. Each was measured on a rating of 1 to 4, where a higher rating represented higher quality or higher prevalence of particular behaviours. Videos were also given a holistic rating across the three main components. For this domain there were no additional indicators of practices from the videos. The artefacts were rated within this domain for *encouraging student self-evaluation*.

The *eliciting student thinking* component measured the extent to which questions, prompts, and tasks elicited detailed student responses. These responses needed to include sufficient detail about the mathematics being worked on, not just short answers that gave the answer or defined a term, for example. Detailed student responses were those that revealed students' thinking processes or rationales. These thinking processes could be revealed by students' step-by-step solving of processes. The component included students' written work but only when there was clear evidence of the nature of that work, for example by one student sharing the steps they used with the whole class.

A rating of 4 indicates that there was a lot of student thinking present and that questions, prompts, and tasks resulted in a mixture of student contributions concerning answers, procedures, the steps necessary for solving a problem, ideas, and concepts. These contributions could be detailed or perfunctory. A rating of 1 indicates that there was no student thinking present.

Teacher feedback measured the extent to which the teacher responded to students' thinking using feedback loops that were focused on why the students' thinking was correct or incorrect, or why ideas or procedures were the way they were. It also measured the extent to which the teacher and student exchanges addressed the mathematics in a complete manner. A feedback loop is a back-and-forth exchange between the teacher and one or more students. Complete feedback loops are responses to students' contributions that address the mathematics at hand in a detailed fashion. A rating of 4 indicates that there were frequent feedback loops where teacher and student exchanges addressed the mathematics in a complete manner. A rating of 1 indicates that there was only one or no feedback loops and teacher and student exchanges addressed the mathematics in a generally limited manner.

Aligning instruction to present student thinking refers to the extent to which the teacher used students' contributions and, if students made errors or struggled mathematically, the extent to which the teacher provided cues or hints to support student understanding. Cues and hints are comments or questions that are intended to move a student's or students' thinking forward given in response to evidence of student thinking, whether that thinking is correct or not. A teacher could draw attention to the student's contribution or features of that contribution, ask a question in response to a student's question or contribution, have students provide the next step in the procedure or process, or acknowledge patterns in student contributions. A rating of 4 indicates that the teacher frequently used students' contributions and that if students made errors or struggled mathematically, the teacher frequently provided cues or hints to support their understanding. A rating of 1 indicates that the teacher did not use students' contributions and when students made errors or struggled with the mathematics the teacher did not provide cues or hints to support their understanding.

How were teachers making assessments of and responding to student thinking?

In England the lessons in the TALIS Video Study included a range of *assessments of and responses to student understanding* with an overall average rating of 2.70. Teachers generally asked questions, gave prompts or used tasks that elicited student responses, responded to students' thinking using feedback loops that focused on why the students' thinking was correct or incorrect or why ideas or procedures are the way they are, used students' contributions, and provided clues or hints to support student understanding when students made errors or struggled mathematically.

Figure 8.1 illustrates the variation in teacher average ratings across the *assessment of and responses to student understanding* domain. More than three-quarters of teachers had an average rating above 3 for *aligning instruction to present student thinking*. However, for *eliciting student thinking* only 25% of teachers had an average rating of 3 or over, and only 1 teacher had an average rating of 3 or over for *teacher feedback*.

Figure 8.1: Distribution of teacher average ratings for the assessments of and responses to student understanding domain
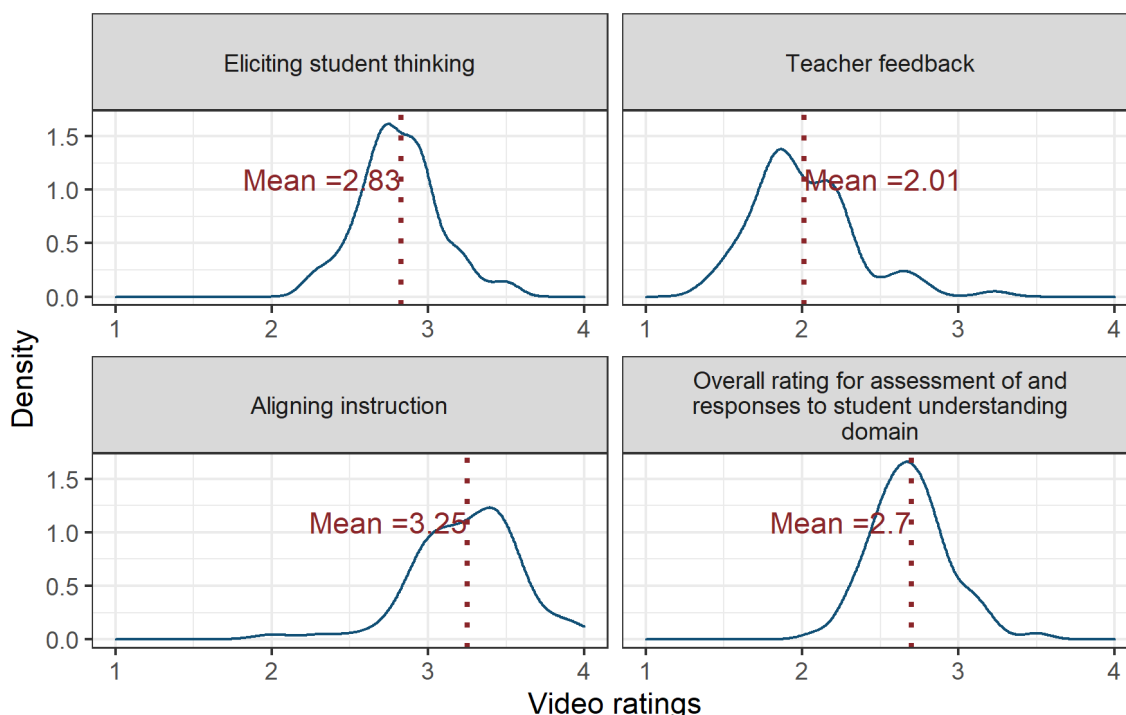
The *assessment of and responses to student learning* domain contains only one component on which the lesson artefacts were rated: *encourage student self-evaluation*.

The encouragement of student self-evaluation

Self-evaluation can help students become more aware of their level of understanding and better able to diagnose their own strengths and weaknesses to target study efforts towards improvement. Lesson artefacts were analysed for the extent to which students were asked to assess their own understanding of the content that was studied, or to reflect on their own learning of the content. Note that tests and quizzes, even if intended as a way for students to check their understanding, were not considered as self-evaluation unless they explicitly asked students to reflect on their understanding.

The average rating of *encourage student self-evaluation* was 1.1, where a rating of 1 indicates that students were not asked to assess or reflect on their understanding of a lesson's content. This overall average indicates that students were very rarely provided with opportunities for self-evaluation that involved explicit reflection on their understanding.

A rating of 2 was achieved if students were asked to broadly assess their understanding, such as assigning a general level to their understanding; whereas a 3 indicates that students were asked to assess their understanding in detail, which might be demonstrated by rating their understanding on specific skills, or by explaining what they learned in a particular lesson. Around a quarter of teachers asked students to broadly assess their understanding in at least one lesson, but only 7% of teachers asked their students to assess their understanding in detail at least once.

Did teachers with different backgrounds assess and respond to their students' understanding in their lessons differently?

There were no significant associations between the *assessment of and responses to student understanding* domain, or any of its components, and teachers' background characteristics. This includes teachers of different genders, with different levels of self-efficacy, or in relation to the number of years' experience a teacher had in teaching mathematics.

Did classes with different characteristics experience different assessment of and responses to their understanding in their lessons?

There were no significant associations between the overall average domain ratings for *assessment of and responses to student understanding* and the average class score for personal interest in mathematics with their previous teacher, the average class score for students' general self-efficacy in mathematics with their previous teacher, or the proportion of students with a first- or second-generation immigrant background within a class. There was significant negative association between the overall average domain ratings for *assessment of and responses to student understanding* and the proportion of students within classes who spoke English at home. There were significant positive associations between the overall average domain ratings for *assessment of and responses to student understanding* and the proportion of students within classes who were female, the average class pre-test score, and the average level of parental education.

There were significant differences between classes with lower average attainment on the pre-test and the teacher average rating for *eliciting student thinking*, with lower-attaining

students experiencing classrooms with lower teacher average ratings. This difference remained after controlling for other class characteristics. There were also significant differences between classes with higher average scores for students' home possessions, personal interest in mathematics with their previous teacher, or general self-efficacy in mathematics with their previous teacher, and the teacher average rating for *eliciting student thinking*. Classes with higher average scores for students' home possessions, personal interest in mathematics with their previous teacher, or general self-efficacy in mathematics with their previous teacher, experienced classrooms with higher teacher average ratings for *eliciting student thinking* than classes with lower average scores. These differences were no longer significant after controlling for the other class characteristics.

Summary

Teachers in the TALIS Video Study in England assessed and responded to students' thinking in a variety of ways, with an average domain rating of 2.70 on a 4-point scale. Most teachers frequently adapted their instruction based on students' level of understanding. Teachers generally asked questions or gave prompts that elicited detailed student responses and sometimes used feedback loops to understand students' thinking; however, students were very rarely provided with opportunities for self-evaluation that involved explicit reflection on their understanding. There were no significant associations between the components within the *assessment of and responses to student understanding* domain and teachers' background characteristics. There were significant positive associations between the overall average domain ratings for *assessment of and responses to student understanding* and the proportion of students within classes who were female, the average class pre-test score, the proportion of students who did not speak English at home, and the average level of parental education.

9. Opportunity to learn

Key findings

- The most common solution method was solving quadratic equations by factorising.

- The most common subtopic in the TALIS Video Study lessons, both as a major and a minor focus, was handling algebraic expressions, which included working with brackets and algebraic terms.

- Fewer than a fifth of lessons included the consideration of applying quadratic equations to real-life contexts.

- Almost all teachers reported content coverage which may have provided *opportunity to learn* in relation to algebraic operations. Just over half of teachers reported content coverage that may have facilitated OTL in relation to real-world applications of quadratic equations.

- Female students reported less *opportunity to learn* than male students during the topic of quadratic equations.

- Students' prior attainment was associated with the extent of *opportunity to learn* reported by students, with students in classes with higher average pre-test scores indicating a greater number of perceived *opportunity to learn* compared to students in classes with lower average pre-test scores.

- Students' perceived level of *opportunity to learn* through the topic of quadratic equations was positively associated with students' attainment on the post-test, their personal interest in mathematics, and their self-efficacy in mathematics for the topic of quadratic equations.

- Students' perceived level of opportunity to use quadratic functions was positively associated with students' attainment on the post-test, their personal interest in mathematics and general self-efficacy in mathematics with their current teacher.

- Students' perceived level of *opportunity to learn* reasoning with quadratic equations was positively associated with students' personal interest in mathematics and their general self-efficacy in mathematics with their current teacher.

