Science Education in England:

Exploring the Evidence *for* and Evidence *of*

Reform

By

Marilyn Hall M.A. Education Studies

Submitted to the University of Sussex in fulfilment of the degree of

Doctor of Philosophy

February 2020

Acknowledgments

Abstract

It is argued that Science as a curriculum subject along with Technology, Engineering, and Maths (STEM) is fundamental in developing a highly-skilled workforce that in turn, drives economic growth and global competitiveness. As such, science education is considered to be of major significance and therefore subject to much government intervention and reform. This study explored the interaction between student attainment in science and teacher classroom practice in response to reform.

The national curriculum reforms introduced in September 2014 have increased the level of challenge encountered by students in their science learning. The ultimate aim of these reforms was to raise attainment in end of key stage assessments for all students and potentially, in global level assessments. The introduction of these reforms and their perceived impact on the way science education is taught and assessed formed the catalyst for embarking on this study.

This research used a mixed-methods design, through a critical realist lens, to explore the associations between policy, historic student attainment and teacher practice. Using secondary data, the quantitative component drew down on the National Pupil Database to analyse English students' attainment in science at the end of key stage 2 and key stage 4, between 2008 to 2018. The voices of 26 secondary and primary science teachers from the South of England, interviewed between October 2017 and April 2018, provided the data for the qualitative component of this study. Using self-completion questionnaires and one-to-one interviews, data were collected that uncovered teacher perspectives on the impact of reform on their classroom practice.

The quantitative findings indicated that whilst attainment at Key Stage 2 and 4 (GCSE) had generally increased over time, the attainment gaps between different groups of students persisted despite reforms to address this. Furthermore, the qualitative analysis found a reliance by primary teachers on purchased schemes of learning, reduced curriculum time and pressure on budgets for primary science. In secondary science, participants outlined that there was less time or support to deliver the increased volume of cognitively demanding content to a wide range of mixed attainment students. Ultimately, at key stage 4, the changes to the accountability measures had constrained teachers' ability to offer an alternative route and enriching experience through GCSE sciences for many of their students.

The theoretical underpinnings of this study positioned teachers in the role of "street-level bureaucrats" and used the Chordal Triad conception of agency to understand teacher practice within the cycle of reform. The contribution made by this theoretical perspective are the insights into teachers' responses to education reform likely to be missed by studies that focus largely on individual teacher knowledge, beliefs or agency alone. Emerging from the analysis were the two key themes of *equity and fairness* and *time and resources* that illuminated our understanding of the impact of policy on practice.

During this period of change and transition participants, had undertaken "translation work", used their past experience and existing schemas to accommodate new understanding and modes of working. The study concluded that while there was a high level of resilience and positive projections for the future which drove teachers' work, science education reform had repercussions for the STEM pipeline, teaching and learning activities, teacher education and continued professional development. Whether to enable collaborative working, across institutions and cross-phase; for covering the curriculum or for embedding the new measures, from both primary and secondary schools, there was a continued call for more time and resources.

Contents

List of Figures

Data Tables

Appendices Tables

Abbreviations

A-level	Advanced level General Certificate in Education qualification usually taken at aged 18.
DfE	Department of Education. Previously also called DfES - Department for Education and Skills; DCSF - Department for Children, Schools and Families.
DIRT	Dedicated Improvement and Reflection Time
EAL	English as an Additional Language
Ebacc	The English Baccalaureate performance indicator linked to the General Certificate of Secondary Education (GCSE)
ERA	Education Reform Act 1988
FSM	Free School Meals
GCSE	General Certificate in Secondary Education qualification usually taken at aged 16.
ITE	Initial Teacher Training
KS2	Key Stage 2 - students aged between 7 and 11 years old.
KS3	Key Stage 3 - students aged between 11 and 14 years old.
KS4	Key Stage 4 - students aged between 14 and 16 years old.
LEA	Local Education Authority
LMS	Local Management of Schools
NPD	National Pupil Database
OECD	Organisation for Economic Co-operation and Development
Ofqual	Office of Qualifications and Examinations Regulation
Ofsted	Office for Standards in Education, Children's Services and Skills
PISA	Programme for International Student Assessment
PP	Pupil Premium
SEN	Special Educational Needs
STEM	Science, Technology, Engineering and Mathematics
TALIS	Teaching and Learning International Survey
TIMSS	Trends in International Mathematics and Science Study

Chapter 1: Introduction

This study explored the interaction between student attainment in science and teacher classroom practice in an era of education reform. STEM (Science, Technology, Engineering, and Mathematics) disciplines are considered, particularly by OECD countries, to be fundamental to the development of the highly-skilled workforce needed for economic growth and global competitiveness (EU STEM Coalition Team, 2016; OECD, 2010). However, the reliance on STEM may offer a one-dimensional solution that fails to consider many of the other factors that influence economic growth. The underrepresentation of women, minority ethnic and low-income groups in STEM careers, also suggests that talented individuals have fewer opportunities to contribute (Billimoria, 2017). Furthermore, the orientation of STEM education towards a neoliberal agenda, risks shifting the goals of science education from collaborative, interdisciplinary working towards the narrow accumulation of human capital (Carter, 2016; Weinstein *et al.*, 2016). Nevertheless, the outcomes of science education, beginning from the early stages of a child's learning, remain a significant focus of continued government action and reform (HM Treasury, 2014).

Driven in part as a response to England's performance in the Programme for International Student Assessment (PISA) tests, (Wheater *et al.*, 2014), a revised national curriculum was put in place September 2014 (DfE, 2014a). At the outset, the reforms aimed to raise the level of challenge for students in science education and by doing so, improve attainment in the end of the key stage assessments and global assessment rankings (DfE and Gove, 2011). It was in the light of the 2014-2016 reforms to the national curriculum, its assessment and the associated accountability measures (Long, 2017; Roberts and Bolton, 2016) that this study was originally conceived. As a secondary science teacher, I became aware that my pedagogic

practice had changed in response to science education reform and as such, was keen to find out if this was consistent with fellow practitioners.

Rooted in strong standardisation and accountability measures that govern the actions of teachers and students (Bowe *et al.*, 2017), there is a high level of government control within the English education system. Student attainment data creates a focal point in policymaking and leads to a justification for framing an understanding of the *why* of the reform agenda (Ball, 2010). This study, therefore, looks at the evidence *for* reform, as generated by the attainment of students in science at key points in their school experience.   Student attainment data from end of key stage 2 and 4 (GCSE) examinations obtained from the National Pupil Database (NPD) was analysed to look for trends within and across the different assessment regimes over time. The broad scope and range of the most recent reforms necessitate scrutiny and critique; therefore, the analysis is critical to generating an understanding of the impact of reform.

If there is evidence *of* reform, it may emerge from developing an understanding of teachers' interaction with the reforms and how they accommodate changes through their practice. A questionnaire designed to elicit data about the use of specific classroom practices is linked to responses to an in-depth one-to-one interview. This integrated data was used to explore the impact of reform on the teachers in my study.

Research suggests that teacher practice is developed through a complex interaction involving beliefs about science (Herrington *et al.*, 2016; Mansour, 2013; Waters-Adams, 2006) and an individual experience which links effective teaching with positive outcomes for students (Rosenshine, 2012; Leander and Osborne, 2008). However, teacher classroom practice can also be shaped by external factors in response to government policy (Wallace and Priestley, 2017; Ryder and Banner,

2013). Reforms that dictate the assessment of the national curriculum (Burgess and Thomson, 2019; Heinrich and Stringer, 2012); the standardisation of teacher professionalism (DfE, 2013a), and the introduction of accountability measures which promote teacher performativity (Holloway and Brass, 2018; Wilkins, 2011; Ball, 2003), serve to constrain teacher action by limiting the options available to them in managing the learning of groups of students. This, I argue, is the outward demonstration of how government reform is embodied by teachers to achieve the raising attainment ambitions. These mechanisms represent the *how* of reform and are reflected in the everyday actions of individuals and groups in response to reform.

Considered to be unfounded, the alarm raised by England's results in PISA tests (Jerrim and Shure, 2016) prompted many of the curriculum and assessment reforms introduced in 2014. I conclude that these reforms which focus on the narrow outcomes of summative, high-stakes tests risk losing what is most enriching about science teaching and learning. The wider, culturally centred view of education is pushed aside in the drive to outperform other countries (Alexander, 2012). I, therefore, suggest that, if the aim of raising attainment for all and closing the gaps between different groups of students is to be met, then wider-reaching evidence is needed ahead of instigating whole-scale reform to the complex, socially integrated education context.

Thesis Structure

Moving forward, the thesis is set out across eleven chapters. Chapters 2 and 3 lay out the context of the study and establish the historical and macrostructure of England's education landscape. Focussing on the part played by three fundamental tenets of the standards-based reform agenda: marketisation, standardisation, accountability, chapter 2 explains how these mechanisms have been used to shape and change science education using student data at the centre. Chapter 3 then

describes the structure of compulsory education, the school population, the curriculum and how it is assessed. The chapter also introduces the terminology and vocabulary currently used in education settings with a particular focus on science.

Chapter 4, the literature review, synthesises the ideas surrounding what determines teacher practice and how it evolves through internal factors associated with the individual and external factors associated with the structure of education. The chapter briefly evaluates the ideas that explain how learning theories determine particular features of teacher practice, particularly in teaching science. The literature on teaching and learning is used to locate the boundaries for the questionnaire, interview questions and subsequent analysis. In chapter 5, in addition to stating the research questions, I explain the theoretical framework through which the data is analysed and interpreted. The chapter makes a distinction between different conceptualisations of agency and its application to explore the internally and externally driven aspects of teacher actions.

The methodological decision-making process is laid out in chapter 6. My justifications for applying a critical realist lens are followed by an explanation of how this particular viewpoint influenced my decision to use a mixed-methods research design. The chapter also describes the processes involved in the use of the National Pupil Database to undertake quantitative analysis of student attainment data that was carried out concurrently with the qualitative data collection. A detailed rationale outlines the decisions taken in the development of the pre-interview questionnaire. To conclude the chapter, I demonstrate how ethical considerations guided my actions in conducting the research and acknowledge the possible impact that my positionality, relative to the participants and their context, may have on the validity of the resultant conclusions.

There are three findings chapters. Chapter 7, focusses on the analysis of student attainment data derived from the end of key stage 2 (KS2) teacher assessment and GCSE science grades at key stage 4 (KS4). The trends and patterns within the data across the years 2008 to 2016 are described for whole cohorts in addition to unpicking the attainment data of students by gender, socio-economic and SEN status. The second and third analysis chapters (8 and 9) focus on the data derived from the pre-interview teacher questionnaire and their interview responses. Chapter 9 situates the responses from the one-to-one interviews within the theoretical framework set out in chapter 5. Teachers' classroom practice is explored from the perspective of their past, present and future actions and attitudes to the new curriculum, assessment and accountability measures are drawn out from the data. Teacher's perspectives of the impact of these reforms on their discretionary-decision making is derived by interpreting the data through the theoretical lens of Street-Level Bureaucracy.

Elements of the study are drawn together in the discussion chapter 10. The knowledge generated by the analysis of the student attainment data is interpreted alongside the understanding arising from the qualitative data. The arguments are built through the deployment of the theoretical framework, reflecting the interactions between mechanism, outcomes and context and thereby, drawing conclusions about the findings and highlighting the originality of the research. Chapter 11 concludes the thesis and summarises the outcomes. Pointing to the study's limitations, the chapter indicates where year on year comparisons in student attainment data were problematic. I also discuss the difficulties encountered in using the NPD and other methodological issues. The implications for further research suggest ways in which the knowledge can be built upon. Revisiting the work, to reassess the impact of the 2016 reforms to science education, in 4- or 5-years' time would speak back to the earlier analysis of teacher practice along temporal lines.

Chapter 2: Three Mechanisms of the English School System: Marketisation, Standardisation and Accountability

Introduction

A "new educational orthodoxy" (Hargreaves *et al.,* 2002, p. 1) has become predominant in Anglo-Western educational systems. Structured through reforms that demand high standards with a narrow focus on literacy and numeracy, these education systems employ a centralised curriculum with aligned assessment monitored via a formal accountability framework. Accordingly, behind the scenes of the specific education policy changes affecting primary and secondary school science, lies an overall change to the discourse of education towards a market economy view, global comparisons and policy borrowing. What Apple (2005, p. 11) referred to as a neo-liberal commitment to the market; a neo-conservative emphasis on active control of the curriculum and "new managerial" policies which install rigorous forms of accountability int schools are now normalised.

The culture, that involves high stakes testing, the publishing of exam results and ranking schools against each other in national league tables using government devised accountability frameworks, is well established (Gilbert, 2012). Seen as a tool to reform education, policymakers use accountability measures as a short cut towards raising standards as opposed to awaiting the rewards of a long term investment in changes to pedagogy, practice and the curriculum (Stobart, 2008). The form of "methodological selectivity" (Alexander, 2014, p. 361) that detaches evidence from the school and classroom lends the top-down policy interventions greater powers to effect improvements. Included is the increased market-driven competitiveness between schools, e.g. free schools vs. academies vs. state schools; the reliance on a standardised core curriculum and the prolific use of data to measure performance, make judgements and to hold schools to account. In the UK

today, we take this environment as the norm with its roots firmly traced back to the 1988 Education Reform Act. This study examines the education context from the post-1988 onwards.