Introduction to opportunity to learn

Opportunity to learn (OTL) focuses on content matter as it is taught and experienced by students[27]. In this chapter, the amount of time devoted to the teaching of quadratic equations, and to the different subtopics, is examined. These subtopics include different methods for solving quadratic equations, such as by factorising, using the quadratic formula, completing the square, or finding the roots on a graph. They also include working with quadratic expressions, using the binomial formula (e.g. $(x-3)^2$), considering the different cases of quadratic equations that depend on which coefficients are present or not (e.g. discriminating between expressions of the form ax^2+bx+c depending on values of a, b, and c), as well as graphing quadratic functions and applying quadratic equations to different contexts. In the TALIS Video Study, *OTL* is part of the *quality of subject matter* domain discussed in Chapter 6.

The four subtopics related to solution methods (solving quadratic equations by completing the square, factorising, using the quadratic formula, and via graphical representation) are discussed in the section concerning content coverage, which addresses the extent to which subject matter associated with the solution of quadratic equations is covered within the topic unit. The remaining subtopics are discussed in the section regarding content emphasis, which examines the prominence of different types of content throughout the series of unit lessons.

The *OTL* across subtopics which students perceived to have experienced during the topic unit, the *OTL* supported by teacher-reported content coverage in the Teacher Log[28], and the *OTL* supported by collected artefacts are discussed in consecutive sections. The relationships between the degree to which *OTL* was perceived by students and the extent to which such opportunities were supported by teachers and artefacts is then examined in the section on correlations between measures of *OTL*. Next the relationship between various student and classroom level factors and students' perceived *OTL* is discussed. Finally, the relationships between students' perceived *OTL* and student outcomes, specifically students' post-test scores, personal interest in mathematics with their current teacher, and general self-efficacy in mathematics with their current teacher, is discussed.

What was students' exposure to the topic content?

Content exposure refers to the amount of time teachers spent on the topic of quadratic equations. In this section, the number of lessons and overall time spent on the topic is considered. The proportion of lessons that include the content coverage subtopics,

---

[27] Burroughs et al. (2019), Klieme (2013), and Schmidt and Maier (2009)
[28] Record of content coverage and emphasis throughout the focal unit

followed by the proportion of lessons that include the content emphasis subtopics, is also discussed.

On average, lessons were 61 minutes long, with the shortest lesson being 30 minutes long and the longest lesson being 120 minutes long, according to the Teacher Logs. Four schools included double-length lessons. The average lesson length as recorded by the videos was 54 minutes. Teachers spent between 4 and 16 lessons on the topic of quadratic equations, with an average of 7 lessons. This means that on average teachers spent 7.6 hours on the topic of quadratic equations.

The most common solution method taught in the TALIS Video Study lessons was solving quadratic equations by factorising. Slightly more than a third of lessons focused specifically on solving quadratic equations by factorising, whilst slightly less than a third of the remaining lessons included this method as a minor focus. Solving quadratic equations by finding roots in a graphical representation appeared as the major focus for the fewest number of lessons, but around a quarter of lessons included this method as a minor focus. Table 9.1 shows the proportion of lessons that included each solution method as a major or minor focus, as indicated by teachers.

Table 9.1: Solution methods taught in the TALIS Video Study unit of work

Subtopic	Minor focus	Major focus	Total proportion
Solving quadratic equations by completing the square	11%	19%	30%
Solving quadratic equations by factorising	27%	36%	64%
Solving quadratic equations by quadratic formula	11%	20%	30%
Solving quadratic equations by finding roots in a graphical representation	24%	14%	38%

Source: TALIS Video Study teacher data file

The most common subtopic in the TALIS Video Study lessons, both as a major and a minor focus, was handling algebraic expressions, which included working with brackets and algebraic terms. This subtopic includes many of the skills students would need in order to use the solution methods discussed above. Fewer than a fifth of lessons included the consideration of applying quadratic equations to real life contexts. Table 9.2 shows the proportion of lessons that include each subtopic as a major or minor focus, as indicated by teachers. It is possible that the inclusion of binomial formulae is under-reported as this terminology is not often used with students studying up to GCSE level.

Table 9.2: Proportion of lesson focusing on each subtopic as a major or minor focus

Subtopic	Minor focus	Major focus	Total proportion
Handling algebraic expressions	43%	30%	73%
Binomial formulae	14%	4%	18%
Introducing one form of a quadratic equation	32%	18%	50%
Discuss different cases depending on the values of the coefficients	28%	14%	42%
Quadratic functions	17%	11%	29%
Real-world applications	12%	6%	17%

Source: TALIS Video Study teacher data file

What content was covered during the topic?

The most common solution method taught in the TALIS Video Study lessons was solving quadratic equations by factorising. This was the case across all lessons at different stages of teaching within the topic unit, with over half of participating teachers focusing on solving quadratic equations by factorising to some extent (minor or major focus) by their final lesson in the unit. Solving quadratic equations by finding roots in a graphical representation was more common during the first lesson and towards the second half of the unit than it was in the lessons in between. More than a quarter of teachers reported focusing on solving quadratic equations by finding roots in a graphical representation to some extent during the initial unit lesson and, whilst the proportion of lessons that included some focus on solution through graphical representation diminished in the immediately ensuing lessons, a similar number of teachers reported focusing on solution through graphical representation in the latter half of their unit lessons. Figure 9.1 shows how the emphasis on different content changed over the sequence of lessons within the unit on quadratic equations, where the average is weighted so that a rating of 1 represents a major focus and a rating of 0.5 represents a minor focus.

Figure 9.1: Content coverage over the duration of the topic unit
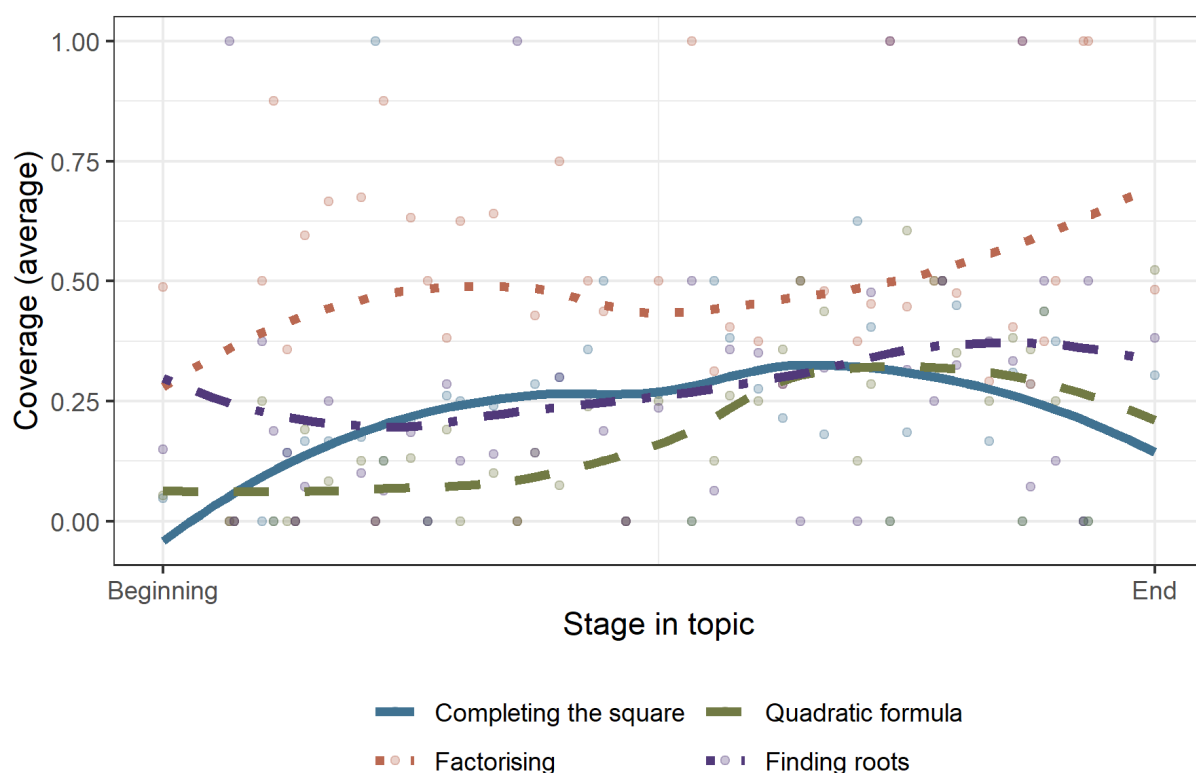
Source: <u>TALIS Video Study teacher data file</u>

What content was emphasised during the topic?

Six of the subtopics relate to content emphasis within the topic of quadratic equations. Content concerning the handling of expressions was emphasised in almost three-quarters of the lessons recorded in the Teacher Log, either as a major or minor focus. Handling expressions was also the most common subtopic included at all stages of the topic, though the inclusion of handling expressions decreased over the course of the topic unit, as illustrated in Figure 9.2. Discussing different cases of ax^2+bx+c depending on values of a, b, and c (e.g., which strategy is best for solving different complete and incomplete quadratic equations) was included slightly more as the topic progressed, as was applying quadratic equations to real-life situations.

Figure 9.2: Content emphasis over the duration of the topic
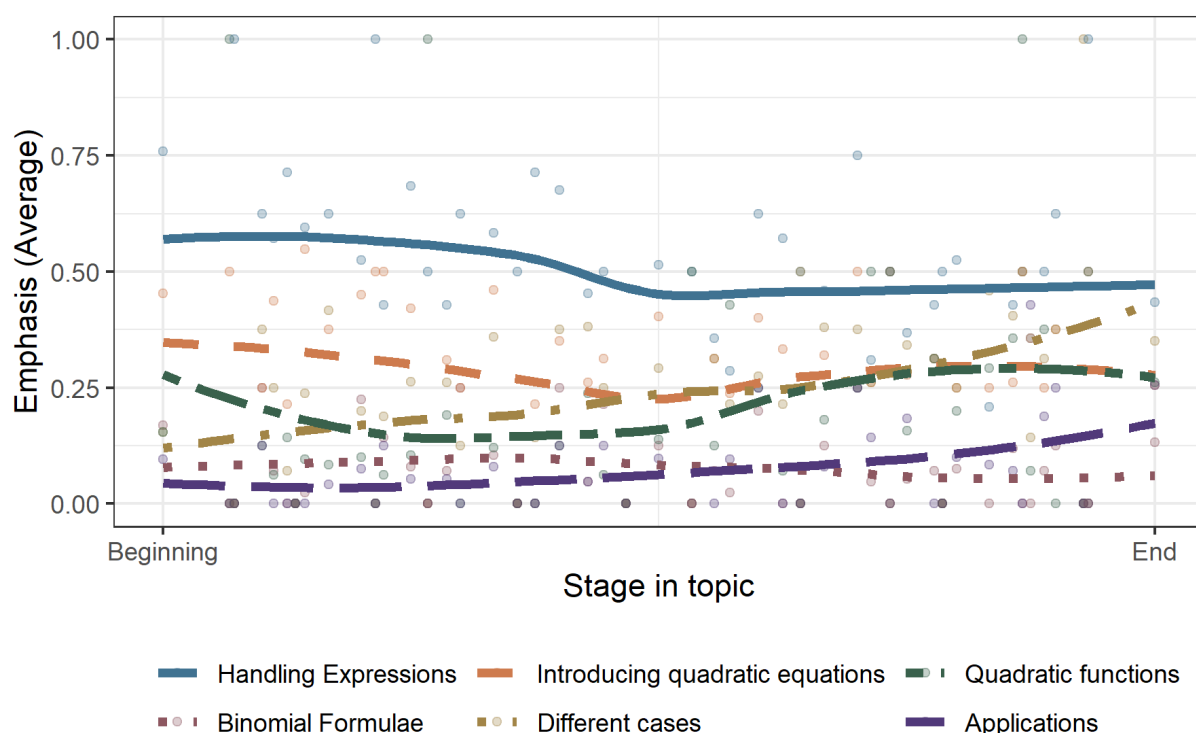
Source: <u>TALIS Video Study teacher data file</u>

These subtopics are grouped depending upon their role in the teaching of mathematics. *OTL algebra* consists of the subtopics handling expressions, binomial formula and the three algebraic methods for solving quadratic equations (factorising, using the formula, and completing the square). *OTL apply* refers to the application of quadratic equations to real-life contexts, and *OTL reasoning* refers to the discussion of different cases of ax^2+bx+c depending on values of *a*, *b*, and *c*. *OTL functions* includes exploring quadratic functions and solving quadratic equations by finding roots in a graphical representation.

What were students' perceptions of their opportunity to learn?

Following completion of the focal unit, students indicated their perceptions of *OTL* experienced throughout the unit, identifying whether they had experienced learning opportunities to develop specific content knowledge across four relevant content subtopics. These constituent subtopics concerned opportunities to use quadratic functions (*OTL functions*), algebraic operations (*OTL algebra*), reasoning about different types of quadratic equations (*OTL reasoning*), and opportunities to apply quadratic equations to real-world contexts (*OTL apply*).

Students indicated experiencing the greatest level of *OTL* amongst subtopics in relation to algebraic operations, on average indicating they had experienced 82% of the *OTL algebra* compared to just 23% of the *OTL apply* as shown in Table 9.3. Notably, whilst

post-unit questionnaires posed a greater extent of prospective *OTL algebra* (five potential *OTL*) than those related to other content subtopics (two potential *OTLs* for each subtopic), students nonetheless on average indicated experiencing a greater proportion of *OTL algebra* than those concerning other *OTL*. Students indicated experiencing the lowest level of *OTL apply*, with only one third of students indicating that they had perceived any *OTL* concerning this subtopic. This is in contrast to the 100% of students who had experienced at least one opportunity to learn algebraic operations.

Table 9.3: Student perceptions of OTL experienced across subtopics

Subtopic	Average Opportunities to Learn	Average proportion of opportunities	Proportion of students
Quadratic functions	1.35	67%	89%
Algebraic operations	4.11	82%	100%
Reasoning	1.44	72%	87%
Application to real-world contexts	0.47	23%	36%

Source: TALIS Video Study student data file

What opportunity to learn did teachers report providing?

Teacher provision of *OTL* across these same content subtopics, opportunities to use quadratic functions, to learn algebraic operations, to learn reasoning about different types of quadratic equations, and to apply quadratic equations to real-world contexts, was derived from Teacher Log records of content coverage and emphasis throughout the focal unit. As such, teacher ratings for *OTL* reflect the extent to which teacher-reported content coverage may have enabled student *OTL* across content subtopics over the course of the focal unit.

Similar to trends apparent in students' perceived *OTL* across subtopics, teacher content coverage enabled, on average, the greatest level of student *OTL* amongst subtopics in relation to algebraic operations, and the lowest level of *OTL* in relation to the application of quadratic equations to real-world contexts, as shown in Table 9.4. Indeed, whereas virtually all teachers reported content coverage which may have enabled *OTL* in relation to algebraic operations, just over half of teachers reported content coverage which may have supported *OTL* in relation to real-world applications of quadratic equations.

Table 9.4: Teacher provision of OTL across subtopics

Subtopic	Average rating	Proportion of teachers
Quadratic functions	1.71	86%
Algebraic operations	2.42	99%
Reasoning	2.10	81%
Application to real-world contexts	0.85	60%

Source: TALIS Video Study student data file

What support of opportunity to learn was there in the artefacts?

Artefact ratings for the support of *OTL* across these same content subtopics – opportunities to use quadratic functions, to learn algebraic operations, to learn reasoning about different types of quadratic equations, and to apply quadratic equations to real-world contexts – were derived based on artefact ratings of subtopic content coverage. Artefact *OTL* ratings therefore represent the extent to which collected artefacts were determined to support student *OTL* across subtopics.

Artefact content coverage, congruent with that of teachers, supported the greatest extent of student *OTL algebra* on average, as shown in Table 9.5. In the artefacts, subtopic coverage rated each lesson's artefacts as including the subtopic (a rating of 1) or not including the subtopic (a rating of 0). Evidence from the artefacts suggested that *OTL reasoning* received the least support (with 45% of artefacts rated as covering this content), in contrast to the evidence from Teacher Logs (which provided the least support for *OTL apply*). However, this diminished support for *OTL reasoning* in the artefacts may, in part, be attributable to the frequently spontaneous, conversational nature of teacher modelling and actions which enable *OTL reasoning*, which may not be adequately represented in static artefacts.

Table 9.5: Artefact support of teacher provision of OTL across subtopics

Subtopic	Average rating	Proportion of artefacts
Quadratic functions	0.26	76%
Algebraic operations	0.47	100%
Reasoning	0.06	46%
Application to real-world contexts	0.18	59%

Source: TALIS Video Study student data file

What associations were there between the different measures of opportunity to learn?

The correlation or association between students' perceived *OTL* and the extent to which teacher-reported and artefact-rated content coverage may have enabled such *OTL* is weak to negligible across all subtopics.

Such a lack of association between *OTL* perceived by students and those ostensibly facilitated by teachers or teaching materials may indicate a disparity or incongruity between teachers' perceptions of content coverage and emphasis across subtopics and that which is realised or experienced by students. Alternatively, however, such an apparent discrepancy between *OTL* facilitated and experienced may be a consequence of issues in the conceptualisation and subsequent measurement of such *OTL*.

What were the relationships between OTL and student and classroom factors?

Regression models[29] were used to investigate the relationship between the extent of *OTL* perceived by students and various student and classroom-level factors. Further detail can be found in the England Technical Report[30].

Students from native and non-native (first- and second-generation immigrant) backgrounds did not differ in the number of perceived *OTL* indicated. Although student home language did not influence overall perceived *OTL*, students whose home language was English had a greater likelihood of indicating a higher number of *OTL reasoning* than their peers who had a home language other than English.

---

[29] Regression models relate an outcome (such as students' attainment, self-efficacy or personal interest in mathematics) to a series of explanatory variables.
[30] McCann, Riggall, Ingram, and Lindorff (2020)

Students' perceived *OTL* did not, on average, differ to a meaningful extent based on parents' or carers' level of educational attainment. The number of students' home possessions was identified as positively associated with perceived *OTL*, whereby students with more home possessions typically indicated experiencing a higher level of *OTL*, both overall and in relation to the algebraic operations subtopic specifically.

The extent of *OTL* that students indicated overall, that is, across all subtopics, was found to be positively associated with students' pre-test scores and class average pre-test scores, such that students attaining higher pre-test scores and students in classes that achieved higher average pre-test scores tended to indicate greater levels of perceived *OTL*. Notably, however, class average score for personal interest in mathematics was negatively associated with *OTL*, intimating that students in classes with, on average, a greater interest in mathematics were likely to indicate a lower level of perceived *OTL*.

The average extent of *OTL* perceived by students within classes was positively associated with class size, such that classes with a greater number of students typically demonstrated a greater average level of *OTL* perceived by students. Classes for which lesson artefacts demonstrated greater *OTL* were associated with higher average student-perceived *OTL*. The Teacher Log reports of *OTL* were not found to be meaningfully associated with the average of that perceived by students within classes except in relation to the subtopic of quadratic functions, wherein greater teacher-reported provision of *OTL* was found to be associated with greater average student-perceived *OTL*.

What were the relationships between OTL and student outcomes?

Regression models were also applied in seeking to identify any relationships between students' perceived *OTL* throughout the focal unit and student outcomes, including student attainment on the post-test, students' personal interest in mathematics with their current teacher, and students' general self-efficacy with their current teacher. The extent of students' overall perceived *OTL* was identified as positively associated with all outcomes, indicating that a greater extent of *OTL* experienced by students was associated with higher attainment on the post-test, greater personal interest in mathematics with their current teacher, and higher general self-efficacy in mathematics with their current teacher.

Students' perceived *OTL functions* was found to be positively associated with all outcomes, whilst perceived *OTL reasoning* was positively associated with students' personal interest in mathematics with the current teacher and students' general self-efficacy with their current mathematics teacher. Perceived *OTL algebra* was positively associated with students' personal interest in mathematics with their previous teacher. However, student perceived *OTL apply* did not exhibit a significant association with any

student outcomes. Class average perceived *OTL reasoning* was identified as positively associated with both students' post-test score and post-unit interest in mathematics.

The Teacher Logs for *OTL reasoning* were negatively associated with students' post-test scores, whilst those for *OTL algebra* were positively associated with students' post-test scores. Whilst artefact rating for the support of *OTL reasoning* was negatively associated with student post-unit general self-efficacy, artefact rating for the support of *OTL algebra* was positively associated with this same outcome measure.

Summary

The most common solution method taught for solving quadratic equations was by factorising. The most common subtopic was handling algebraic expressions. Fewer than a fifth of lessons included the consideration of applying quadratic equations to real-life contexts. Students' prior attainment was associated with the extent of *OTL* reported by students, with students in classes with higher average pre-test scores indicating a greater extent of perceived *OTL* compared to students in classes with lower average pre-test scores. Students' perceived *OTL* throughout the topic of quadratic equations was also positively associated with their attainment on the post-test, their personal interest in mathematics, and their self-efficacy in mathematics with their current teacher.

10. Relationships between student characteristics, teaching practices and student outcomes

Key findings

- Students with higher pre-test scores tended to have higher attainment on the post-test, with pre-test scores explaining half the variation in post-test scores.

- After accounting for individual pre-test scores and student characteristics, class average pre-test scores had a significant negative relationship with individual post-test scores for students in the lower three quartiles of class average pre-test score, meaning that students in lower-attaining classes had lower post-test scores even after accounting for their individual pre-test scores.

- Students who were provided with more opportunity to learn quadratic functions tended to have higher post-test scores.

- On average, students who had higher general self-efficacy in mathematics with their previous teacher also had higher general self-efficacy in mathematics with their current teacher; previous self-efficacy scores explained about one-fifth of the variation in students' current self-efficacy scores.