The first part of this chapter outlines the characteristics of three control mechanisms that the literature (Sahlberg, 2006; Scoppio, 2002) suggests frame the education landscape - *marketisation, standardisation and accountability*. Charting the historical context of England's education system explains how these mechanisms were brought into prominence through the Education Reform Act in 1988. The remainder of the chapter deals with *marketisation, standardisation and accountability,* in turn, to argue that these mechanisms continue to impact and shape the education system today. Student attainment data permeates and unifies all three control mechanisms and is used to inform government decision-making and education policy. Therefore, the chapter also explores the role played by the *datafication* of education by looking at the types, uses and implications of student attainment data. This chapter is deliberately broad as it explores the structural boundaries of the education system since 1988 and sets the scene into which the science curriculum is positioned. The following chapter, Chapter 3, gives a deeper explanation of the development of the science curriculum and examines the part played by various stakeholders in determining the direction of science curriculum reform.

2.1 Three Control Mechanisms

2.1.1 Marketisation

The marketisation of education is the introduction of market forces to state-funded services that generate greater school autonomy and result in increased competition for pupils, resources and funding (Bartlett, 1993). The OECD explored the emergence of education as a consumer-based commodity as early as 1994 (Hirsch, 1994). The

growing diversity of school types and increasing autonomy within school management was being witnessed in many different countries, often brought about by governments with a strong pro-market stance (Hirsch, 1994). Ultimately, marketisation of education redefined the relationship between government, school and parents (Power and Whitty, 1996). Through reforms which created devolved systems of schooling, the application of market forces increased the emphasis on parental choice and competition between increasingly diversified types of school (Whitty and Power, 2000). Marketisation is seen as beneficial where the increased choice and diversity in education, related to student-funding, leads to increased student attainment. Less beneficial, however, is the loss of social cohesion once brought about through central control (Gorard and Fitz, 1998). Parents with greater "cultural capital" more readily use this advantage to secure their choice, leading to an increase in inequality and possible greater social stratification (Noden, 2000).

2.1.2 Standardisation

"Standards are a type of social technology which comprises a discursive apparatus of codified abstract rules or norms" (Gronn, 2003, p. 8). This definition implies that standards engender uniformity of conduct in preference to variation in performance, ultimately leading to compliance and verification. All education systems are, to a lesser or greater degree, standardised in order to bring about equality of standards nationwide (Bol and Van de Werfhorst, 2011). The term "standardisation" in the context of education, encompasses any structures that dictate measurable expectations, including how and what children are expected to learn and their expected levels of attainment and progress. It also covers what and how teachers are expected to teach, how they should behave and what schools should advocate in order to raise attainment levels.

The aim to achieve universal education for all with an emphasis on student learning and high expectations are some of the inherent positive effects of standardising education (Sahlberg, 2016). These policies, however, rely on high-stakes standardised tests; benchmarking and the use of targets for attainment; performance indicators and an inspection regime (Ball, 2004). Although, not a new phenomenon in the measurement of student knowledge and skills, over time standardised assessment has become "a key instrument for policy reform" in education within developed countries (Mons, 2009, p. 5).

2.1.3 Accountability

As the third control mechanism, accountability ranges from the broad notion of a government's accountability to deliver a high standard of education (OECD, 2018), to the personal accountability of individual teachers in their classroom (Ryder, 2015). In short, accountability assigns responsibility for particular actions to individuals or groups involved in the process of change (Sahlberg, 2010). In an earlier categorisation of accountability which I suggest still holds true, Halstead (1994) used the terms "Contractual Accountability" and "Responsive Accountability". These terms distinguished between actions that tend to be more measurement-driven, demonstrating that students are learning what they are supposed to, and those actions which emphasised the educational process and decision making, giving and explaining an account of one's actions. Anderson (2005) argued that three main types of accountability system run simultaneously in the education field; the accountability for compliance with regulations, for adherence to professional norms and accountability that is results-driven. From a government perspective, accountability and the focus on specific measures is not only common sense but inevitable, given the financial investment in education, distribution and deployment of public resources (Winter, 2017; Sachs, 2016).

2.2 Education Reform Act and the rise of marketisation, standardisation and accountability

The political and economic arguments for marketisation in England's education system can be traced back to the 1970s (Whetton, 2009, pp. 138–139; Ball, 2007, pp. 17–19). The reimagining of public services with a commitment to the restructuring of education provision along market principles (Power, 1992) has been achieved through individual policy decisions that have centralised power upwards, to the Department for Education and devolved power downwards to schools and parents. Concurrently, the role and involvement of the Local Education Authority (LEA) in England has been downgraded. The first step in the marketisation of public education in England was considered to be the introduction of the Assisted Places Scheme by Margaret Thatcher's government in 1980 (Whitty *et al.,* 1998). Introduced through the 1980 Education Act, the Assisted Places Scheme provided public funding to enable academically able children from poor homes to attend some of the country's elite private schools, opening up parental choice (Whitty *et al.,* 1998). Seen as undermining the principles behind comprehensive education, the Assisted Places Scheme was quickly scrapped by the new Labour government when they came into power in 1997. In England, mechanisms of marketisation were brought into sharper focus by the introduction of the Education Reform Act in 1988. The introduction of Local Management of Schools (LMS) brought new roles, accountabilities, and management structures into the education landscape. The 1980s and 90s saw schooling reconfigured as a competitive quasi-market from which parents could choose from a variety of different options (Institute for Government, 2012). The education marketplace now included Specialist Schools, City Technology Colleges, Foundation Schools, Voluntary Aided and Community schools; and the continued existence of Grammar Schools under the Labour and Conservative governments (Long, 2015; Institute for Government, 2012; Taylor, 2003; Whitty et al., 1998). Open enrolment removed the restrictions imposed by the LEA on school size, allowing

schools to attract as many students as possible and bringing about wider parental choice (Whitty and Power, 2000; Whitty and Edwards, 1998).

The 1988 Education Reform Act (ERA) not only changed how schools were managed but also what was taught and how schools were measured. The National Curriculum standardised the education provision for five to 16-year-olds in England by offering a broad and balanced curriculum with a strong bias towards the basics of reading, writing, mathematical and scientific literacy (HMSO, 1988). Ensuring children's equal access to knowledge and reducing variation from poor teaching, the national curriculum has been repeatedly (1995, 1999 and 2008) revised, reviewed and slimmed down, new testing arrangements have been introduced and teachers given increased autonomy (Roberts, 2018; House of Commons, 2009). Some criticise the national curriculum, citing it as the enforced alignment of education provision around a core curriculum, "… a monument to central planning and control" (Cullingford, 2017, p. 45). However, the standardised curriculum and assessment, with all its benefits and drawbacks (Oates, 2011), remains optional in academies and free schools but mandatory for all other state maintained schools (Roberts, 2018).

With the national curriculum came national curriculum assessments, the history of which is multi-layered and multi-staged (Whetton, 2009). The ERA established the term "key stage" to represent groups of children across a specific age range and detailed the arrangements for integrated assessments of children against attainment targets in the curriculum. Although the responsibility for the tests moved through a series of statutory bodies, generating a lack of coherence in the system (Whetton, 2009), the effects of formal standardised testing have been wide-ranging. Standard Assessment Tasks, commonly known as SATs, were introduced to primary classes in 1989. These attempted to match elements of classroom-based assessment but were

found to be unmanageable in a mass testing system. Testing was introduced to secondary schools to measure student attainment in English, maths and science at the end of key stage 3 (KS3), but was later withdrawn in 2008 due to controversy over marking accuracy and teacher workload (Whetton, 2009). Science tests at KS2 were also removed in 2010, while reviews to national curriculum testing (e.g. Bew, 2011) continued to alter its emphasis and scope.

The outcomes of standardised tests generate information about student learning. In the UK, data on student attainment in state schools, particularly in the core subjects of English, maths and science have been collated and tracked by the Office for National Statistics since the introduction of national testing in 1988 (House of Commons, 2009). The publication of school performance tables, commonly known as league tables and introduced in 1992, shared this information more widely (Goldstein, 2001). Thereby, the outcomes of the end of key stage assessments provided the information to enable parents, local authorities and government to judge performance and make comparisons between schools. As a driver to raising standards overall, it has been debated whether the information is valid, objective and reliable (Goldstein, 2001). Nevertheless, a target culture followed which demanded ever-increasing proportions of students to reach the required benchmarks (Isaacs, 2010) with added scrutiny on the attainment of students from disadvantaged backgrounds (Jarrett *et al.,* 2016; Goodman and Burton, 2012). Student attainment data is at the heart of the accountability measures which use constant comparison of the outcomes from classrooms, schools and internationally to govern education (Ozga, 2013).

2.3 Present-Day Manifestations of the Three Mechanisms

This next section explores how the three mechanisms have evolved since the ERA and touches upon how each mechanism manifests within the English school system.

2.3.1 Marketisation of Education

Much has changed in the 30 years since the ERA came into being, however, two principles remain, that of raising attainment through a National Curriculum and the view of education as a commodified marketplace through which to drive up standards (Burgess *et al.,* 2015; Barker, 2008). With the responsibility for education in the UK now devolved to each constituent country, this discussion reflects the English education system. Wilkins (2015) suggested that of all countries adopting a neo-liberal, market-driven ideology, the English system has undergone the most extensive marketisation with an intensive performative regulatory framework. This section explores how aspects of marketisation have evolved and discusses the implications for different stakeholders.

Marketisation invites the involvement of profit-orientated business organisations into public-sector education; effectively letting the private sector run a public service funded by tax payer's money (Whitty *et al.,* 1998). In England, arising out of LMS and the apparent move towards privatisation of state education, is the expansion of what Ball (2009, p. 85) described as the plethora of companies eager to sell "improvement" and "innovative solutions" to schools struggling to achieve their targets or compete against other institutions. The changes to accountability measures, confusion over policy and the prospect of "failure" present themselves as business opportunities in which the private company acts as a saviour for the public good (Ball, 2009). These companies are based on market principles driven by the interests of business and not by the best interest of children (Stevenson and Wood, 2014). The wide range of education consultancies, foundations and publishers or

"edu-businesses" (Verger *et al.,* 2017) represent another face of education reform. Support previously offered by the local education authority or national strategy consultant to governing bodies, schools and teachers (DfE, 2011; Hatcher, 2006), is now offered at a price by companies covering everything from CPD for classroom practice (CPD for Teachers, 2018; IRIS Connect, 2018) to revision workshops (Hillcrest, 2017) and mock Ofsted inspections (Creative Education, 2018; Wright, 2013). New markets appear that did not exist before and often many other functions formerly routine within public sector organisations are now outsourced to profit-making companies (Gunter, 2017; Connell, 2013).

2.3.1.1 Marketisation: Schools and Teachers

Although diversity in school provision had been in existence since the 1980s, the academy model introduced by New Labour in 2002 built on the previous Conservative policy of the City Technology College (Institute for Government, 2012). Designed to shape social and economic outcomes by transforming underperforming schools in disadvantaged areas, newly built academies were independently managed schools with additional autonomy outside of LEA control, often sponsored by business or a charity. From 2010, the Conservative government widened the academies programme, allowing successful and failing schools to convert to academy status and enabling a direct partnership between business and education (Institute for Government, 2012). Introduced in 2011, the "Free School", new schools not previously LEA maintained (Roberts, 2016a), further opened the education market. Assumptions were that marketisation and the adoption of a business model would bring a high degree of autonomy and improve fiscal efficiency by freeing schools to manage their own finances, staffing, recruitment, and curriculum.

The market model allowed schools the independence and freedom to differentiate themselves as providers of education with unique features (Wilkins, 2012). Furthermore, schools, such as academies, with greater autonomy, assume authority over their admissions, with the potential to select the students who best enable them to raise attainment (Mansell, 2016). The academic performance and socio-economic make-up of the school can therefore act as important determinants of parent choice, within the limitations of the school's proximity to their home (Burgess *et al.,* 2015; Allen *et al.,* 2014). Where a school's performance fails to meet the required standards, parental choice is leveraged, resulting in the fall of pupil numbers, reduced school funding and the increased likelihood of inspection (Allen *et al.,* 2014; Stevenson and Wood, 2014). Despite education in the UK being amongst one of the most market-driven in the world, the privatisation of public education still creates controversy (Hicks, 2015) with no clear evidence that this model raises attainment for children (House of Commons Education Committee, 2015).

Individual teachers may find themselves employed in a variety of school organisational structures with differing levels of autonomy, depending on their relationship with central or local government (Long *et al.,* 2018; Roberts, 2016a). Therefore, teacher experience of marketisation can be conceptualised in different ways placing them as both object and subject. As objects of marketisation teachers can be viewed as a valuable commodity. A highly qualified physics graduate can be awarded a substantial bursary to commence Initial Teacher Training (ITE) (DfE, 2018a). Individuals can exploit the market; negotiating a better salary if they teach in a shortage subject (Vignoles *et al.,* 2018; Sims, 2017); take advantage of the opportunities to teach overseas (COBIS, 2018) or opt to teach in a leafy suburb with none of the perceived difficulties that come with working in a school based in a deprived area (Foster, 2018). Alternatively, teachers as the subjects of marketisation are expected to make a positive contribution to the wider life and ethos of the

school (DfE, 2013a). They come as part of the "package" on offer from the school, the smiling face seen on open-evening, the driving force behind raising attainment in the classroom (MacBeath *et al.,* 2004; Maguire *et al.,* 2001).

New structures which linked teachers' performance to their salary progression (DfE, 2018b; Marsden, 2015) further cemented the teaching profession into the frame of a market ideology driven by results (Stevenson and Wood, 2014). This performative policy technology affected teachers work from the outset (Sinclair *et al.,* 1996) by shifting teacher practice towards meeting the needs of government targets through using constraining performative instruments (Wilkins, 2015). Autonomy for schools and teachers alike is earned when the level of performance satisfies the imposed standards, enabling the school (and teacher) to remain in a strong position in the "market".

2.3.2 Standardisation in Education

The next section looks at how standardisation continues to frame discourse in education and is encompassed by the standardisation of assessment, attainment and, the standardisation of teacher practice and professional identity.

2.3.2.1 Standardisation of Curriculum and Assessment

The adoption of a national curriculum ushered in an era of standardisation across education, which continues to this day. It has been suggested that the revised 2014 national curriculum, with its narrow focus on the "core knowledge", facts and ideas that students are expected to master, was inspired by the work of E.D Hirsch (Steers, 2014). Within the formal curriculum, the academic subject-based curriculum organises knowledge, bringing together the current state of understanding in particular fields and setting out what is deemed appropriate for study by government

and learned societies. Consequently, whilst also representing knowledge that has been accumulated over time and providing continuity with the past (McCulloch, 2016), the standardised curriculum can be conceptualised as a vehicle for promoting change in education to meet the demands of society. It has been argued that a *hidden* curriculum exists and works covertly to reproduce social differences through the influence of culture, values and relationships (Hargreaves, 2011; Kentli, 2009). However, it can also be said that, the formal curriculum made statutory by government policy and in part driven by economic imperatives (Apple, 2012), overtly dictates the range and scope of what is accepted and valued as knowledge to be taught (White, 2016).

Standardised assessments can be regarded as an equitable means to compare the attainment of large cohorts of students across the range of socio-economic and ethnic groups (Stobart, 2008).   Assessments are also used at classroom level to judge teacher performance; at school level to hold schools to account and at macro-policy level to evaluate policy (Mons, 2009; Stobart, 2008). As a link between school and government to evaluate reforms, standardised assessments must therefore deliver across a range of objectives and act as an information source to others, especially parents.

2.3.2.2 Standardisation of Teaching Practices

Hailed as "one of the most ambitious change management programmes in education" (DfE, 2011, p. 2), the National Strategies introduced the pedagogical control mechanisms of the National Literacy and National Numeracy Strategies to primary schools and in 2001, the National Strategies for English, Maths and Science at KS3 to secondary schools. Advocating interactive whole-class teaching in the literacy hour and numeracy hour, the primary strategies intended to shape and standardise the classroom experience of children to raise attainment in literacy and numeracy (Smith

et al., 2004). Supported by a wide-ranging national infrastructure and employing 2,000 consultants, CPD and training materials (DfE, 2011; DfES, 2006a), the prescriptive structure of the Literacy Hour guided teachers to divide the lesson into three parts, ending with a plenary session to revisit objectives and reflect on learning (Machin and McNally, 2008).

Whilst not statutory, the nature of the inspection and accountability system and the added funding that came with them meant that schools felt compelled to adopt the National Strategies direction. Inevitably, evaluation found that the strategies had altered classroom practice, although the ability to foster learning, through the one size fits all approach, was less clear (Smith *et al.,* 2004; Earl, Watson, Levin, *et al.,* 2003; English *et al.,* 2002). The changes brought about by the NLS and NNS frameworks endured through their successful implementation which relied on teachers taking on board the training and applying the given materials consistently (Ofsted, 2008; Webb and Vulliamy, 2006). Whole-class teaching, structured lessons and the use of objectives to plan learning are elements of classroom practice that teachers have adopted through their engagement with the national strategy frameworks (Gill, 2017). At present, the standardisation of teaching may appear to be less prescriptive, with schools now guided to refer to evidence-based, "what works" practice through government documents such as the Literacy and Numeracy Catch-up Strategies (DfE, 2018c) or to organisations such as the Education Endowment Fund (EEF, 2018a), but teachers in state schools are subject to a set of professional standards.

2.3.2.3 Standardisation of Teacher Identity

Teacher Standards (DfE, 2013a) require all teachers to demonstrate continued development of knowledge, skills and understanding. The document also lists the expected values and behaviours that must be demonstrated throughout a teaching

career. This move toward the standardisation of teachers' classroom behaviour and pedagogy (Sachs, 2016) is not new; the first standards for qualified teacher status (QTS) (DfEE, 1998) were introduced in 1998. The OECD (2005) considered teacher quality to be the single most important variable in influencing student attainment so it is understandable why many governments pursued a means to systemise this variable.

Teachers have become the targets of government-initiated professionalism-shaping mechanisms with the current model of "*demanded* teacher professionalism" (Evans, 2011, p. 852) heavily focussed on the behavioural aspects of teachers in meeting the government standards-based reform agenda. To adverse effect, the de-professionalisation of teachers is enabled through the performative axis of managerialism, appraisals and performance-related pay linked to student outcomes (Stevenson and Wood, 2014; Ball, 2003). Viewed as detrimental to collegiality and collaboration crucial to core professional values (Wilkins, 2011) the standardisation creates compliant teachers, coerced into spending their time and energy chasing targets.

Alternatively, judging teachers against a pre-established, peer-reviewed, and evidence-based system of outcome measures (Coates, 2011) is thought to bring a greater degree of equity ensuring that the children receive quality teaching irrespective of the school context, regulating their experiences and rooting out underperforming teachers (Evans, 2011) so that the attainment targets can be rightfully achieved. Subsequently, the standardisation of the teaching profession and the generation of frameworks to measure teacher effectiveness may generate positive outcomes (Coe *et al.,* 2014). Forde (2016) suggested that teachers need to become skilled in the use of standards as reflective and planning tools for exploring

their contextualised practice rather than as part of a reductive, quality assurance process.

It has also been suggested that over time, an understanding of what constitutes a performative education environment will no longer be debated with such a hankering for the pre-target orientated times (Wilkins, 2011). The younger teachers entering the profession have only ever experienced education as associated with performative structures; end of key stage tests, target grades, high-stakes assessment and Ofsted inspections (Wilkins, 2011). Fully embracing the performative culture, the young newly qualified teachers enjoy the micro-autonomy of the classroom yet are comfortable with the wider performative framework and accountability culture seeing it as an effective way to raise student attainment (Wilkins, 2015).

In England, it is evident that standardisation has been applied widely in education. Policies that outline the expectations for curriculum coverage, student assessment and attainment, teacher practice and professional behaviours are embedded into the education context and discourse. Winter (2017) argues that prescriptive practices, standardisation and rigour hinder other ways of being and thinking. However, alignment with the given standards is ensured through the prescribed accountability measures, performance tables and school inspection; this is explored in the next section.

2.3.3 Accountability within Education

The term accountability can be applied across many contexts and in England's schools it may be characterised through a number of different relationships between the school and local authority; an academy and central government; the headteacher and the schools' governors, and an individual teacher and their students' learning (West *et al.,* 2011). For the government, accountability is a means to justify the

spending of public money (The Children, Schools and Families Committee, 2010).
This section considers how accountability mechanisms enshrined in education policy
support the continued drive towards a standards-based system.

2.3.3.1 Accountability Measures

School accountability measures make a difference to practice and as such the
measures that the government policy chooses to value will have major implications
for schools, teachers and students (Leckie and Goldstein, 2017; Acquah, 2013).
Predominantly linked to examinations at the end of KS2 and KS4 (GCSEs), the
publication and wide reporting of a school's performance form the core of the
accountability framework (DfE, 2019a). The different metrics used to measure
absolute attainment and progress have been found to increase the pressure on
schools to adopt a range of practices to maximise attainment (Acquah, 2013). For
instance, teaching to the test (Stobart, 2008), judicious selection of the
qualifications taken (Parameshwaran and Thomson, 2015), GCSE early entry (Taylor,
2013) and the manipulation of controlled assessment marks (Ofqual, 2012) or by off-
rolling poorly performing students (Nye, 2017). Poor performance against these
metrics has wider consequences for the school and community, possibly triggering an
Ofsted inspection and ultimately leading to school closure or forced academisation.
Furthermore, parents may compare the school unfavourably with neighbouring ones,
leading to a reduction in cohort numbers and therefore a reduction in funding
(Roberts and Bolton, 2016).

Accountability measures use literacy and numeracy as proxies to represent learning
and progress in primary education (Alexander and Armstrong, 2010), which as
research has shown narrows the curriculum (House of Commons Select Committee,
2008). Whilst the importance of literacy, as a foundation for learning in all subjects,

is undisputed (DfE, 2015a; Morrisroe, 2014) evidence suggests that the curriculum

focus on standardised assessments in the core subjects, literacy and numeracy at

KS2, reduced the available time for other subjects including science (Wilshaw, 2016;

Wellcome Trust, 2014) and has done for some time (Harlen, 2012; Burton, 2010;

Boyle and Bragg, 2005).

2.3.3.2 Accountability and Inspection

Although educational accountability exists in other countries England's regime is

regarded as one of the more extreme due to its scale and the far-reaching intended

and unintended consequences that result (Stobart, 2008). "Policing" this system

with the responsibility to ensure that all schools are inspected regularly through a

rigorous and transparent process (Elliott, 2012) is the *Office for Standards in*

*Education (*Ofsted). To measure the overall effectiveness of schools the new

inspection framework, applicable from September 2019, judges schools in four areas,

"quality of education", "behaviour and attitudes", "personal development",

"leadership and management". Quality of education is an holistic combination of the

previous strands, "teaching, learning and assessment" and "outcomes" (Ofsted,

2019). Critics remark on the negative impact of Ofsted inspections on teachers'

well-being (Barton, 2015, Education Support Partnership, 2015) more significantly

researchers question the positive impact that inspections have on raising students'

attainment. Doubts about its effectiveness have long been debated here and abroad

and although there was some improvement in the attainment of females and

selective schools, Ofsted inspection had no positive effect on examination

attainment for maintained schools, in fact, it is more likely to result in negative

effects (Altrichter and Kemethofer, 2015; Cullingford and Daniels, 2013; Rosenthal,

2004; Shaw *et al*., 2003).

2.4 Datafication in Education

The impetus for marketisation, standardisation, and accountability in education has been variously but widely argued as the means to raise student attainment (DfE, 2010a; Sahlberg, 2006). Linked to funding, pupil numbers and accountability measures, student attainment data creates the metrics through which government policy is driven (Winter, 2017; Lingard, 2011). From this perspective, this section describes the purpose of student attainment data and how this is woven into and supports each mechanism.

2.4.1 Uses of Data

The data collected on or about students' attainment is varied, wide-ranging and may serve several purposes (Stokes, 2016). Shared with a variety of stakeholders, both in and out of the classroom (Kelly and Downey, 2011) attainment data has been used for example to explore differences in groups of students by their characteristics (Black-Hawkins *et al.*, 2017; Jerrim, 2017), carry out question-level analysis of test papers (Thomson, 2015, 2019) and to monitor school and teacher performance (Perry, 2016). Although not straightforward (Strand, 2014a, 2015), attainment data can also be used to identify need, allocate resources, funding and support (Goldstein, 2001).

Drawing attention to the differences in student attainment also serves to identify which students are underperforming and highlights any attainment gaps. Defined as the "disparity in performance on various educational measures between different groups" (Goodman and Burton, 2012, p. 500), attainment gaps are persistent and widen as a child moves from primary to secondary school (Andrews *et al.*, 2017). Among the complex combination of factors that determine outcomes, it is difficult to isolate the exact causes of the attainment gaps between students of different socio-economic status, ethnic group or gender. Having said this, tackling attainment gaps

between different groups of students remains a priority for schools and has been for some time (EEF, 2018b; Sharp *et al.,* 2015; Wilson, 2014).

In addition to identifying the attainment gaps between different groups of students, student-level attainment data provides some of the information required to judge school performance. However, attributing the improvements in student attainment to specific actions has not been straightforward (Goldstein, 2001), any change in attainment could be associated with a range of different factors including teacher intervention, changes to assessment methods or changes in cognitive demand of the examination papers. The steady rise in KS2 English and maths results from 1995 to 2003 and the changes in the variation in the grade boundaries between the different level outcomes are a case in point. In this instance, the impact of the test technique and teaching to the test were found to be contributory factors in the rise of standards (Tymms, 2004). Similarly, Gove (2009) suggested that the continued rise in GCSE attainment was driven by secondary schools "gaming the system" in order to meet to reach their 5A*-C GCSE targets. As discussed earlier, key accountability measures linked to student attainment and used to judge school performance may directly impact school exam entry practices.