- After accounting for self-efficacy in mathematics with their previous teacher, as well as student characteristics, students in classes with a higher class average parental education had lower general self-efficacy in mathematics scores with their current teacher.

- Students who were provided with more opportunity to learn reasoning about different types of quadratic equations tended to have higher general self-efficacy in mathematics with their current teacher.

- On average, students who had a higher personal interest in mathematics with their previous teacher also had higher personal interest in mathematics with their current teacher, with their previous interest in mathematics explaining approximately one-eighth of the variation in students' current interest.

- After accounting for personal interest in mathematics with their previous teacher, students who were first- or second-generation immigrants, were in Year 8, or had more home possessions, had higher personal interest in mathematics with their current teacher.

- Students who experienced more opportunity to learn algebraic operations, opportunities to use quadratic functions, and opportunity to learn reasoning about

different types of quadratic equations tended to have a higher personal interest in mathematics with their current teacher.

- Before accounting for student and class characteristics, there were no significant relationships between teaching practice (as investigated through classroom management, social-emotional support, discourse and assessment, or mathematics instruction) and the student outcomes of post-test attainment, general self-efficacy in mathematics with the current teacher, and personal interest in mathematics with the current teacher.

- After accounting for student and class characteristics, there exists a significant interaction between classroom management and post-test score in the 4-domain models for students in the second- lowest quartile of class average pre-test scores, in that higher classroom management average ratings corresponded to increased post-test scores, whereas for other quartiles the trend was either reversed or flat.

- After accounting for student and class characteristics, there are significant interactions in the 4-domain model for students in the lowest quarter of class average general self-efficacy in mathematics with the previous teacher, where higher classroom management or higher mathematics instruction average ratings corresponded to higher general self-efficacy in mathematics with the current teacher.

- For students in classes with teachers who had the lowest average ratings for discourse and assessment, personal interest in mathematics with the current teacher was higher for students in classes with the lowest average personal interest in mathematics with the previous teacher than for students in classes with higher average personal interest in mathematics with the previous teacher, while this relationship was reversed for students in classes with teachers who had higher average ratings for discourse and assessment.

Introduction

This chapter builds on the results presented in previous chapters, to examine the relationships between teaching practices, student and class characteristics, and student outcomes in further detail using regression models. Such models allow for the investigation of the relationship(s) between the outcomes that were the focus of the TALIS Video Study (student attainment, general self-efficacy in mathematics, and personal interest in mathematics) and one or more explanatory variables relevant to student and class characteristics as well as the measures of teaching practices

discussed in Chapters 3 to 8. Further information on the modelling approach can be found in the accompanying Technical Report[31].

The chapter is separated into three main sections. The first focuses on how each student outcome (attainment on the post-test, general self-efficacy in mathematics with the current teacher, and personal interest in mathematics with the current teacher) varies by student characteristics and by class characteristics. The second section focuses on how teaching practices relate to each of the same three student outcomes, and how these relationships vary by student and school characteristics. In this second section, teaching practices are defined in terms of four domains based on the video component average ratings for each teacher.

1. *Classroom management* (as defined as in Chapter 3)

2. *Social-emotional support* (as defined in Chapter 4)

3. *Discourse and assessment*: a combination of *discourse* and *assessment of and responses to student understanding* (as defined in Chapters 5 and 8)

4. *Mathematics instruction*: a combination of the *quality of subject matter* and *student cognitive engagement* domains (as defined in Chapters 6 and 7)

The *discourse and assessment* domain is configured in this way on the basis that these are teaching practices that are widely used by teachers in a range of curriculum areas, while the *mathematics instruction* domain includes measures of teaching practice that are more specific to the teaching of mathematics. Results from alternative 3-domain models, in which domains include *classroom management*, *social-emotional support*, and *instruction* (*discourse, assessment of and responses to student learning, quality of subject matter, and student cognitive engagement*) are also discussed in the second section.

The third section examines the relationships between the different measures of teaching practices. The focus is first on the relationships between teaching practices as measured within a particular type of rating (video components, video indicators, or artefact components), and then on the relationships between teaching practices as measured across different types of ratings (for example, relationships between video component and video indicator ratings of teaching practices within the same domain).

Within the first two sections a step-by-step approach is used to build up the models, with relationships examined in successive groups of variables. This approach yields results that show relationships between each group of variables and the different outcomes, over and above the relationships involving variables that were in the model at the previous step. So, for example, in models for which the outcome of interest is students' self-

---

[31] McCann, Riggall, Ingram, and Lindorff (2020)

efficacy in mathematics with their current teacher, and if students' self-efficacy in mathematics with their previous teacher was included as a predictor at one step, when students' gender and socio-economic status are included at the next step, any significant relationship between socio-economic status or gender and self-efficacy in mathematics with the current teacher is *over and above* or *independent of* the relationship between self-efficacy in mathematics with the previous teacher and that with the current teacher. The regression models thus tease out the "net" effect of particular variables and groups of variables, controlling for the effects of other variables in the models. This regression approach enables establishment of the direction, strength and statistical significance of different variables in accounting for variation in the different student outcomes.

Groups of variables were entered as follows:

1. Teaching practices based on video domains (for models concerning teaching practices only). Results are reported based on the 4-domain and 3-domain models, with domains defined as outlined above. Additional 6-domain model results are given in the accompanying Technical Report[32] and data tables.

2. Pre-unit measures corresponding to the relevant post-unit outcomes (student attainment on the post-test, general self-efficacy in mathematics, or personal interest in mathematics). That is, when the outcome of interest was students' general self-efficacy in mathematics with their current teacher, then students' general self-efficacy in mathematics with their previous teacher was entered.

3. Student characteristics as reported on the student questionnaires, including gender, year group, first- or second-generation immigration background, home possessions, parental education (GCSE or lower, A-level, non-university tertiary, university), and language spoken at home (English or not).

4. Class characteristics including the pre-unit class average corresponding to the relevant post-unit outcome (student attainment on the post-test, general self-efficacy in mathematics or personal interest in mathematics) split into Highest(1)/High(2)/Low(3)/Lowest(4) quartiles, proportion of the class who were female, proportion of the class who were first- or second-generation immigrants, class average home possessions, class average parental education, and class size (as reported by the teacher).

5. *Opportunity to learn* (based on items from the student questionnaires) for subtopics including opportunities to use quadratic functions (*OTL functions*), opportunities to learn algebraic operations (*OTL algebra*), opportunities to learn reasoning about different types of quadratic equations (*OTL reasoning*), and opportunities to apply quadratic equations to real-world contexts (*OTL apply*).

---

[32] McCann, Riggall, Ingram, and Lindorff (2020)

6. Interactions between student and/or class characteristics (e.g. class average pre-test score and individual pre-test score). Interactions allow for the relationship of one variable to the outcome to vary according to another variable.

Relationships between teaching practices are examined using correlations between the video component ratings, video indicator ratings, and artefact ratings.

Detailed definitions of each variable used in this chapter are not presented here as these have been discussed at length in previous chapters, except where variables have been modified specifically for use in the modelling process.

What were the relationships between student and class characteristics and student outcomes?

This section reports results from regression models that use attainment on the post-test, general self-efficacy in mathematics with their current teacher, and personal interest in mathematics as the outcome measures, relating these to student and class characteristics based on the student pre-questionnaires. Full model results are available in the data tables of the accompanying Technical Report[33].

Relationships between student and class characteristics and student attainment on the post-test

There was a significant positive relationship between students' pre-test scores and their post-test scores, meaning that students with higher scores on the pre-test tended to score higher on the post-test, and vice versa, as shown in Figure 10.1. Pre-test score explained half of the variation in post-test scores, showing the strength of this relationship. All subsequent variables, including socio-economic status and language spoken at home, by contrast, explained very little of the variation in post-test scores.

---

[33] McCann, Riggall, Ingram, and Lindorff (2020)

**Figure 10.1: Relationship between attainment on the pre-test
and attainmnet on the post-test**
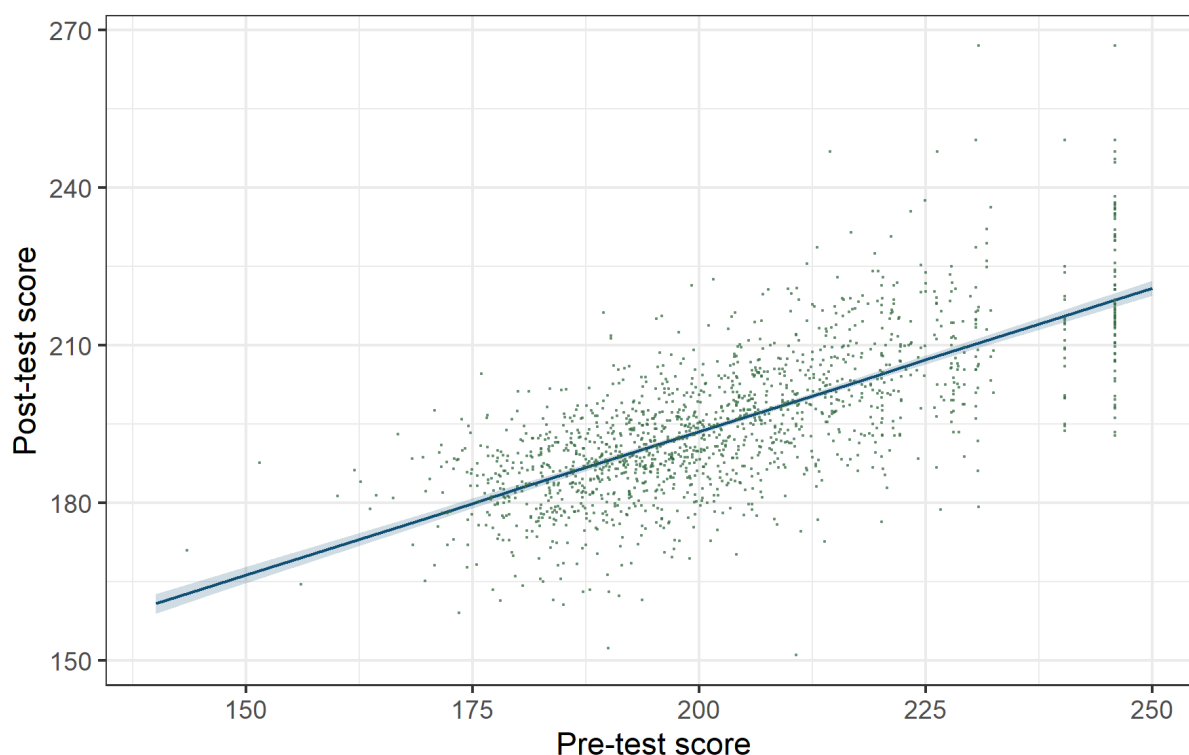
After pre-test scores had been accounted for, no student characteristics showed any significant relationships with the post-test scores except for the home possessions measure, which was positively and significantly related (that is, students with higher scores on the home possessions scale tended to have higher scores on the post-test). The lack of significant relationships between other student characteristics and attainment on the post-test might seem counter-intuitive, as many of these characteristics have been widely established to have relationships to attainment in the academic literature[34]. However, in this instance, because the duration between pre-test and post-test was only the length of one unit of taught content (rather than, for example, a full academic year or more), it is important to be cautious in interpreting these results. Here, the pre-test score is likely to be already accounting for relationships that exist between student demographic characteristics and attainment in general, and the results of this analysis show that these demographic characteristics are not significantly related to students' progress within the single taught unit of interest beyond any pre-existing achievement gaps.