Student attainment data also plays a part in the performance management process of teacher appraisal and evaluation. Contrasting arguments suggested that, even where value-added measures were taken into account, student examination outcomes were not reliable measures of teacher effectiveness (Goldstein, 2001). However, despite its limitations, when compared to classroom observation or pupil survey, student attainment data was argued to be the best way to judge teacher effectiveness (Murphy, 2013). There are inherent difficulties in measuring the progress of students year on year as they move from one teacher to the next or from one key stage to the next (Wilkins, 2011; Goldstein, 2001). Furthermore, the inclusion of student

attainment data in the decisions around performance-related pay can also distort teacher practice and narrow the curriculum experiences for students, in addition to having a divisive impact within the school environment, (NASUWT, 2016; Marsden, 2015; Baker *et al.,* 2010).

2.4.2 Data and the STEM Agenda

The analysis of student attainment data can play a significant part in monitoring the progress of students in particular fields of study and the participation of students in the STEM subjects has been of interest to the government for some time (Hyam, 2006). Since science and innovation were at the heart of the UK's economic plan, the intention to close the STEM skills gap was addressed by taking direct action to increase the uptake of single sciences at GCSE and A-level sciences (Archer, 2015; HM Treasury, 2014). The government announced that, from September 2008, all 14-year-olds achieving a Level 6 score in the KS3 national tests would be entitled to follow a triple science GCSE course (Tomei *et al.,* 2015). However, numerous factors whether at school-level, involving parental science capital or gender role models contribute to student's decisions to continue studying science post-16 (DeWitt and Archer, 2015; Bennett *et al.,* 2013; Hampden-Thompson and Bennett, 2013).

The long-term wider goals for STEM, to raise the number of students studying sciences, used changes to examination and curriculum as its principal drivers. The coalition government used school performance and accountability to drive widespread curriculum reform, as seen with the introduction of the new national curriculum with more demanding content and the revised A-level specifications (HM Treasury, 2014). National data sets pointed to the year on year increase in the number of students achieving A and A* grades in GCSE sciences with the resultant conclusion that the exams were getting easier (Gove, 2011). The Royal Society

review of science examination papers across a three-year period confirmed that there was insufficient challenge for the more able students (Osbourne, 2011). As such, more rigour was introduced, multiple-choice examination papers withdrawn, opportunities to retake exams reduced and certain vocational qualifications were no longer given GCSE equivalent status (e.g. BTECs). The resulting changes to the accountability measures published in the performance tables have had an impact in the subjects that students take at GCSE (Parameshwaran and Thomson, 2015). Structures of control have been put into place through the accountability framework that requires schools to manage their curriculum offer in science. Whether the changes to the curriculum will ultimately lead to a greater number of students joining the STEM pipeline remains to be seen.

2.4.3 Data and International Comparisons

It was England's declining performance in the PISA tests in 2012 that was to have a direct impact on government education policy as evidenced in the speech given by Michael Gove, the then Secretary of State for Education, to the House of Commons (DfE and Gove, 2012). Through "policy borrowing", Gove set out to replicate in the UK the educationally successful structures established in other countries. Something that Alexander (2012, p. 5) considered to be a "naïve belief".   Using the PISA results, Gove claimed that England was falling behind South East Asian countries, "leaving our children behind in the global race" (DfE and Gove, 2012). The reforms that he outlined in the speech have to some extent been adopted and have not only changed the curriculum but the accountability frameworks in England too. Jerrim (2013) demonstrated that the alarm bells raised over England's failing education system were unwarranted, as the evidence lacked the necessary strength to justify the sweeping changes that the Conservative government introduced. Amplified by the media, the panic resulting from England's perceived poor performance in the

PISA tests is symptomatic of the keen interest in its outcomes, this lays the data open to misinterpretation and misuse (Tomei *et al.,* 2013). Furthermore, others have challenged the validity and reliability of PISA scores, particularly the way in which they have been used as a measure of a whole cohort learning across all subjects rather than merely a sample of 15-year-old's performance in three subjects (Meyer and Benavot, 2013; Alexander, 2012).

The current education environment is one in which data is an integral tool to a teacher's knowledge of their students but is also used as a tool by a range of stakeholders with differing agenda. It is used in policy decisions fundamental to the instigation of many of the changes that we have seen in education as a whole. As a means of control, student attainment data and the accountability measures that report the outcomes of standardised tests, determine to some extent what happens in the classroom (Winter, 2017). The prime use of student data in this study is to examine what it reveals about the changes in science attainment over time and to compare the outcomes of students with different background characteristics in response to the changes to curriculum and assessment.

2.5 Summary

It is evident that the collective responsibility for student learning is impacted by accountability, standardisation and marketisation. The current neo-liberal, neo-conservative cultural framing of education shifts the focus from a collegiate model of schooling to one in which the individual actor or institution competes with others. Sharing good practice across schools, departments or individual teachers is replaced by a reworking of professionalism and an erosion of trust (UNESCO, 2018; Ball, 2016). Additionally, it can be conceived that by perpetuating a system of quantifying output to make comparisons against external standards (Bol and Van de Werfhorst, 2011),

standardisation and marketisation reduce teacher autonomy and narrows the focus of learning.   As such, the prescribed curriculum aligned to national tests and accountability measures defines success in terms of high-stakes tests with the overall "flattening of complex human and social processes into crude representations" (Ball, 2004, p. 17), creating an environment that challenges creativity and results in risk-averse teachers and pupils (Robertson, 2015). A high level of control in a complex system such as education does not always result in the intended outcomes, principally because it relies upon the integration of many different factors; people, place, and space (Cullingford and Oliver, 2017). The next chapter looks in more detail at the development and reform of the science curriculum in schools over the past 30 years and briefly outlines the current profile of the students to whom the curriculum applies.

Chapter 3: Science Education in England

Introduction

This chapter begins by putting forward arguments that briefly explore the place of science as an academic subject in the curriculum, locating the major drivers of reform within the socio-political context. A general overview of the structure of national curriculum is given along with a brief outline of previous reforms to the science curriculum from 1988 to present day. This is followed by a more detailed description of the most recent 2014 reforms that have changed the ways in which school science is assessed. Before the chapter summary there is a final section that describes the student population for whom the science curriculum is a compulsory element in their schooling.

3.1 A Brief Overview of the Socio-Political Influences in The Development of The Science Curriculum

Science matters because it provides a means to explore ourselves and our world in a systematic way, building an understanding that diminishes past superstition and enabling decision-making to plan for a better future (Millar and Osborne, 1998). However, debates around science in the curriculum have been taking place for some time, with the importance of science education having gained momentum after the Second World War (Donnelly and Ryder, 2011; Waring, 1979). For example, the 1950s and 60s saw the Cold War and "space race" between the US and the Soviet Union that highlighted the link between scientific advances and perceived national power. This competitive drive marshalled efforts to raise the profile of science education, which in the England at the time focussed on preparing grammar and public-school students for further study of science at university and entry to professional occupations. For students attending secondary modern or technical schools the locally determined curricula prepared them for "practical" applications of science (Donnelly and Ryder, 2011). During the post-war period, dissatisfaction with science

education was being expressed not only in the House of Commons but also by teaching unions such as the Science Masters' Association, Federation of British Industry, learned societies (Royal Society of Biology (RSB), Royal Society of Chemistry (RSC) and Institute of Physics (IOP)), and universities (Waring, 1979). Moreover, all reform takes place through the ideological lens of whichever political party is in power at the time (Smith, 2018), so whilst science courses such as Nuffield and Salters had been developed for school certification (Bennett et al., 2005; Waring, 1979), it was the change to the comprehensive school system and the replacement of the two-tier certification, Ordinary Levels (O-levels) and Certificate in Secondary Education (CSEs), at 16 which heralded significant standardisation of science education (Donnelly and Ryder, 2011). The newly introduced national curriculum brought with it the new General Certificate of Secondary Education (GCSE) qualification and the positioning of science as a core subject.

From an historical perspective, it can be suggested that in England, the start of industrialisation in the late eighteenth century and the increased use of technology to generate wealth opened the way for scientific innovations (Bell and Skiebe-Corrette, 2016). Within the various debates that link science, education and economic growth (e.g. Billimoria, 2017; Weinstein et al., 2016; Hanushek et al., 2008), it has been argued that the creation, distribution and exploitation of knowledge with the rapid application of scientific advances are important drivers of economic performance (OECD, 2000). To this end, the Labour government under Tony Blair, in their plan for growth, placed science, innovation and research at the heart of a policy aimed at achieving global economic competitiveness (HM Treasury, 2004, 2014). The ambitious goals set out to support the supply of scientists by improving the quality of science teachers and lecturers at every level; raise attainment in science GCSEs; increase the proportion of minority ethnic and women

participants studying science in higher education. Additionally, the ambition was that the UK would maintain and build on its centres of research excellence; develop greater responsiveness between research and the end user; increase awareness of scientific research and innovation in society as a whole and, see an increase in UK business investment in Research and Development (R&D) from 1.25% of GDP to 1.70% (HM Treasury, 2004).

Knowledge in science is always advancing and new knowledge being uncovered (Holman, 2018; Royal Society, 2017), thereby necessitating on-going review of the science curriculum over time (Wong, 2019). For instance, advances in genetic modification and cloning in biology; the structure and uses of Buckminster fullerenes and nanotechnology in chemistry are now established topics in the GCSE specifications (e.g. AQA, 2019). Alongside the advancements taking place within the academic disciplines of science, there were argued to be three distinct aims for school science education: 1) Enhancing student interest by promoting wonder and curiosity. 2) Supporting the development of 'scientific literacy'. 3) Preparation for further scientific study (Millar and Osborne, 1998, p. 12). Each of these threads has been represented in subsequent versions of the science national curriculum in different ways (Wong, 2019).

In addition to the government departments with a specific remit for education e.g. DfE, Ofsted and Ofqual, other stakeholders have vied for the right to influence the science curriculum (Bell and Skiebe-Corrette, 2016). These have included the learned societies (RSB, RSC and IOP mentioned previously) under the banner of SCORE (Science Community Representing Education based at the Royal Society of Chemistry); industry partners (e.g. BAE Systems, BP and Vodaphone) in the form of Project ENTHUSE (EdComms, 2016); charitable research organisations such as the

Wellcome Trust, Sutton Trust and the Education Endowment Foundation (EEF) and eminent academics from UK universities (e.g. (Harlen, 2012; Oates, 2011; Alexander and Armstrong, 2010; Osborne et al., 2003). Ofqual and its predecessor, the QCA, have conducted reviews of GCSE coursework and controlled assessments which directly impacted the way in which practical science is assessed (Ofqual, 2013; QCA, 2005a). Similarly, Ofsted published its findings and recommendations for science education prior to the introduction of the 2014 curriculum (Ofsted, 2013). The Wellcome Trust made recommendations for primary school science in its various reports (Wellcome Trust, 2011, 2014, 2017), and in response to the proposals for the newly introduced science curriculum (Wellcome Trust, 2013). Furthermore, SCORE produced numerous detailed responses to government proposals for the 2014 curriculum, which reflected their vision for science education (SCORE, 2013d, 2013c, 2013b, 2013a). It appears that there are competing demands on the purpose for science education creating a contested arena for school science in which tensions exist between multiple stakeholders with multiple aims (Wong, 2019). The different directions advocated by each stakeholder shift between one with an emphasis on the traditional academic content. This position supports the drive towards increasing the take-up of science post-16 to meet the need for a scientifically knowledgeable workforce. The alternative position emphasises the socio-political and ethical aspects of science, with the aim to make the subject relevant to all students irrespective of their future intentions (Ryder and Banner, 2011).

Referring to the KS4 science curriculum, Ryder and Banner (2011) suggested that the 2006 iteration was heavily influenced by the Nuffield Foundation funded project work of Millar and Osborne which resulted in the Beyond 2000 report (1998). These reforms were the first to introduce the terms "How Science Works" and "scientific literacy" to the science curriculum and, included an emphasis on developing

student's decision-making skills in the face of contradictory scientific data and in exploring the social, political and ethical issues relating to science in the real world (Donnelly and Ryder, 2011; Zolle, 2006).   In November 2010, plans were confirmed that a complete revision of the national curriculum would take place with changes to KS2 assessment alongside reforms to the assessment of the curriculum through the GCSE and A-level qualifications (Long, 2017; Roberts, 2016b). Consultation took place between the DfE and various stakeholders including teachers, schools, local authorities and university academics (DfE, 2013b, p. 2). Attracting a mixed response, the first draft of the proposals was issued in January 2012, arguments that the new curriculum was too "fact-based" were levelled and respondents to the consultation on science were concerned that the increased content would lead to superficial learning (DfE, 2013b). As a result, the whole-scale reform was pushed back and a new national curriculum published in 2014 (Roberts, 2018). However, Wong (2019) argues that the learned societies and professional scientists (SCORE) were most influential in deciding the content of the 2014 science curriculum. Wong (2019) concluded that SCORE were instrumental in increasing the scientific rigour and maths content of the GCSE curriculum with the aim to better prepare students for the study of science at A-level and beyond.

In primary science, the topics remained relatively unchanged apart from the introduction of a new topic on evaluation and inheritance in Year 6, and changes to scientific enquiry – now reframed as "working scientifically" (Naylor, 2013). For GCSE sciences, the curriculum reform came into force for year 10 pupils in September 2016 (DfE, 2014b) with the three subjects Biology, Chemistry and Physics made more distinct and, complex mathematical data handling, graphical and algebraic skills expected (Wong, 2019; SCORE, 2013b). This chapter adds detail to the description of the current education structure set out in chapter 2, with a particular reference to

science, I outline the nature of the curriculum and its assessment, and begin by identifying the student populations this applies to.

3.2 Overview of the National Curriculum

The Department for Education (DfE) simply explains the National Curriculum (NC) as a set of subjects and standards used by primary and secondary schools so children learn the same things, covering the subjects taught and standards that children should reach (DfE, 2014a). The ERA divided compulsory schooling into the now familiar Key Stages (KS). Key stages 1 to 4 cover compulsory education in schools, Key Stage 5 corresponds to non-compulsory education in schools and colleges (see Table 1).

Table 1: School Key Stages by Age

Key Stage	Age range	School Years
Key Stage 1	5 to 7	1 and 2
Key Stage 2	7 to 11	3 to 6
Key Stage 3	11 to 14	7 to 9
Key Stage 4	14 to 16	10 and 11
Key Stage 5	16 to 18	12 and 13

The ERA also designated curriculum subjects to be either Core or Foundation. Foundation subjects e.g. Humanities, Arts, Design Technology and Modern Languages are taught up to and including aged 14, end of KS3. The core subjects, English, mathematics and science are compulsory and taught from Key Stage 1(KS1) up to age 16, the end of KS4 (DfE, 2014a). This too has remained largely unchanged over the years except for computing, which replaced Information and Communication Technology (ICT) (Appendix Table A, p320).

The foundation subjects are not compulsory at KS4 but all students have a statutory entitlement to be able to study a subject in each of those four areas. This

entitlement has begun to feature more heavily in the light of the changes to the accountability measures, of which more will be discussed later. The demise of the national curriculum tests (also known as SATS) for 14-year olds at the end of KS3 in 2008, brought about a change to this phase of education in secondary schools. Teacher Assessment became the main form of statutory assessment, using the scale of attainment levels, numbered 1 to 8. The requirement for schools to publish their science KS3 Teacher Assessments ended in 2013, and the use of attainment levels was removed and not replaced (DfE, 2013c). There is now no formal acknowledgment of KS3 as representing a significant phase in a child's schooling. Schools are free to develop their own approaches to the 11-to-14 curriculum. Instead of following the traditional 3-year period, schools have the freedom to adopt a 2-year KS3 (Noden *et al.*, 2007, p. 4; DfES, 2006b), thereby creating the option to commence study of the KS4 curriculum in Year 9.

The National Curriculum Framework document also sets out the corresponding Attainment Targets (the knowledge, skills and understanding) and Programmes of Study (the matters, skills and processes) which pupils of different abilities and maturities are expected to have by the end of each Key Stage. These are closely tied to the national curriculum tests, assessments and GCSE examinations. As such the national curriculum for science aims to ensure that all pupils:

- develop **scientific knowledge and conceptual understanding** through the specific disciplines of biology, chemistry and physics
- develop understanding of the **nature, processes and methods of science** through different types of science enquiries that help them to answer scientific questions about the world around them
- are equipped with the scientific knowledge required to understand the **uses and implications** of science, today and for the future.
(DfE, 2014a, p. 168)

In addition to specific teaching of science content, teachers are expected to use their lessons to develop student's literacy and numeracy skills; as these are seen as preconditions to students' success across the entire national curriculum.

Although this study focusses on the most recent changes subsequent to the introduction of the revised national curriculum in 2014, the table below (Table 2) briefly sets out the changes to the science curriculum since the 1988 Education Reform Act made this a core subject with associated high-stakes examinations (see also Appendix Table B (p321). A more detailed account of science in the curriculum can be found in Black (1995) and Donnelly and Ryder (2011).

Table 2: Summary of the Major Reforms to Science in the National Curriculum

Date	Reform
1988	National Curriculum for science introduced consisting of 17 Attainment targets (AT) and 409 statements in the Programme of Study (PoS).
1989	Science curriculum revised, and the number of ATs reduced to 5 with 178 PoS statements.
1991	Science curriculum reduced to 4 ATs. Attainment Levels (1 to 10) introduced in KS1, KS2 and KS3. Double and single science introduced at GCSE.
1995	Levels for KS1 to KS3 run from 1 to 8 with exceptional performance introduced at for the end of KS3.
2002	GCSEs in Applied Science introduced and count as two GCSEs
2004	KS1, KS2 and KS3 remain unchanged but KS4 is restructured with "How Science Works" replacing the Scientific Enquiry strand. Single or Double science now replaced by GCSE Science and GCSE Additional Science or three separate sciences.
2006	A wider range of GCSE Science qualifications introduced, schools can no longer disapply student from science and all students have an entitlement to study two GCSEs.
2009	Modular GCSEs introduced.
2010	KS2 science end of key stage tests end.
2011	Wolf Report (2011) changes the weighting and value of vocational qualifications including Applied Science. GCSE examined 75% externally examined and 25% internal controlled assessment.
2014	New, more rigorous science curriculum introduced for first teaching in primary school.
2016	New science curriculum for KS4 introduced for first teaching to year 10. 100% externally examined.

The next sections explore the changes to the national curriculum and science attainment from 2008 onwards.

3.3 Changes at Key Stage 4 from 2008

3.3.1 Accountability Measures

The General Certificate in Secondary Education (GCSE) represents the culmination of the study of the KS4 national curriculum for 16-year olds. Student performance in GCSE examinations is published using headline figures which summarise absolute attainment and relative progress from KS2 (Leckie and Goldstein, 2017). The measures which counted the percentage of students who achieved 5 or more A*-C grades and the percentage of students who made Expected Progress in English and maths remained unchanged up to 2016 (Roberts and Bolton, 2016). The recent revision to the national curriculum has led to the introduction at KS4 of two new headline progress measures: Progress 8 and Attainment 8 replacing 5A*-C E&M, these are used to judge secondary schools in addition to other measures as follows

- the % of pupils being entered for and achieving the Ebacc[1] measure;

- the % of pupils achieving grade 5 or better in both English and maths GCSEs

- the % of pupils staying in education or employment after KS4.

Attainment 8 is a point score calculated from a pupil's best eight grades across three subject-based categories. Progress 8 compares a pupil's Attainment 8 score to the national average for pupils who scored the same in English and maths tests at primary school. A school's results are the average across all its eligible pupils (DfE, 2019a).

The EBacc is not a qualification in its own right, it is the name given to a group of curriculum subjects. As a new government accountability measure, it records the percentage of students who have followed and gained qualifications in a prescribed set of academic subjects, which are: English, Mathematics, Science,

---

[1] Ebacc The English Baccalaureate performance indicator linked to student entry and achievement in a core number of GCSE subjects

History/Geography and a Modern Foreign Language (MFL). Citing the decline in the take up of MFL and science, the need to reduce the attainment gap between richest and poorest and desire for more performance measures and publicly available data, the EBacc measure, though not compulsory, acts as a mechanism for the government to manage the school curriculum and students' subject choices (Long and Bolton, 2017).

3.3.2 Curriculum Content

A simple comparison of the number of statements in the 2007 National Curriculum Programme of Study (PoS) for KS4 with those in the 2014 Programme of Study demonstrates the increase in subject content. For instance, Biology lists one subheading, "Organisms and Health", with five statements in 2007. Whereas, the 2014 PoS gives a more detailed breakdown, spread across seven sub-headings and 45 statements. The terminology used to describe experimental and practical science also changed from "How Science Works" (HSW) in 2007 to "Working Scientifically", with subsequent changes in emphasis. The 2007 curriculum contained statement 4b

> to consider how and why decisions about science and technology are made, including those that raise ethical issues, and about the social, economic and environmental effects of such decisions. (QCA, 2007, p. 223)

Whereas, the discuss of ethical implications for science is missing from the 2014 curriculum document, a new statement reads,

> explaining every day and technological applications of science; evaluating associated personal, social, economic and environmental implications; and making decisions based on the evaluation of evidence and arguments (DfE, 2014c, p. 5).

Although revision of the national curriculum affected all key stages, the attainment at KS2 and KS3 are no longer formally published for individual schools and students, therefore the table below (Table 3) focusses on the changes to the KS4 science PoS and compares the number of detailed statements contained in the 2007 and 2014 national curriculum for science that reflect the increase in subject content at GCSE.

Table 3: Comparison of National Curriculum Science statements for KS4 2007 vs. 2014

	2007[1]		2017[2]	
	Sub-headings	Statements	Sub-headings	Statements
HSW/WS	4	14	4	28
Biology	1	5	7	45
Chemistry	1	4	8	37
Physics	2	7	9	41

(1:QCA, 2007; 2:DfE, 2014c)

There are now 28 statements, under four new subheadings within the Working Scientifically (WS) strand in 2014, where previously there were only 14 statements across four HSW strands in 2007. Significantly, the 2014 curriculum specifically details the requirements for the use of appropriate vocabulary, units, symbols and nomenclature associated with the increased demands in mathematical skills written into the curriculum, e.g. calculating the concentrations of solutions from the mass of solute and volume of solvent (DfE, 2014c, p. 11). The increase in the number of detailed statements can be seen across all three sciences, Chemistry and Physics increasing from 4 to 37 and from 7 to 41 respectively.

3.3.3 Grading

Students in KS4 opt to study their preferred foundation subjects at GCSE in addition to core subjects English, maths, and science. Previously, the grading of GCSE qualifications ranged from A* the highest to G, the lowest qualifying grade with a U grade signifying Ungraded. The letter grades changed to a numbered system from 2017, from the highest grade 9 to lowest grade 1 (Ofqual, 2016). In line with government ambitions to raise standards, a C grade was previously deemed a good pass at GCSE, its new equivalent is a grade 4 and is termed a "standard pass" (see Figure 1). The level of knowledge required to achieve a grade 4 in the new science

GCSEs represents the bottom to middle of the old grade C. A grade 5 used in the accountability frameworks is termed a "strong pass" (Long, 2017) and within this grade, there is overlap between the old and new grade boundaries which constitutes a level of attainment nearer to the demands of the old B grade (Ofqual, 2014). It is expected that the new grades 8 and 9 would provide a greater degree of discrimination at the top of the grade range than the previous A* grade.

Source: (Ofqual, 2016)

Figure 1: Diagram giving a comparison of old GCSE grades vs. new GCSE grades

3.3.3 Trends in Attainment

Since their introduction, and despite the variability in the percentage of students achieving A*-C grades, GCSEs have been criticised for perceived grade inflation partly due to the continued improvements in the number of students achieving the highest grades (Baird et al., 2013, 2019). It was the strongly held view that GCSEs were too easy (Gove, 2013; Shepherd, 2010) that led in part to their overhaul. As a

result, a range of reforms were introduced post-2010 which meant that the GCSE coursework component was replaced by controlled assessment and any modular examinations were replaced by linear terminal examinations at the end of a two-year course (Long, 2017). From 2013/14, reforms were also implemented which altered the calculation of KS4 performance measures, these changes impacted on the subject entries at GCSE. Firstly, Professor Alison Wolf's Review of Vocational Education (Wolf, 2011) prevented any qualification from counting as larger than one GCSE and capped the number of non-GCSEs included in performance measures at two per pupil. This prompted a move away from the vocational qualifications in science which no longer counted in the performance tables (DfE, 2019b; Allen and Thompson, 2016). Secondly, for subjects counted in the Ebacc (Burgess and Thomson, 2019) the early entry policy only counts a student's first attempt at a qualification and placed restrictions on the opportunities for retaking exams and on how schools used the subsequent data resulting from any improvement.

Prior to 2016, the suite of GCSEs in science covered a range of certifications including Double Award, Core Science, Additional Science, Applied and Additional Applied Science, Biology, Chemistry, Physics, Environmental Science, Geology, Astronomy (Gill, 2012), BTEC Level 1 and 2 qualifications (Pearson, 2018). The number of GCSE entries in science has changed due to the introduction of the EBacc accountability measure (see Table 4). The EBacc only recognises the new Double Award Combined Science, which replaced the Core and Additional Science qualifications, or separate sciences as counting towards the student's Progress 8 or Attainment 8 scores (DfE, 2019b). The 2016 science specifications go further, eliminating controlled assessments altogether, replacing the assessment of How Science Works investigative skills with the requirement to conduct 12 pre-set practical activities which are assessed via the terminal exams at the end of year 11,

these make up 15% of the overall grade (Ofqual, 2015a). For example, in GCSE

Combined Science, there are two tiers of entry (Higher and Foundation) available

across six externally examined papers, each over an hour long (AQA, 2019; Pearson,

2019).

Table 4: Number of Entries to Science GCSE courses between 2008-2018 all schools in England

GCSE Route	Mean No. Entries 2008-2017	Nº· Entries 2018
Core	389,886	739,406
Additional	300,672	
Biology	124,630	159,516
Chemistry	121,443	156,224
Physics	120,524	154,768

Sources: (DfE, 2016a, 2019c; DCSF, 2009, 2010) See Appendix Table N, p365

In 2018, there were over 739,000 entries to the new Double Award Science, each

entry equated to two GCSEs (DfE, 2019b) and represented an increase of around 7%

on the mean number of students (690,558), who had taken the Core and Additional

route combined for the period 2008 to 2017. There was also an increase in the

number of students following the triple science route. The average number of

student entries to separate science courses between 2008 and 2017 was 122,199,

however, this rose to over 156,836 in 2018 an approximate increase of 28% across all

three sciences. Overall in 2017/2018, 68.0% of students entered the combined

science GCSE and 27.5% took triple science (DfE, 2019b).