---

[34] Lenkeit, J., Caro, D.H., and Strand, S. (2015) or Strand, S. (2014)

Once both the pre-test scores and student characteristics had been accounted for, most class characteristics (for example, class proportion of students who were female, class average home possessions, class average parental education score, class proportion of students who were first- or second-generation immigrants, and class size) showed no significant relationship to attainment on the post-test. The only exception to this was the class average pre-test score, illustrated in Figure 10.3. Being in a class with average pre-test scores in any of the lower three quartiles had a significant negative relationship to students' individual post-test scores, over and above the relationship between their individual pre- and post-test scores. In other words, students taught in classes with lower pre-test performance on average tended to have lower scores on the post-test even after taking into consideration their individual pre-test attainment. Individual home possessions was no longer significantly related to the post-test score once class characteristics were included, which suggests that class characteristics may be associated with the relationship between home possessions and students' attainment on the post-test.

Figure 10.2: Relationship between class prior-attainment and student attainment on the post-test
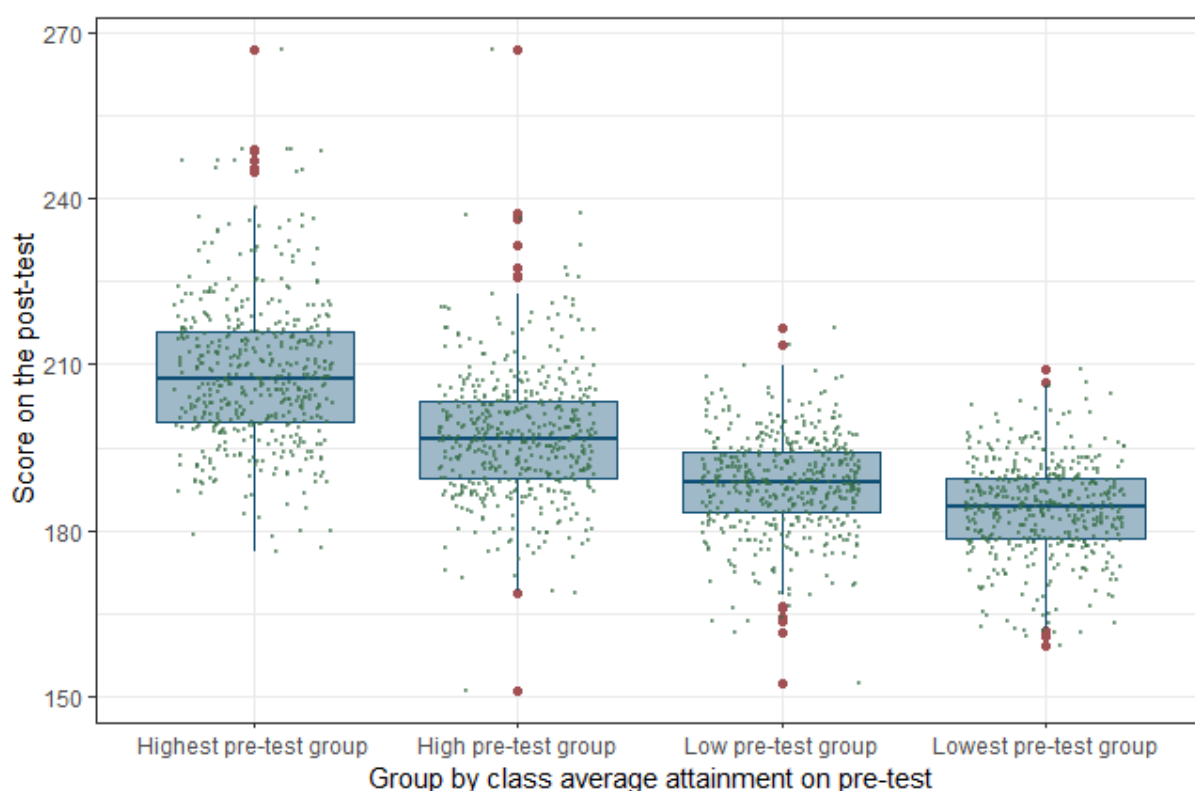
Opportunity to learn (*OTL*) was significantly related to student attainment on the post-test, but only for some subtopics. There were no significant relationships between *OTL algebra*, *OTL reasoning*, or *OTL apply* and post-test attainment. There was a significant positive relationship between *OTL functions* and post-test attainment; see Chapter 9 for

descriptive information about these subtopics and the OECD Technical Report[35] for further information about how each of these *OTL* scores was measured.

There was a significant positive interaction between class average pre-test score and individual pre-test score for the lowest and second-lowest quartiles of class average pre-test attainment. Furthermore, once this interaction was accounted for, only the lowest quartile of class average pre-test attainment was significantly related to post-test attainment. This significant interaction is illustrated in Figure 10.3. For students with a lower individual pre-test score (towards the left side of the graph), the average level of pre-test attainment in their class is less strongly related to their post-test attainment. For students with higher individual pre-test scores, the average level of pre-test attainment in their class "makes more of a difference" giving a boost to their post-test attainment. For example, for the lowest-attaining students at pre-test, there is little difference in their post-test scores according to their class's average pre-test scores (which can be seen on the left side of the graph in Figure 10.3). For the highest-attaining students, there is more of a pronounced difference in post-test attainment according to the average pre-test attainment in the class (which can be seen on the right side of the same graph); these students who were higher-attaining at pre-test tended to have higher post-test scores the higher their class's average pre-test attainment.

Figure 10.3: Interaction between class average pre-test attainment and individual pre-test attainment for student attainment on the post-test
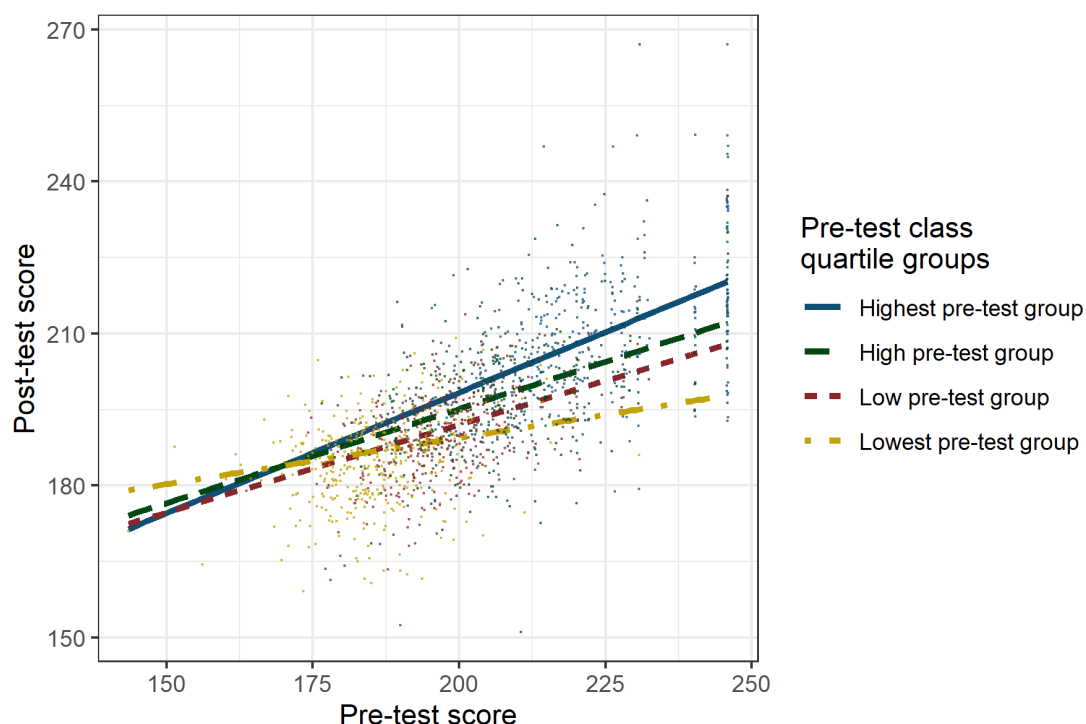
Source: TALIS Video Study student data file

---

[35] OECD (2020b)

Relationships between student and class characteristics and student general self-efficacy in mathematics with their current teacher

There was a significant positive relationship between students' general self-efficacy in mathematics with their previous teacher and their general self-efficacy in mathematics with their current teacher, meaning that students with higher self-efficacy with their previous teacher tended to have higher self-efficacy with their current teacher and vice versa. General self-efficacy with students' previous teacher explained 22% (or about a fifth) of the variation in general self-efficacy with their current teacher. All subsequent variables included in the models explained very little of this variation.

After general self-efficacy in mathematics with the previous teacher had been accounted for, most student characteristics did not show any significant relationships with general self-efficacy with the current teacher. There were two exceptions: first- and second-generation immigrant status was significantly and positively related to general self-efficacy in mathematics with the current teacher, as was non-university tertiary parental education (compared to parental education at GCSE level or lower). As noted above for student attainment on the post-test, the lack of significant relationships between other student demographic characteristics and general self-efficacy in mathematics with the current teacher may mean that pre-existing relationships between student characteristics and general self-efficacy were already reflected in the relationship between general self-efficacy in mathematics with the previous teacher and that with the current teacher.

Once both general self-efficacy with the previous teacher and student demographic characteristics had been accounted for, most class characteristics (for example, class average self-efficacy with the previous teacher, class proportion of students who were female, class average home possessions, class average parental education score, class proportion of students who were first- or second-generation immigrants, and class size) showed no significant relationships to general self-efficacy in mathematics with the current teacher. The only exception to this was the class average parental education, which was significantly and negatively related. This suggests that higher average parental education was associated with lower general self-efficacy in mathematics with the current teacher. Having more home possessions was also significantly and positively related to general self-efficacy with the current teacher once class characteristics had been accounted for.

OTL was significantly and positively related to general self-efficacy in mathematics with the current teacher, but only for some subtopics. There were no significant relationships between *OTL algebra* or *OTL apply* and general self-efficacy in mathematics with the current teacher. There were significant positive relationships between *OTL functions* as well as *OTL reasoning* and general self-efficacy in mathematics with the current teacher, meaning that greater (student-reported) opportunities to learn within these subtopics were associated with higher general self-efficacy in mathematics with the current teacher.

There was no significant interaction between individual general self-efficacy in mathematics with the previous teacher and the class average level of this same measure, nor any significant interaction between class average pre-test score and individual home possessions. However, there was a significant negative interaction between class average parental education and year group, specifically for Year 8 and Year 11 (compared to Year 10). For students taught in classes with the lowest average parental education, general self-efficacy in mathematics tended to be higher for students in Years 8 and 11 than for students in Years 9 and 10. For students in classes with the highest parental education, the reverse was found – that is, students in Years 8 and 11 tended to have lower general self-efficacy in mathematics than those in Years 9 and 10. This may relate to the fact that there were greater numbers of Year 9 and Year 10 students in the sample for this study, and the Year 9 and 10 groups appeared to be more diverse than the Year 8 and 11 groups (see Chapter 2).

Relationships between student and class characteristics and student personal interest in mathematics

There was a significant positive relationship between students' personal interest in mathematics with their previous teacher and their personal interest in mathematics with their current teacher, meaning that students with higher personal interest in mathematics with their previous teacher tended to have higher personal interest in mathematics with their current teacher and vice versa. Personal interest in mathematics with students' previous teacher explained 12% of the variation in personal interest in mathematics with their current teacher. All subsequent variables included in the models explained very little of the variation in personal interest in mathematics with their current teacher.

After personal interest in mathematics with the previous teacher had been accounted for, there were significant positive relationships between being in Year 8 (compared to Year 10; bearing in mind that there were very few Year 8 classes so it is important to interpret this result with caution), having more home possessions, and being a first- or second-generation immigrant, and personal interest in mathematics with the current teacher. There were no other significant relationships between student demographic characteristics and this outcome measure. As noted above for student attainment on the post-test and general self-efficacy, the lack of significant relationships between other student demographic characteristics and personal interest in mathematics with the current teacher may mean that pre-existing relationships between student characteristics and personal interest in mathematics were already reflected in the relationship between personal interest in mathematics with the previous teacher and that with the current teacher.

Once both personal interest in mathematics with the previous teacher and student demographic characteristics had been accounted for, most class characteristics (for example, class proportion of students who were female, class proportion of students who

were first- or second-generation immigrants, class average home possessions, class average parental education score) had no significant relationships to personal interest in mathematics with the current teacher. The only exceptions to this were class average personal interest in mathematics with the previous teacher, which was significantly positively related to students' personal interest in mathematics with the current teacher, but only for the lowest two quartiles of class personal interest with the previous teacher; and class size, which had a positive relationship to personal interest in mathematics with the current teacher, but this relationship was very small (this should be interpreted carefully as it corresponds to a very small difference in post-test score). Students taught in classes with lower average personal interest in mathematics with their previous teacher(s) tended to have higher personal interest in mathematics with their current teacher than those in classes with the highest average personal interest with their previous teacher(s).

OTL was significantly and positively related to general self-efficacy in mathematics with the current teacher for all subtopics except for *OTL apply*. In other words, students who reported greater *OTL algebra*, *OTL functions* and *OTL reasoning* tended to have higher personal interest in mathematics with their current teacher.

There was no significant interaction between class average personal interest in mathematics with the previous teacher and individual personal interest in mathematics with the previous teacher.

What were the relationships between teaching practices and student outcomes?

This section reports results from regression models that use attainment on the post-test, general self-efficacy in mathematics with their current teacher, and personal interest in mathematics as the outcome measures, relating these to teaching practices discussed in Chapters 3 to 8. Full details of models can be found in the England Technical Report[36] and accompanying data tables. This section further examines how these relationships vary by student and class characteristics based on student pre-questionnaires.

Relationships between teaching practices and student attainment on the post-test

In models with student attainment on the post-test as the outcome, there were no significant relationships between *classroom management*, *social-emotional support*, *discourse and assessment*, or *mathematics instruction* and student attainment on the post-test before accounting for student and school characteristics. The same was the case for *classroom management*, *social-emotional support* and *instruction* in the 3-

---

[36] McCann, Riggall, Ingram, and Lindorff (2020)

domain model. These results come from modelling teaching practices together rather than separately, which explains some differences between them and the OECD Policy Report's[37] findings from models which only considered teaching practices (*classroom practice*, *social-emotional support* and *instruction*) separately.

How do relationships between teaching practices and student attainment on the post-test vary by student and class characteristics?

Still using student attainment on the post-test as the outcome, after accounting for student and class characteristics as well as *OTL*, there was a significant interaction between the class average pre-test attainment and teaching practice, but only in relation to classroom management, and only for students taught in classes in the second-lowest quartile of average pre-test attainment. This significant interaction is illustrated in Figure 10.4. As the graph shows, this meant that class average pre-test attainment "made more of a difference" to students in classes with teachers who had lower ratings for classroom management. No such interaction was significant for the 3-domain model.

Figure 10.4: Interaction between class average attainment on the post-test and teacher classroom management

Source: TALIS Video Study student data file

---

[37] OECD (2020a)

Relationships between student attainment on the pre-test and teaching practices

Students' attainment on the pre-test was not significantly associated with teacher combined average ratings for the instructional domains where instructional domains were treated as the outcome variable, nor for the discourse and assessment or mathematics instruction domains. The OECD Policy Report[38] found that classes with an above-median average pre-test score and classes with a below-median average pre-test score had significantly different average teacher ratings across instructional domains. These findings may appear contradictory; however, testing differences between classes that were above or below the overall median (where some classes might be just slightly above or below the median) loses the detail necessary to understand how those relationships manifest in the context of the full range of variation between students within classrooms.

Relationships between teaching practices and general self-efficacy in mathematics

In models with general self-efficacy in mathematics with the current teacher as the outcome, there were no significant relationships between *classroom management*, *social-emotional support*, *discourse and assessment*, or *mathematics instruction* and general self-efficacy in mathematics with the current teacher before accounting for student and class characteristics. The same was found for *classroom management*, *social-emotional support* and *instruction* in the 3-domain model.
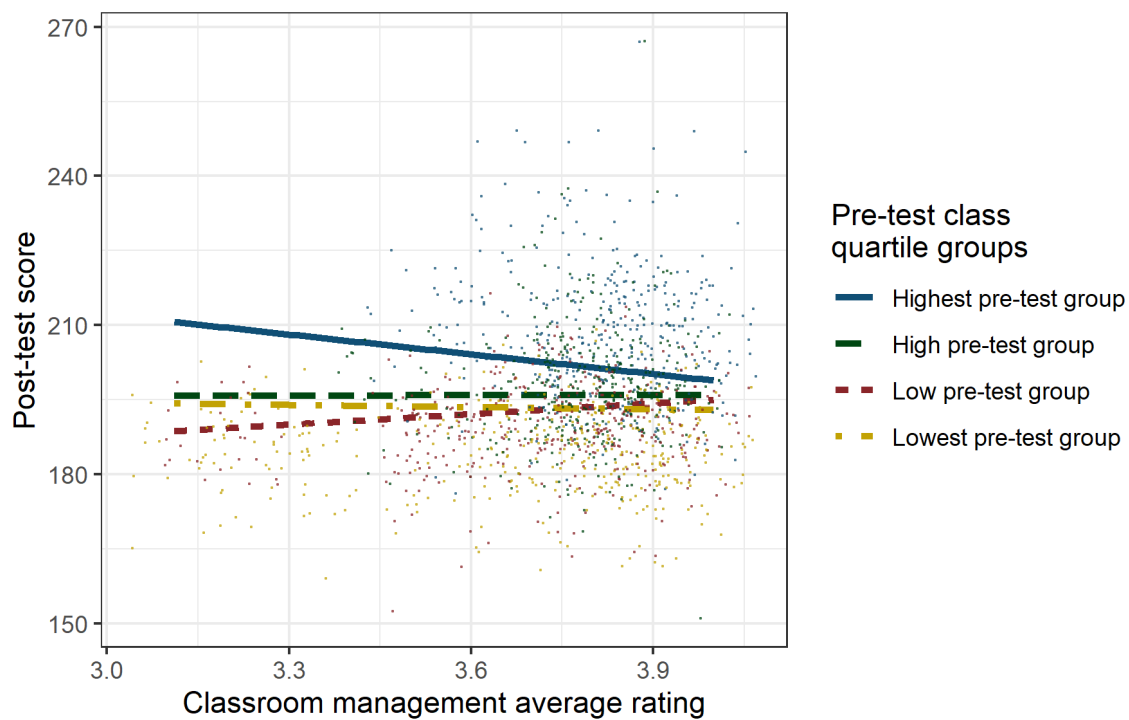
How do relationships between teaching practices and general self-efficacy in mathematics vary by student and class characteristics?

Still using general self-efficacy in mathematics with the current teacher as the outcome, after accounting for student and class characteristics and *OTL*, there was a significant interaction between the class average general self-efficacy in mathematics with the previous teacher and teaching practice, but only in relation to classroom management and mathematics instruction, and only for students taught in classes in the lowest quartile of average general self-efficacy in mathematics with the previous teacher, within the 4-domain model. This meant students in classes with the lowest average general self-efficacy with their previous teachers had lower general self-efficacy with their current teacher than students in other groups, if the teacher's classroom management or mathematics instruction ratings were low; and higher general self-efficacy with their current teacher than students in other groups, if the teacher's classroom management or mathematics instruction ratings were high. In the 3-domain model, such interactions were not significant.

---

[38] OECD (2020a)

Relationships between teaching practices and personal interest in mathematics

In models with personal interest in mathematics with the current teacher as the outcome, there were no significant relationships between *classroom management*, *social-emotional support*, *discourse and assessment*, or *mathematics instruction* and personal interest in mathematics with the current teacher before accounting for student and school characteristics. The same was found for *classroom management*, *social-emotional support* and *instruction* in the 3-domain model.

How do relationships between teaching practices and personal interest in mathematics vary by student and class characteristics?

Still using personal interest in mathematics with the current teacher as the outcome, after accounting for student and class characteristics and *OTL*, there was a significant interaction between the class average personal interest in mathematics with the previous teacher and teaching practice, but only in relation to discourse and assessment, and only for students taught in classes in the lowest quartile of average personal interest in mathematics with the previous teacher. Such interactions were not significant in the 3-domain model. For students in classes with the highest teacher ratings for discourse and assessment, the class average personal interest in mathematics with the previous teacher made less of a difference to individual personal interest in mathematics with the current teacher except for the lowest quartile (who were likely to have substantially lower personal interest in mathematics with the current teacher). For students in classes with the lowest teacher discourse and assessment ratings, the higher the class average personal interest in mathematics with the previous teacher, the lower the individual personal interest in mathematics with the current teacher.

What were the relationships between teaching practices?

This section examines the relationships between teaching practices as measured within a particular type of rating (video components, video indicators, or artefact components) as well as the relationships between teaching practices as measured across different types of ratings (for example, relationships between video component and video indicator ratings of teaching practices within the same domain).

Relationships are investigated using correlations. Domains are defined as outlined in the introduction to this chapter (with results reported according to both the 4-domain and 3-domain versions used for modelling above). Except in rare instances where negative relationships are explicitly indicated in the text below, all significant relationships between teacher practices were positive. Indicators discussed in this section exclude those

identifying the use of particular types of technology such as graphical calculators or smartboards as there was very little variation in their use within England. Further information can be found in the England Technical Report[39].