Until 2017, a good GCSE pass grade covered the range from A*-C and the data in

Table 5 uses the Statistical First Release data and shows the year on year trends in

the percentage of students who achieved an A*-C in science courses between 2008

and 2017. Generally, the percentage of students who achieved an A*-C grade in

Core, Additional or the triple science shows modest increases and decreases of less

than + or – 2% each year.

Table 5: Percentage of students achieving A*-C in selected science subjects in all schools in England 2008 to 2018

Year	Core Science	Additional Science	Biological Sciences	Chemistry	Physics	Other Sciences
2008	59.3	67.8	91.4	94.8	94.5	48.6
2009	60.8	66.8	93.0	94.7	93.9	57.5
2010	60.9	68.4	94.1	93.9	94.0	59.8
2011	63.0	69.6	94.3	93.4	94.1	60.6
2012	64.5	68.8	93.9	93.4	93.7	62.3
2013	60.1	65.2	91.6	90.9	91.4	62.1
2014	62.0	65.8	91.0	91.2	91.8	55.9
2015	59.2	64.6	91.6	91.5	92.3	58.9
2016	55.0	59.8	91.0	90.6	91.0	64.2
2017	53.1	58.2	90.9	90.1	90.9	67.4
2018	54.7		90.3	90.0	90.9	70.4

Sources: (DfE, 2014d, 2017, 2019c)
    2007/08 to 2012/2013: Subject time series tables SFR01_2014
    2013/14 to 2017: Subject times series tables SFR57/2017
    2018: 2018 Subject timeseries table

In 2013 all science subjects at GCSE saw a decline in results at A*-C of over 2%, when compared to the previous year. The results for Additional Science fell by more than 5% in 2013 from 2012, and the percentage of students who achieved A*-C in Core Science also fell by over 4%.   Although the attainment at A*-C remained relatively stable for the separate sciences since 2014, the Core and Additional Science percentage A*-C show year on year declines between 2014 to 2017. The 2018 grading for GCSEs has changed to number grades 9 to 1, meaning that the attainment data for 2018 is not directly comparable but it has been included here for completeness. The government published attainment data (DfE, 2019b) shows that in 2018 almost 55% of students achieved an equivalent of a C grade in two sciences (new grades 4,4).

3.4 Changes at Key Stage 2 from 2008

3.4.1 Accountability Measures

The inclusion of science as a core subject, consequently led to science national curriculum tests being administered to the majority of students at the end of KS2 (Bew, 2011). The Bew Review (2011) focussed on testing at KS2 and recognised the need for a change to the accountability system to make it more transparent, fairer and with an equal emphasis on progress and attainment. The findings from the report resulted in the changes to the assessment of the primary curriculum (DfE, 2013d).

Before to 2009, the results of the science tests at KS2 were published to parents but not used as accountability measures (Bevan *et al.,* 2009). After 2009, the national science tests at the end of KS2 were removed and replaced by a small but representative national test sample of 11-year-olds (Wellcome Trust, 2011). Identified as a more constructive way of monitoring national standards, the sampling methodology allowed year on year comparisons to be made (Stobart, 2008). A 5% sample of the student population in England is tested in science biennially. During 2010 to 2012 interim policies were put into place to deliver an annual measure of student performance in science using 750 schools as a sample (Standards and Testing Agency, 2016). Though there were no tests in 2013, the biennial KS2 science sampling assessment took place in 2014 (Standards and Testing Agency, 2017a). For 2016, approximately 9,500 students were randomly selected, based on five pupils from 1,900 schools (DfE, 2016b). The different sampling methodologies used in 2010, 2012, and 2014 and changes to national curriculum frameworks make year on year comparison of student outcomes from the external examination samples far more problematic. The science sampling tests are not published to parents or used to hold schools accountable but designed solely to give an estimate of attainment nationally (DfE, 2016b).

3.4.2 Grading

The outcomes from the science examinations and teacher assessments at KS2 were previously reported as levels, 1 to 6, with the key measure relating to the % of students achieving level 4 and above (Ofsted, 2013). From 2016, levels were no longer used, science in primary school is assessed using the Teacher Assessment Framework for Science (Standards and Testing Agency, 2017b). This stipulates a range of "pupil can" statements through which teachers make judgements on their pupil's progress against the set criteria. For science teacher assessment, valid results are:

- has not met the expected standard (HNM)
- working at the expected standard (EXS)
- absent (A)
- disapplied (D)
- maladministration (Q). (DfE, 2018d, p. 9)

Evidence that students have made expected progress in each of the scientific statements can be drawn from across the entire length of the key stage. Teachers are encouraged to undertake internal and external moderation of students work to ensure consistency and reliability of the assessments (Standards and Testing Agency, 2017b). As a result, there are a large number of learned societies such as the Royal Society of Chemistry (2014), government STEM initiatives (STEM Learning, 2018) and a wide range of commercial organisations (e.g. Empiribox, 2018; Siemens, 2018; TES, 2018) who provide resources and Continued Professional Development (CPD) to support primary science teaching and assessment.

3.4.3 Trends in Attainment

The attainment of students in end of KS2 from 2008 to 2015 is shown in Table 6 as the percentage of students who achieved level 4 and above (4+) and level 5 and above (5+) in the end of KS2 teacher assessment levels for science. Attainment at level 4 represents the expected standard set by the government and is associated

with progress targets and whole school performance measures (DfE and National

Statistics, 2010). The mean attainment outcomes across the 8-year period, shows

that 87% of students achieved the expected standard or above (level 4+). The mean

number of students who achieved at the higher ability, level 5+, across the same

period was 38% (see Table 6).

Table 6: Percentage of students achieving teacher assessment level 4 and above and
level 5 and above in science by year

	All	
	Level 4+	Level 5+
2008	85	38
2009	86	38
2010	85	37
2011	85	35
2012	87	36
2013	88	38
2014	88	39
2015	89	40
Mean	87	38

Source: (DfE, 2015b) See Appendix Table L, p363

The percentage of students achieving level 4+ and level 5+ at the end of KS2 has

risen steadily since 2011. Although not strictly comparable, it appears that overall

fewer students have met the expected standard following the changes to the primary

national curriculum introduced in 2014, which incorporated the increased scientific

demand. Table 7 gives the percentage of students who have met the expected

standard versus those who did not.

Table 7: Percentage of Students meeting expected standard in science by gender and year

	All		Males		Females	
	Met	Not	Met	Not	Met	Not
2016	81	19	79	21	83	17
2017	82	18	79	20	84	16
2018	82	17	80	20	85	15
Mean	82	18	79	20	84	16

Source: (DfE, 2018e, p. 7)

Where the mean for 2008-2015 was 87% of pupils reaching level 4+, for the period 2016-2018 the mean for students reaching expected standard fell to 82%, well below 2015 figures.

3.5 The Characteristics Student Population taught the science curriculum

Today the UK state education system is complex and diverse, reflecting the open market economics strived for through the ERA. As of January 2019, there were about 2,300 independent, fee-paying private schools who receive no government funding but often take on charitable status; 16,770 state-funded primary schools and 3,450 state-funded secondary schools (DfE, 2019d). Within the state sector, which also includes over 1,000 special schools, there is further variation with a range of community schools, foundation schools, voluntary aided, voluntary controlled, academies and free schools (Miller, 2011). The total student population in England has increased by 7.5% since 2008 (Appendix Table K, p362) with just under 4.73 million primary and 3.33 million secondary pupils attending stated funded schools in 2019. A description of the student population by school type and phase is not as straightforward as would be expected. There is an overlap between the ages of students attending middle schools (aged between 8 and 13 years old) and their designation as either a primary or a secondary institution. Whereas, some secondary

schools are designated as all through schools with cohorts of children from nursery through to post-16 (GOV.uk, n.d.).

It is useful to summarise the demographic profile of the student population as this study explored the trends in attainment over time and the impact of education policy reform on *all* children. Reporting the composition of the student population by their characteristics illustrates the increasingly diverse nature of primary and secondary classrooms to which education reform applies and sets the backdrop for chapter 7. Tables 8 and 9 below, detail the changes in the student population in primary and secondary schools and was taken directly from the DfE data collected via the school census, 2008 to 2018. The data shows the percentage of students by key student groups; gender, in receipt of Free School Meals (FSM), with English as an Additional Language (EAL), a statement of Special Educational Need (SEN) and by major ethnic group. The characteristics of the school population in terms of gender have remained relatively static between 2008 and 2018, with 51.0% males in primary and 50.3% in secondary school (Tables 8 and 9). The percentage of students with statements (SEN) has changed very little over the period in primary schools but show a small decline at secondary. The percentage of students who are known to be eligible for Free School Meals (FSM) has steadily decreased since 2013/14. Students with EAL have made up less than 20% of the student population in primary settings, with an even smaller proportion in secondary schools. Finally, the ethnic mix of primary and secondary schools has changed with an overall increase in the percentage of students from non-white backgrounds.

Table 8: State Primary Schools Student Characteristics by Year 2008 – 2018

Year	% Males	% Females	% FSM	% White	% Mixed	% Asian	% Black	% Chinese	% Other	% EAL	% SEN with statement
2008	51.1	48.9	16.6	80.0	3.9	8.9	4.8	0.3	1.3	14.4	1.5
2009	51.1	48.9	17.1	79.2	4.1	9.3	4.9	0.3	1.4	15.2	1.4
2010	51.0	49.0	18.5	78.5	4.3	9.6	5.1	0.3	1.5	16.0	1.4
2011	51.0	49.0	19.2	77.7	4.6	9.9	5.3	0.4	1.5	16.8	1.4
2012	51.0	49.0	19.3	76.9	4.8	10.3	5.4	0.4	1.6	17.5	1.4
2013	51.0	49.0	19.2	76.3	5.1	10.4	5.6	0.4	1.6	18.1	1.4
2014	51.0	49.0	18.0	75.8	5.3	10.5	5.6	0.4	1.7	18.7	1.4
2015	51.0	49.0	16.5	75.4	5.5	10.6	5.7	0.4	1.8	19.4	1.4
2016	51.0	49.0	15.2	75.0	5.7	10.6	5.7	0.4	1.8	20.1	1.3
2017	51.0	49.0	14.7	74.7	5.9	10.7	5.6	0.4	1.9	20.6	1.3
2018	51.0	49.0	14.2	73.9	6.2	11.1	5.5	0.5	2.0	21.2	1.4
Mean	51.0	49.0	17.1	76.7	5.0	10.2	5.4	0.4	1.6	18.0	1.4

Source: (DfE, 2019f, 2019e; DCSF, 2008) See Appendix A4, p362.

Table 9: State secondary schools student characteristics by Year 2008 - 2018

Year	% Males	% Females	% FSM	% White	% Mixed	% Asian	% Black	% Chinese	% Other	% EAL	% SEN with statement
2008	50.4	49.6	14.2	82.5	3.0	7.4	4.1	0.4	1.1	10.8	2.1
2009	50.4	49.6	14.5	81.9	3.3	7.7	4.3	0.4	1.2	11.1	2.0
2010	50.4	49.6	15.4	81.2	3.5	7.9	4.4	0.4	1.2	11.6	2.0
2011	50.4	49.6	15.9	80.6	3.7	8.3	4.6	0.4	1.3	12.3	2.0
2012	50.4	49.6	16.0	79.8	3.9	8.7	4.8	0.4	1.3	12.9	1.9
2013	50.4	49.6	16.3	79.0	4.0	9.1	4.9	0.4	1.4	13.6	1.9
2014	50.3	49.7	15.7	78.1	4.2	9.5	5.1	0.4	1.4	14.3	1.9
2015	50.3	49.7	14.9	77.1	4.4	9.9	5.3	0.4	1.5	15.0	1.8
2016	50.3	49.7	14.1	76.3	4.7	10.3	5.5	0.4	1.6	15.7	1.7
2017	50.2	49.8	13.8	75.1	5.0	10.7	5.6	0.4	1.7	16.2	1.7
2018	50.2	49.8	13.3	74.2	5.2	11.0	5.8	0.4	1.8	16.6	1.6
Mean	50.3	49.7	14.9	78.7	4.1	9.1	5.0	0.4	1.4	13.6	1.9

Source: (DfE, 2019f, 2019e; DCSF, 2008) See Appendix A4, p362

3.6 Summary

The combined effect of the above-mentioned developments will prove too complex to be fully evaluated within the scope of this research. The nature of education reform, and its universal application to maintained schools covering the vast majority of school-aged children, means that pilot schemes are rarely carried out (Earl, Watson and Katz, 2003). Instead, government-sponsored reviews are undertaken which are normally chaired by eminent academics accompanied by a panel of experts representing universities and learned societies (for instance: Bew, 2011; Rose, 2009; Noden *et al.,* 2007). So, a change to education policy is difficult to evaluate and judge in the same manner as a smaller scale intervention programme, as it relies on large scale, homogenised enactments.

The analysis of student attainment from the end of KS2 assessment and GCSE examinations in this research will be steered through a lens which explores changes over time and attempts to pinpoint how key aspects of policy change impact outcomes and teacher practice. In this research, particular attention is paid to how the changes manifest themselves within the secondary and primary national curriculum, testing and assessment regime. This leads towards an understanding of the policy environment under which schools and teachers find themselves working; and places student attainment in science within the context of the wider agenda of raising standards to meet national and international measures.