Relationships between teaching practices across domains: video components

There were strong associations between most of the components for the *classroom management* domain, with the exception of a weak association between *routines* and *monitoring*. The strongest association was between the teacher average ratings for *disruptions* and the overall teacher average rating for the domain. All of these associations were significant.

Within the *social-emotional support* domain, all components were significantly positively related to one another except for *respect* and the overall teacher average rating for the domain. The relationships between *respect* and *encouragement and warmth* as well as between *respect* and *risk-taking* were weak, while the relationship between *risk-taking* and the overall teacher average rating for the domain was moderate. Relationships between *risk-taking* and *encouragement and warmth*, and between *encouragement and warmth* and the overall teacher average rating for *social-emotional support*, were strong.

Classroom management components were all significantly and positively related to *social-emotional support* components, except for a non-significant relationship between *risk-taking* and *routines*. The strongest relationships were between *respect* and *disruptions* as well as the overall average teacher rating for *classroom management*, and the weakest relationships were between *encouragement and warmth* and *routines* as well as *disruptions* and between *risk-taking* and *disruptions*. This suggests that teachers with higher average ratings for *classroom management* components often also had higher average ratings for *social-emotional support* components.

Classroom management components were also significantly and positively related to all *discourse* components, with moderate relationships between *explanations* and *routines* as well as the overall average teacher rating for *classroom management* and between the overall average teacher rating for *discourse* and *routines*, *disruptions* and the overall *classroom management* average teacher rating. All other relationships between *classroom management* and *discourse* components were weak. Most of the relationships between *classroom management* and *assessment of and responses to student understanding* components were significant, and most were weak. *Eliciting student feedback* was, however, moderately and positively related to *routines*, *disruptions* and the overall *classroom management* average teacher rating, as were the overall average teacher ratings for *classroom management* and *assessment of and responses to student understanding*. This suggests that teachers who had higher ratings for *classroom*

---

[39] McCann, Riggall, Ingram, and Lindorff (2020)

management tended to also have higher ratings for *discourse* as well as, to some extent, *assessment of and responses to student understanding*.

All of the relationships between *classroom management* and *quality of subject matter* components were weak and positive, but significant. Most of the associations between *classroom management* and *student cognitive engagement* components were significant, positive, and weak. Exceptions to this included non-significant relationships between *engagement in cognitively demanding subject matter* and *routines*, and between *understanding of subject matter* and all *classroom management* components including the overall average teacher rating. There was also one significant weak negative relationship between *using multiple approaches* and *monitoring*. As was the case for other domains, this means that teachers who had higher ratings for *classroom management* tended to have somewhat higher ratings for *quality of subject matter* and *student cognitive engagement*.

Social-emotional support components were, in general, significantly and positively related to *discourse* components except for a non-significant relationship between *encouragement and warmth* and *explanations*. The strongest relationship was between *risk-taking* and the overall average teacher rating for *discourse*, and the weakest was between *encouragement and warmth* and *questioning*. Most relationships between *social-emotional support* and *assessment of and responses to student understanding* components were positive, significant and weak to moderate, except for insignificant relationships between *aligning instruction* and all *social-emotional support* components including the overall average teacher rating. The strongest (still moderate) significant relationship was between *eliciting student feedback* and the overall average teacher rating for *social-emotional support*, and the weakest was between *teacher feedback* and *respect*. In general, teachers who had higher ratings for *social-emotional support* often also had higher ratings for *discourse* and for *assessment of and responses to student understanding*.

Relationships between *social-emotional support* and *quality of subject matter components* tended to be significant, weak and positive. Exceptions to this included non-significant relationships between *explicit patterns and generalisations* and *respect* as well as *risk-taking*, and between the overall average teacher *quality of subject matter* rating and *encouragement and warmth* as well as *risk-taking*. Most of the associations between *social-emotional support* and *student cognitive engagement* components were also significant, weak, and positive. *Risk-taking*, however, was significantly, positively and moderately associated with *cognitively demanding subject matter*, *understanding of subject matter* and the overall *student cognitive engagement* overall average teacher rating. There were non-significant relationships between *cognitively demanding subject matter* and *respect*, as well as between *understanding of subject matter* and *respect*, *encouragement and warmth* and the overall *social-emotional support* average teacher rating. This suggests that teachers with higher *social-emotional support* ratings tended to have higher ratings for *quality of subject matter*. This was the case for *student cognitive*

engagement to some extent, particularly with regard to teachers' ratings for the *social-emotional support* component *risk-taking*.

Within the *discourse and assessment* domain, all *discourse* components were significantly and positively associated with one another, as were all components within *assessment of and responses to student understanding*. These associations were all moderate to strong. Between *discourse* and *assessment of and responses to student understanding* components, all relationships were significant and positive except for a non-significant relationship between *aligning instruction* and *nature of discourse*. *Aligning instruction* was, however, positively and moderately related to other *discourse* components and positively and strongly related to the overall teacher average rating for *discourse*. *Teacher feedback* was weakly and positively associated with *nature of discourse* and *questioning*, and moderately and positively associated with *explanations* and the overall teacher average rating for *discourse*. *Eliciting student feedback* was moderately and positively related to *nature of discourse* and strongly positively related to other *discourse* components. Overall teacher average ratings for *assessment of and responses to student understanding* were moderately and positively related to *nature of discourse* and *questioning*, and strongly and positively related to *explanations* and the overall teacher average rating for *discourse*. In general, teachers who had higher ratings for *discourse* often had higher ratings for *assessment of and responses to student understanding*.

Within the *mathematics instruction* domain, all *student cognitive engagement* components were significantly and positively related to one another. These associations were strong except for a moderate relationship between *understanding of subject matter* and *multiple approaches*. Within the *quality of subject matter* components, the relationship between *explicit patterns and generalisations* and *explicit connections* was positive, moderate and significant. The overall teacher average rating for *quality of subject matter* was not significantly related to *explicit connections* nor to *explicit patterns and generalisations*. All *student cognitive engagement* components, as well as the overall *student cognitive engagement* teacher average rating, were significantly and positively associated with *explicit connections*, *explicit patterns and generalisations*, and the overall teacher average rating for *quality of subject matter*. These relationships were all moderate except for a strong relationship between the overall teacher average ratings. This suggests that despite not all of the *quality of subject matter* components being related to one another, teachers with higher ratings for *quality of subject matter* often also had higher ratings for *student cognitive engagement*.

In the broader *instruction* domain (encompassing *discourse*, *quality of subject matter*, *student cognitive engagement* and *assessment of and responses to student understanding*), most components were significantly and positively related, in addition to those already discussed above within *discourse and assessment* and *mathematics instruction* domains. Overall teacher average ratings for *quality of subject matter* were not significantly related to any of the *discourse* components, but *explicit connections* and

explicit patterns and generalisations were significantly and moderately related to *questioning*, *explanations* and the overall teacher average *discourse* rating, as well as being significantly and more weakly associated with *nature of discourse*. All *student cognitive engagement* components were significantly and moderately-to-strongly associated with all *discourse* components. Relationships between *student cognitive engagement* and *assessment of and responses to student understanding* components were also all significant, and all of these relationships were moderate to strong except for weak relationships between *teacher feedback* and all *student cognitive engagement* components and overall teacher average ratings, and between *aligning instruction* and *multiple approaches*. Most relationships between *quality of subject matter* and *assessment of and responses to student understanding* were moderate and significant, except for *explicit patterns and generalisations* which was not significantly related to *aligning instruction*, and weakly (though significantly) related to all other *assessment of and responses to student understanding* components and the overall teacher average rating. Teachers with higher *discourse* ratings did not necessarily tend to have higher *quality of subject matter* ratings, but they did often have higher ratings for *student cognitive engagement*. Teachers with higher ratings for *assessment of and responses to student understanding* also tended to have higher ratings for *student cognitive engagement* as well as for *quality of subject matter*.

Relationships between teaching practices across domains: video indicators

Within the *classroom management* domain, there were positive weak and significant relationships between *small group* and *whole group* as well as *pair* activity structures being used in lessons. On the other hand, *whole group* activity structure was negatively and weakly related to time on task, as was *individual* to *small group* and *pair* activity structures.

Within the *social-emotional support* domain, there was a weak, positive and significant relationship between the average percentage of segments with *persistence* present and *requests for public sharing*. That is, to some extent, classrooms in which *persistence* was more often observed also tended to have more *requests for public sharing*.

Between the *classroom management* and *social-emotional support* domains, most relationships were not significant. Exceptions included a positive, weak relationship between *individual* activity structure and the average lesson rating for *persistence*, and weak negative relationships between the average lesson rating for *persistence* and *time on task* as well as *small group* and *pair* activity structures, plus a weak negative relationship between the average percentage of segments with *persistence* present and *whole group* activity structure. This suggests that overall there was not much of a pattern in teachers' ratings for *social-emotional support* based on their *classroom management* ratings.

Classroom management indicators were largely not significantly related to indicators in the *instruction* domain. Exceptions to this included weak positive relationships between *whole group* activity structure and *discussion opportunities*, *connecting mathematical topics* (average maximum rating), *mathematical summary* (average maximum rating), *technology for understanding* (average rating), and *software use for learning* (average maximum rating), as well as between *accuracy* (average rating) and *time on task*. There were further weak, positive relationships between *pair* activity structure, using *equations* as representations, and *technology for understanding* (average rating), as well as between *individual* activity structure and *persistence* (average rating). Again, as was the case for *social-emotional support*, this suggests that in general having higher ratings for *classroom management* did not mean that a teacher was particularly likely to have higher ratings for instruction.

Similarly, most *social-emotional support* indicators were not significantly related to indicators in the *instruction* domain, though there were a few exceptions. The average lesson rating for *persistence* was weakly and negatively related to *discussion opportunities* and to *metacognition*, while the average percentage of segments with *persistence* present was weakly and positively related to the average maximum rating for *metacognition*, and weakly and negatively related to *repetitive use opportunities* (average maximum rating). *Requests for public sharing* was weakly and positively related to *discussion opportunities* and to *metacognition* (average maximum rating), and weakly and negatively related to *technology for understanding* (average rating) and *software use for learning* (average maximum rating). In general, this means that there was not an overall tendency for teachers with higher *social-emotional support* ratings to have higher or lower *instruction* ratings, though there were a few trends across specific components in these domains.

The only indicator in the *instruction* domain that was also relevant to the *discourse and assessment* domain (in the 4-domain model) was *discussion opportunities*. This was only significantly related to a small number of indicators in the *mathematics instruction* domain, including weak positive associations with *real-world connections* (average maximum) and *connecting mathematical topics*, and weak negative associations with *explicit learning goals* and using *equations* as representations. This suggests that while there was not a strong overall pattern between teachers' ratings for *discourse* and *assessment* and for *instruction*, higher ratings for *discussion opportunities* were at least to some extent likely to accompany higher ratings for some *mathematics instruction* components and lower ratings for others.

The majority of indicators relevant to the *quality of subject matter* were not significantly related to one another, but there were some exceptions to this. There were weak positive relationships between using *equations* as representations and *explicit learning goals*; between *real-world connections* (both average rating and average maximum rating), *connecting mathematical topics*, and *organisation of procedural instruction*; between using *graphs* and using *tables* as representations; between *mathematical summary*

(average maximum rating) and using *graphs* as well as *connecting mathematical topics* (both average rating and average maximum rating); and between *connecting mathematical topics* (average rating) and *real-world connections* (both average rating and average maximum rating). There were weak negative relationships between *organisation of procedural instruction* and using *tables* and *objects* as representations; between using *equations* and using *objects* and *drawings*; between using *drawings* and using *graphs* and *tables*, as well as *accuracy* (average minimum rating); between *accuracy* (average minimum rating) and *mathematical summary* (average maximum rating) as well as *real-world connections* (both average rating and average maximum rating); and between *accuracy* (average rating) and *explicit learning goals*.

Most indicators relevant to *student cognitive engagement* were also not significantly related to one another. Exceptions included weak negative relationships between *repetitive use opportunities* (average maximum rating) and *technology for understanding* (average rating) as well as *metacognition* (both average rating and average maximum rating).

Similarly, most indicators were not significantly related across *student cognitive engagement* and *quality of subject matter*. The only exceptions were several weak positive relationships between *metacognition* (average rating) and *explicit learning goals* as well as *mathematical summary* (average maximum rating), and several weak negative relationships between *metacognition* (average rating) and using *graphs* as representations as well as organisation of procedural instruction, and between *repetitive use opportunities* (average maximum rating) and *explicit learning goals*. This suggests that if a teacher had higher ratings for *student cognitive engagement*, that did not necessarily mean that they were more likely to have higher ratings for *quality of subject matter*.

Relationships between teaching practices across domains: artefact components

None of the artefact components were relevant to the *classroom management* or *social-emotional support* domains.

Within the *discourse and assessment* domain, *asking for explanations* and *encouraging student self-evaluation* were not significantly related. *Asking for explanations* was also not significantly related to most artefact components relevant to *quality of subject matter* or *student cognitive engagement*. Exceptions to this included significant and weak associations between *asking for explanations* and *using multiple mathematical methods* (positive) as well as *opportunities to practice a skill or procedure* (negative). There were also significant associations between *asking for explanations* and *connecting mathematical representations* (moderate, positive) as well as *real-world connections* (weak, positive).

Within the *mathematics instruction* domain, the significance, direction and magnitude of relationships was quite variable. Within artefact components relevant to *quality of subject matter*, the only significant and moderate relationships were between *connecting mathematical representations* and *real-world connections*, and between *explicit learning goals* and *addressing diverse student needs*. Significant, weak positive relationships existed between *explicit learning goals* and *real-world connections*, as well as between *connecting mathematical representations* and three other components: *explicit learning goals*, *addressing diverse student needs*, and *explicit patterns and generalisations*. There were significant, weak and negative relationships between *accuracy of materials* and *explicit learning goals* as well as *addressing diverse student needs*, and between *explicit patterns and generalisations* and *real-world connections*. Within artefact components relevant to *student cognitive engagement*, *opportunities to practice a skill or procedure* was related positively and significantly to *using multiple mathematical methods*, and significantly, weakly and negatively to using *technology for understanding*. Relationships across *quality of subject matter* and *student cognitive engagement* components were similarly varied. There was a moderate, significant and positive relationship between *connecting mathematical representations* and *using multiple mathematical methods*. However, all other significant relationships were weak, including positive associations between *using multiple mathematical methods* and *explicit learning goals*, *addressing diverse student needs*, and *explicit patterns and generalisations*; and negative associations between *using multiple mathematical methods* and *accuracy of materials*, between *opportunities to practice a skill or procedure* and *connecting mathematical representations* as well as *real-world connections*, and between using *technology for understanding* and *addressing diverse student needs*.

Together, these results suggest that although the relationships between teachers' ratings on specific artefact components had some patterns, overall having higher ratings for a particular domain did not mean that teachers were more likely to have higher ratings for other domains.

Relationships between teaching practices within domains across different types of rating (video components, video indicators, and artefact components)

There were four significant relationships between teachers' ratings on video components and video indicators relevant to *classroom management*. Teachers' ratings for the indicator *time on task* were significantly, moderately and positively related to the video component rating for *routines*, and significantly, positively but weakly related to the video component rating for *disruptions* as well as the overall average rating for the *classroom management* domain. The rating for the amount of time spent in lessons on *whole group* instruction was significantly and positively but weakly associated with the video component rating for *disruptions*. In other words, teachers whose classes had more time on task had higher ratings for routines and fewer (or more-efficiently handled) disruptions, and overall higher classroom management ratings (in video components).

126

Having more whole-group instruction was also to some extent related to having fewer or more-efficiently handled *disruptions*.

There were several significant associations between teachers' ratings on video components and video indicators relevant to the *quality of subject matter*. The use of graphs as representations had a significant, strong and positive relationship to the video component rating for *explicit connections*, a significant, moderate and positive relationship to the overall video component rating for the *quality of subject matter*, and a significant, positive but weak relationship to the video component rating for *explicit patterns and generalisations*. The use of tables as representations was significantly and positively but weakly associated with the video component rating for *explicit patterns and generalisations*, and the use of drawings as representations was significantly, moderately and positively associated with both the video component ratings for *explicit connections* and for the *quality of subject matter* overall. This suggests that teachers who had higher ratings for the *quality of subject matter* for video components often had higher ratings for video indicators for the *quality of subject matter* as well.

Most of the relationships between teachers' artefact and video component ratings relevant to the *quality of subject matter* were not significant. The *mathematical connections* artefact rating was significantly strongly and positively associated with the *explicit connections* video component rating as well as significantly moderately and positively associated with the overall *quality of subject matter* video component average rating. There was also a significant, moderate positive association between the artefact rating for *real-world connections* and the *explicit connections* video component rating. This means that having higher ratings overall for *quality of subject matter* for artefacts did not necessarily mean a teacher was more likely to have higher ratings for *quality of subject matter* for video components.

There were a number of significant relationships between the artefact and video indicator ratings relevant to the *quality of subject matter*. The indicator rating for *explicit learning goals* was significantly, moderately and positively associated with the artefact rating for *plans and learning goals*, as were the indicator and artefact ratings for *real-world connections*. The use of graphs as representations was significantly, strongly and positively related to the artefact rating for *mathematical connections*, and significantly, positively but weakly related to the artefact rating for *real-world connections*. The use of drawings as representations was significantly and positively but weakly associated with the artefact ratings for both *mathematical connections* and *real-world connections*. Teachers' average maximum ratings for *real-world connections* and *connecting mathematical topics* were significantly and positively related to the artefact rating for *real-world connections* (moderately and weakly, respectively). The average maximum rating for the *mathematical summary* video indicator was also significantly, moderately and positively related to the artefact rating for *patterns and generalisations*. At least to some extent, this means that having higher ratings for the *quality of subject matter* for artefacts

often meant that teachers also had higher ratings for the *quality of subject matter* for video indicators.

Summary

Students with higher pre-test scores tended to have higher attainment on the post-test, and students' pre-test scores explained half the variation in post-test scores. After accounting for individual pre-test score and student characteristics, students in lower-attaining classes had lower post-test scores even after accounting for their individual pre-test scores. Class average pre-test score made more of a difference for students who were higher-attaining on the pre-test. On average, students who had higher general self-efficacy in mathematics with their previous teacher also had higher general self-efficacy in mathematics with their current teacher. Students' previous self-efficacy scores explained about one-fifth of the variation in students' current self-efficacy scores. After accounting for self-efficacy in mathematics with their previous teacher, as well as student characteristics, students in classes with a higher class average parental education had lower general self-efficacy in mathematics scores with their current teacher. Overall, students who had a higher personal interest in mathematics with their previous teacher also had higher personal interest in mathematics with their current teacher, with their previous interest in mathematics explaining approximately one-eighth of the variation in students' current interest. After accounting for personal interest in mathematics with their previous teacher, students who were a first- or second-generation immigrant had higher personal interest in mathematics with their current teacher. Before accounting for student and class characteristics, there were no significant relationships between teaching practice and the student outcomes of post-test attainment, general self-efficacy in mathematics with the current teacher, and personal interest in mathematics with the current teacher. There were some significant interactions after accounting for student and class characteristics in relation to the *classroom management*, *mathematics instruction*, and *discourse and assessment* domains.

References

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., and Schmidt, W. (2019). *Teaching for excellence and equity: Analyzing teacher characteristics, behaviors and student outcomes with TIMSS*. New York: Springer

Department for Education (2016). *Standard for teachers' professional development*. Department for Education. Accessed 15/04/2020 from https://www.gov.uk/government/publications/standard-for-teachers-professional-development

Jerrim, J. and Sims, S. (2019). *The Teaching and Learning International Survey (TALIS) 2018*. Department for Education Research report. Accessed 15/04/2020 from https://www.gov.uk/government/publications/teachers-in-primary-and-secondary-schools-talis-2018

Kaput, J. J. (1999). Teaching and learning a new algebra. In: E. Fennema and T. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 145–168). New York: Routledge

Klieme E. (2013). The Role of Large-Scale Assessments in Research on Educational Effectiveness and School Development. In: von Davier M., Gonzalez E., Kirsch I., and Yamamoto K. (Eds.), *The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research*. Dordrecht: Springer

McCann, E., Riggall, A., Ingram, J. and Lindorff, A. (2020). *TALIS Video Study Technical Report*. London, United Kingdom: DfE

Mehan, H. (1979). *Learning lessons*. Cambridge, MA: Harvard University.

OECD (2020a). *Global Teaching InSights: A Video Study of Teaching*. Paris: OECD

OECD (2020b). *Global Teaching InSights Technical Report.* Paris: OECD

Sammons, P., Sylva, K., Melhuish, E., Siraj, I., Taggart, B., Toth, K., and Smees, R. (2014). *Influences on Students' GCSE Attainment and Progress at Age 16*. London: Department for Education

Schneider, S. (2008). *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries*. Manheim: MZES.

Schmidt, W. and Maier, A. (2009). Opportunity to learn. In: Sykes, G., Schneider, B., and Plank, D. N. (Eds.), *Handbook of Education Policy Research* (pp. 541–559). New York: Routledge

Sinclair, J. and Coulthard, M. (1975). *Towards an analysis of discourse*. Oxford: Oxford University Press

Sizmur, J., Ager, R., Bradshaw, J., Classick, R., Galvis, M., Packer, J., Thomas, D., and Wheater, R. (2019). *Achievement of 15-year-olds in England: PISA 2018 results*. Department for Education Research report. Accessed 15/04/2020 from https://www.gov.uk/government/publications/pisa-2018-national-report-for-england

Strand (2014). School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11, *Oxford Review of Education*, 40:2, 223–245, DOI: 10.1080/03054985.2014.891980