Chapter 4: Literature Review

This research involved conducting semi-structured interviews with teachers of science in secondary and primary school settings to determine whether their classroom practice had altered in response to changes in education policy. The literature review examines, what research suggests is, effective classroom practice and explores the factors that shape it. To achieve this, I briefly explain the theoretical foundations, originating from psychology, that inform an understanding of the process of learning and which are used to develop common pedagogical practices across a range of contexts. The literature review goes on to explore how the theory translates into classroom practice with the potential to promote student learning and considers some of the factors that may shape or change teacher practice, whether internally derived or externally imposed. The review also informed the development of suitable questionnaire items on science teaching practices used for data collection.

4.1 What is Teacher Practice? From Theory to the Classroom

4.1.1 Theories of Learning

This section briefly introduces, what are argued to be, the most prominent theories of learning based on psycho-dynamic research (Wang, 2012; Vosniadou, 2001). An explanation of how theory is translated into pedagogical practice then follows.

Learning is something that can happen naturally, planned or unplanned, as an individual interacts with their environment (Pritchard, 2017). However, theorists, researchers and practitioners have been unable to agree on a precise definition of what learning is (De Houwer *et al.*, 2013). Although there are different beliefs about how humans learn (Wang, 2012), educators generally accept that learning involves change, endures over time and occurs through experience (Schunk, 2012). It

is not possible to consider all theories of learning within the context of this study therefore, it focuses on the three key theories that connect learning and practice, behaviourism, constructivist and social constructivist (see brief notes in Appendix A2, p326).

Behaviourism emphasises the role of environmental factors in influencing behaviour (McLeod, 2017). Generally, behaviourism supports teacher-centred approaches where the teacher is the sole authority figure, and knowledge is parcelled out (Westbrook, 2013). The learner is largely passive and dependent on the teacher, reacting to conditions in the environment rather than taking an active role in discovering the environment. Classroom activities that require students to recall facts, define and illustrate concepts, automatically perform specified procedures with little choice or interaction fall within behaviourist models of teaching. Assessment in behavioural teaching and learning is often exam-oriented and high-stakes, without teachers' direct involvement (Muijs and Reynolds, 2017; Westbrook, 2013; Stewart, 2012). Furthermore, the use of clearly defined learning outcomes with specified observable objectives are a key part of behavioural teaching within a stimulus and reward environment.

With constructivism, unlike behaviourism, children create their own meaning, actively making sense of the world through their conception of reality (Vosniadou, 2001). The teacher's role is to create activities which require students to rethink their ideas. The encounters should be challenging enough so that students go through the process of accommodation – adding new information to change existing knowledge - and make progress. Social constructivist theories lead to pedagogic practices which prioritise student-teacher or student-student interactions and adaptable tasks for students across any age range (Muijs and Reynolds, 2017; Stewart, 2012). Activities include working in small-groups, pairs and whole-class

interactive work, extended dialogue with individuals, higher-order questioning, problem-solving, teacher modelling, and co-operative learning (Westbrook, 2013).

Research into learning and teacher effectiveness has been carried out widely (Ko *et al.,* 2014; Hanna *et al.,* 2010) despite this, there are no consistent, all-pervading answers, which when applied by teachers will enable all students to develop the high-level creative metacognitive skills described by Wilson (2016) and Krathwohl (2002). In fact, research and classroom practice often exist in an awkward relationship (Boxer and Bennett, 2019; de Corte, 2010) and understanding how learning happens is a complex and ever-evolving process (Pritchard, 2017). Having said this, the abundance of research and empirical studies do describe key aspects visible in the classrooms of effective teachers (Muijs *et al.,* 2014; Dunlosky *et al.,* 2013). The next section describes the translation of the theories of learning into classroom practice with a particular focus on science teaching. There are six key elements thought to form the basis of instructional practice in science namely; demonstrations; explanations; questioning; forms of representation; group and collaborative work and inductive-deductive learning cycle (Treagust, 2013, p. 373), these practices are explored, loosely grouped into two categories, teacher-centred and student-led.

4.1.2 Teacher-Centred

Whole class teaching, also known as direct instruction, is founded on behavioural theory and has been a significant pedagogical practice in schools for some time (Magliaro *et al.,* 2005). Not to be confused with didactic, "chalk and talk" teaching, the term refers to a style of teaching in which the teacher actively teaches the content of the lesson directly to the whole class (Muijs and Reynolds, 2017) placing the authoritative voice of the teacher at the centre of the social interaction in the classroom. Through lectures, explanations, dialogue and questioning, teachers

scaffold student's understanding; presenting new material in small steps that facilitate processing in the working memory (Fisher and Frey, 2014; Rosenshine, 2012). It has been found that students learn more through direct interactive teaching than through inquiry learning or working on their own (Muijs and Reynolds, 2017; Hattie, 2009). Through direct instruction, teachers can monitor the class and quickly respond to indications of pupil's (lack of) engagement. But there are limitations, its effectiveness is dependent on students' characteristics particularly when teaching more complex ideas (Muijs and Reynolds, 2000). Furthermore, critics of Hattie's Visible Learning meta-analysis argue that the evaluation approach which drives the conclusions are flawed both conceptually and methodologically (Rømer, 2019). McKnight and Whitburn (2020) add that whilst the teaching strategies put forward in Hattie's work (2009) are useful, its neoliberal style supports a surveillance culture in schools. A particular subset of direct instruction is Direct Instruction (capitalised DI), an approach developed by Engelmann and Becker (1980). This structured, scripted approach to teaching has proven effective but meets with resistance from those advocating constructivist teaching pedagogies (Boxer and Bennett, 2019).

4.1.2.1 Explanations and Lectures

Teacher explanations play a major part in developing student understanding by making use of language, modelling, demonstrations and imagery to create meaning and moving student thinking from the concrete to the abstract (Treagust, 2013). If given sufficient time to predict and record the outcomes, a teacher-led demonstration can engage students enabling them to find meaning, effectively support their understanding (Miller *et al.,* 2013; Crouch *et al.,* 2004), and is a cheaper and safer way to conduct laboratory experiments. Lecture-style explanations assume that all students learn at the same pace and offer little opportunity for feedback on learning; students took a passive role, in which their attention and engagement often waned (Schwerdt and Wuppermann, 2011).

However, using the 2003 Teacher Questionnaire from TIMSS, Schwerdt and Wuppermann (2011) found no evidence of a detrimental effect of lecture-style teaching on overall student learning as measured by TIMSS scores. They concluded that lecture-style teaching was no less effective than teaching based on in-class problem solving.

4.1.2.2 Sharing Learning Goals

Student's ideas are said to develop through the intervention and guidance of the teacher (the More Knowledgeable Other – see Appendix A2, p327) (Vygotsky, 1978) and the explicit sharing of teaching and learning goals (Schwerdt and Wuppermann, 2011). As a result, a key component of whole-class teaching and a long-established tenet of AfL (Black and Wiliam, 1998) includes clearly stating the objectives for learning in terms of student outcomes (Magliaro *et al.*, 2005). Confusingly, there are a variety of terms that describe what will be taught in any given lesson, terms like, *Teaching Goals*; *Learning Goals*; *Learning Intentions*; *Learning Outcomes*; *Success Criteria* and *Lesson Objectives* are often used interchangeably but have subtly different meanings (Marzano, 2010). Therefore, unless stated explicitly within school policy, different teachers, in different curriculum subjects may have different interpretations and expectations of these terms. When allied to a behaviourist model, learning objectives can be articulated to demonstrate observable outcomes; from a cognitive perspective, learning objectives can be framed to engage student thinking (Chizhik and Chizhik, 2016), while focused participatory objectives can emphasise socio-cultural aspects of learning (Havnes and Prøitz, 2016).
Through the use of active verbs (Havnes and Prøitz, 2016; Hattie, 2009), teachers state what students will do in a lesson but find it more challenging to be specific about what they are going to learn (Wiliam, 2011). Moreover, the objectives, often written as "Students will be able to…" (Chizhik and Chizhik, 2016, p. 211) when focussed on what students will do, rather than what they should be thinking,

overlook the development of student's cognitive skills. Furthermore, statements are often tokenistic and poorly defined, too narrowly focussed on learning intentions limiting student's ability to transfer the learning to a different context (Wiliam, 2017).

4.1.2.3 Teacher Questioning

Questioning forms a significant part of teacher-led direct instruction (Ko *et al.,* 2014), is a cornerstone of assessment for learning (Wiliam, 2017) and "is an integral part of good teaching" (Chin, 2007, p. 839). Although the effects of questioning vary due to the type and depth of the questions asked (Hattie, 2009), as a classroom strategy, seminal research found that teachers spend between thirty-five to fifty per cent of their classroom dialogue conducting questioning (Cotton, 2001). More recently, it has been suggested that this has changed very little, as teacher questioning dominated classroom social interactions, with 2-3 questions asked every minute (Dong et al., 2019; Albergaria-Almeida, 2010a; Almeida and Neri de Souza, 2010).

Teacher questioning can be diagnostic, to check student understanding and ascertain how well the material has been learnt (Fisher and Frey, 2014) or can support students to construct meaning and practice new information through teacher-student interaction (Pearsall, 2018; Gillies, 2013). Questioning can engage students in reflective thinking to promote higher-level scientific understanding (Smart and Marshall, 2013; Brookhart, 2010; Oliveira, 2010) particularly if aligned and targeted through, for instance, Bloom's Taxonomy (Krathwohl, 2002). Despite agreement that questioning promotes student learning, many teachers used a limited range of low-level questions that did not develop the students' cognitive skills (Smart and Marshall, 2013; Erdogan and Campbell, 2008; Chin, 2006) and as the focus of much classroom dialogue, questioning was not always planned in a way to support student

learning (Black and Wiliam, 2010). Neither did teachers provide sufficient wait-time after a question was asked and before a response was given (Albergaria-Almeida, 2010b). It has long been argued that a longer wait time allowed students to formulate their ideas and reduced the number of "don't know" responses (Rowe, 1986). Shorter wait times were often associated with low-level cognitive questions with teachers waiting less than 3 seconds before continuing (Iksan and Daniel, 2016). For questioning to draw out understanding effectively, teachers must avoid the temptation to close down student's conceptual thinking work, resist using questions where the answer is already known (Iksan and Daniel, 2016; Heritage and Heritage, 2013) or asking questions where the students are required to "guess" the correct answer (Oliveira, 2010).

4.1.2.4 Feedback

Coupled with questioning, feedback is considered to be an effective tool in developing student engagement and understanding (Shute, 2008). However, due to the wide range of different types of feedback interaction, the impact on student learning was variable (Voerman *et al.*, 2012). Formative feedback has been defined as "information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving learning" (Shute, 2008, p. 154) and integral to assessment *for* learning (Wiliam, 2017). Whereas, summative feedback can be conceived of as assessment *of* learning often occurring at the end of significant episodes of instruction. Generally associated with high-stakes and used to hold teachers, pupils, schools and local authorities to account (Roberts, 2016b) summative feedback can be used formatively, to identify topics that require revisiting (Dixson and Worrell, 2016).

Effective feedback occurs when connected to showing a student where they are now; what the goal is and specifying the steps needed to get there. Learning is enhanced

as the gap between what is understood and what is intended is reduced (Hattie and Timperley, 2007). However, feedback, oral or written, positive or negative, specific or non-specific, grade or process focussed, varied in its usefulness as teachers sometimes mixed several different feedback types in one message (Gamlem and Smith, 2013). Voerman (2012) suggested that much of a teacher's feedback to students tended to fall into two categories. *Discrepancy feedback,* detailed the difference between the current level and the next steps in student understanding. Whereas *progress feedback* provided students with information on how far they have come from their initial starting point. The study, with 78 Dutch secondary teachers, found that feedback was mostly non-specific and comprised less than 20% of all observed interventions (Voerman *et al.*, 2012). Discrepancy feedback was provided by a higher number of teachers than progress feedback, thereby emphasising what was not learned rather than what was achieved.

Research studies tended to agree that in order to be effective, feedback should be specific, timely, guide improvement but not overly elaborate (Gamlem and Smith, 2013; Shute, 2008). Moreover, emotions, particularly those associated with positive responses, also played a part in how students related to feedback. Giving praise, the most frequent form of feedback intervention, had an impact on learning but only when specific and goal-related (Voerman *et al.*, 2014). Ultimately, comment-only written feedback without grades, combined with correctional review that is acted upon, was considered more "powerful" than feedback that is personal (Black *et al.*, 2004) or effort based such as "well done, you've worked hard" (Hattie, 2009, p. 177).

4.1.2.5 Summary

In summary, when whole-class teaching is used interactively, with a varied selection of activities to engage and challenge student thinking, cognitive gains are said to take place. Teacher-led classroom practices involve structured lessons with teacher direct instruction, clear objectives, modelling of the required outcomes, questioning for understanding and consolidation, on-going formative assessment and feedback (Muijs and Reynolds, 2017; Rosenshine, 2012; Hattie, 2009). However, other modes of teaching, involving students working with each other and teachers as facilitators, have been found increasingly effective in supporting student learning.

4.1.3 Student-Centred

Student-centred learning encompasses a range of pedagogical practices sometimes known as "flexible learning", "self-directed learning" and "personalised learning" (Wolfe *et al.,* 2013; Robinson and Sebba, 2010). All of which reflect the ways in which teachers structure lessons, enabling students to take an active role in learning and co-construct understanding. Students' active involvement in their learning supports the likelihood that any new knowledge can be applied to other contexts (Muijs and Reynolds, 2017). In science, student-centred learning might include holding debates on issues such as genetics, carrying out practical work or investigations in pairs or presenting posters from their research (Rosenshine, 2012).

4.1.3.1 Working in Groups

Small group work is an example of how teachers structure lessons to support student-centred learning. While there are similarities and differences, overlap and divergence between the practices, research shows that group work, collaborative and cooperative learning are linked to improved student learning outcomes and socialisation (Mercer and Howe, 2012; Galton *et al.,* 2009; Hattie, 2009). However,

clear distinctions between collaborative and cooperative learning are difficult to make and have long been sought (Panitz, 1999). Collaborative learning features activities intentionally structured to ensure that all participants actively engage (Barkley *et al.,* 2014). Whereas, cooperative learning is described as "the instructional use of small groups so that students work together to maximise their own and each other's learning" (Johnson and Johnson, 2014, p. 841). Not all group work is cooperative; to be so, students must interactively work toward a common goal. Under the banner of "group work", these pedagogical practices reflect social constructivist underpinnings and can be used to structure learning across different activities (Slavin, 2015).

How students work together, *as* a group or *in* the group, underpins the differences in practice and reflects how accountability for learning is divided, either across the group or individually (Scheuermann, 2018; Wiliam, 2017).   Although it may take longer to cover the material (Muijs and Reynolds, 2017), for effective group work, teachers need to adequately prepare their students. Setting out the conditions and expectations for group interactions and outcomes helps to avoid the students focussing only on the end product. Careful preparation also helps deter students from shirking responsibilities for the group learning or engaging in off-task talk yet being awarded the final group grade (Slavin, 2015; Barkley *et al.,* 2014).

Mixed ability grouping places students with differing expected attainment levels together and can be effective in raising attainment (Boaler, 2013). Albeit requiring a time commitment, teachers use experience and knowledge in order to allocate groups and structure lessons (Blatchford and Russell, 2018; Brown, 2017) with students aware of their assigned groups, even from a young age (Boaler, 2013). Nevertheless, this grouping may also lead to adverse reactions from academically

able students who feel that they are being held back or from less assertive students who feel ignored by others in the group (Tsay and Brady, 2010). Homogeneous groups, consisting of the more academically able students may promote a supportive yet challenging learning environment. Yet, this may also induce anxiety in students that find the pressure to perform at the highest level on every task, too stressful (Dweck, 2007). Equally, grouping all students with lower target grades together may result in the negative effects of lowering the expectations of the group or fuelling feelings of inadequacy (Dweck, 2014). Whilst collaborative and cooperative learning are seen to have a positive impact on student understanding, equally, small group work may limit independent learning and reinforce misconceptions (Muijs and Reynolds, 2017).

Gillies and Boyle (2010) concluded that effective implementation of cooperative learning pedagogies is complex requiring teacher decision making across many dimensions including class organisation, task type, mode of instruction and expectations. Recent studies (Raviv *et al.,* 2019; Buchs *et al.,* 2017) further highlight the barriers for teachers in introducing this form of learning, specifically group dynamics, motivation, managing group talk and assessing student contribution. Their conclusions suggested that teachers found cooperative learning challenging to implement and limited by curriculum time constraints.

4.1.3.2 Self- and Peer- assessment

Pair and small group work might also involve students in peer- and self-assessment. As a component of assessment for learning (Black and Wiliam, 2010), these classroom pedagogies can be utilised by teachers in the feedback process and are said to enhance collaborative learning (Topping, 2013). Peer assessment can be defined as "a set of activities through which individuals make judgements about the work of

others" (Reinholz, 2016, p. 301), as opposed to self-assessment which uses the act of questioning oneself to make judgements about progress and next steps (Boud, 2013). Although not well understood, peer-assessment is thought to support self-assessment through the involvement of students in an assessment cycle and learning dialogue focused on closing the gap between goal and current performance (Reinholz, 2016). Reliant on teachers creating a supportive learning environment, the ability of students to make accurate judgements based on their understanding was one of the potential gains of peer- and self-assessment (Sadler and Good, 2006). Students became aware of strengths and progress and, improved the productivity and cooperative nature of the learning by generating a positive attitude towards the activity. Finally, students peer- and self- marking saved teacher time and ensured the swift return of feedback, although Sadler and Good (2006) found that students gained greater benefits from self-marking compared to peer-marking.

There are numerous ways in which peer- and self-assessment can be structured, concerns around accuracy, a reluctance to criticise friend's work and the possible detrimental psychological effects of students publicly sharing outcomes were possible drawbacks (Panadero and Brown, 2017; Harris and Brown, 2013; Topping, 2013). Sharing success criteria, examination mark schemes or grading rubric were ways that teachers explicitly specified the learning expectations and therefore ensured students did not, under or over, inflate the level of achievement (Reinholz, 2016). Allotting time for reflection and response in the form of Dedicated Improvement and Reflection Time (DIRT) was considered good teacher practice (Hill, 2014; Beere, 2012) ensuring that feedback was not only valued but valuable; acted on and learned from (Muijs and Reynolds, 2017; Quinton and Smallbone, 2010).

4.1.3.3 Practical work

Practical work in school science, although commanding a large share of the science department budget (SCORE, 2013a), was considered to motivate students, enhance their learning of concepts, teach manipulation and processing skills and was integral to the overall experience of the science curriculum (SCORE, 2013a; Toplis and Allen, 2012; Woodley, 2009). Appropriately planned practical activities should scaffold student learning from observation and evidence gathering toward explanation and understanding. "Hands-on" practical activities were effective at teaching laboratory skills and techniques, whilst open-ended investigations were better used as tools in the development of student's understanding of the scientific methods of inquiry (Osborne and Dillon, 2010). Although the evidence for the effectiveness of practical work in developing student conceptual understanding has been previously questioned (Abrahams and Millar, 2008), teachers and students continued to view practical work as effective and affective in the learning and enjoyment of science (Abrahams and Reiss, 2012).

Practical work was usually carried out by students in small groups or pairs (SCORE, 2013a), with teachers following different approaches. Either, an inductive approach which involved collecting data then formulating a rule-based theory to govern and explain the results. e.g. on the angle of incidence and angle of reflection. Or, a deductive approach, starting from the theory which is confirmed by collecting data, e.g. increasing the temperature by 10°Kelvin doubles the rate of reaction. Students often recalled the "whizz-bang" element of a practical lesson but had little understanding of why they had observed the task or what they had learnt from the scientific ideas. Thus, practical work anchored the descriptive nature of the science and though there was evidence that teachers planned practical activities in order to produce the intended observable learning, in the manipulation of equipment and

conduct of procedures, it appeared that the overt planning for the teaching of ideas was not always apparent (Abrahams and Millar, 2008).

Practical work carried out in school laboratories often involved far more recipe style tasks, teacher-led instruction, rather than student-centred open-ended investigations (Abrahams and Reiss, 2012). This research suggested that open-ended tasks presented increased pedagogical challenges to teachers, the bias towards recipe style tasks was, to some extent, due to the relatively short length of lessons and availability of limited resources such as equipment and consumables. At key stage 4 in particular, a student's real experience of open investigations – undertaking a task for which they do not immediately see the answer- is limited and may be overly structured to meet the needs of the assessment regime (Ofsted, 2013; Toplis and Allen, 2012). More recently, the 2016 reforms to the assessment of scientific inquiry at GCSE required new styles of pedagogy to increase the knowledge retention of the assessed practical procedures. Where pre-laboratory preparation outside of lesson time was used, the learning opportunities and familiarity with the practical procedures was extended thereby supporting student's retention of new knowledge (Hennah, 2018).

4.1.3.4 Summary

Surveys such as those conducted by TIMSS (Trends in International Mathematics and Science Study) (IEA, 2014) aim to capture the essence of what makes a good teacher; they acknowledge that understanding the effectiveness of instructional quality is complex. High-quality teaching depends on establishing a positive, supportive classroom climate, an element of direct instruction through well-paced, well-structured lessons; support for student autonomy and self-determination and opportunities for students to face cognitive challenges and undertake problem-

solving (Vieluf *et al.,* 2012). International organisations researching the global teaching landscape look for evidence when conducting teacher questionnaires in searching for "what works" (Mullis and Martin, 2013; Vieluf *et al.,* 2012). Despite the large-scale assessments such as PISA identifying successful aspects of classroom climate, they are less clear on identifying particular types of teaching practices which work for all or acknowledging that teachers must adapt and differentiate their teaching to meet the needs of specific classes and pupils. The next step in this review examines what the literature says about factors that contribute to changing teacher practice.

4.2 What Shapes and Changes Teacher Practice?

The previous section evaluated some of the teaching practices used in science classrooms to promote learning. Correspondingly, there are a multitude of different conceptions around what shapes teachers' practice. In theorising teachers as learners, Clark and Hollingsworth (2002) described six perspectives to explore the internal and external drivers in the change process :

- Change as training—change is something that is done to teachers; that is, teachers are ''changed''.

- Change as adaptation—teachers ''change'' in response to something; they adapt their practices to changed conditions.

- Change as personal development—teachers ''seek to change'' in an attempt to improve their performance or develop additional skills or strategies.

- Change as local reform—teachers ''change something'' for reasons of personal growth.

- Change as systemic restructuring—teachers enact the ''change policies'' of the system.

- Change as growth or learning—teachers ''change inevitably through professional activity''; teachers are themselves learners who work in a learning community.

(Clarke and Hollingsworth, 2002, p. 946)

Change as learning and growth, the continued development of one's practice is the actualised expectation of the Teacher Standards (DfE, 2013a) in the professional activity of a teacher. In exploring what factors shape teacher practice and influence what they do in the classroom, this section first describes the role played by the interconnected constructs of beliefs and self-efficacy. These constructs are then linked to the pedagogical content knowledge (PCK) inherent in the teacher through their subject knowledge and learning experiences. I have brought these three constructs together as the *Internal* factors that shape an individual's teaching practices. I go on to explore several *External* factors, which arguably, contribute to changing teachers' practice. These include elements derived from the stratified social layers found within and outside the school, namely professional development, school leadership, and policy reform.

4.2.1 Internal Factors

4.2.1.1 Beliefs

Beliefs are the best indicators of the decisions people make throughout their lives (Bandura, 1986). The global construct of "belief" has long been the subject of research and philosophy across diverse fields and has therefore acquired different definitions as discussed in Pajares' (1992) well-cited review. Often described as the personal convictions or the ideas one holds, clusters of beliefs are said to form attitudes or intentions to action Centrally held "core" beliefs are thought to be formed early in life and are more difficult to change than the "peripheral" beliefs accumulated through education (Glackin, 2016). Therefore, peripheral beliefs systems are said to be dynamic and permeable mental structures susceptible to change in the light of experience (Muijs and Reynolds, 2002). Glackin (2016) suggested that a teacher might have a core belief about how children learn but hold a peripheral belief on how children should be taught; meaning that the belief about

learning theory is more stable and influential than the belief about pedagogy (2016, p. 412). Arguably, there are other beliefs that are relatively static, less susceptible to change (Dweck, 2008; Nespor, 1987) and include those allied to one's emotions (Mercer, 2010) or religion (Graham and Haidt, 2010). The interest in teacher's beliefs has been developing over the past 60 years (Ashton, 2015) during which time authors, for example, Buehl and Fives (2011) have attempted to bring more clarity and a better understanding to the construct.

Research suggests that teacher's beliefs are multifaceted and draws upon a range of contexts (Fives and Buehl, 2016; Hutner and Markman, 2016; Skott, 2015). Among these are epistemological beliefs, defined as "the beliefs that teachers hold about the nature of scientific knowledge and knowing" (Chen *et al.,* 2014, p. 4); pedagogical beliefs about how students learn and beliefs about the way the curriculum is structured (Enderle *et al.,* 2014). Beliefs, based on a teacher's learning experience and personal life, are thought to filter how an individual views or incorporates new knowledge into their practice (Davis, 2003).   Accordingly, beliefs about the different instructional processes which impact on instructional practices can vary depending on an individual's experience, both as a learner and a teacher (Fives and Buehl, 2016). Furthermore, teachers hold a range of beliefs about their professionalism, the purposes of education and schooling, and the impact of context on outcomes for children (Vieluf *et al.,* 2012; Waters-Adams, 2006).   It is also suggested that whilst a teacher's beliefs shapes their practice, there is a reciprocal relationship in that via reflection, practice influences beliefs (Hutner and Markman, 2016; Enderle *et al.,* 2014). For some time it has been thought that a teacher's beliefs forms part of their embodiment as individuals exerting a more powerful influence over their teaching than their subject knowledge (Nespor, 1987).

Linked to their experiences, a teacher's beliefs are individual, subjective, value-laden mental constructs, forming the lens through which classroom practice is interpreted and conducted (Skott, 2015). For instance, the belief that males are better at physics and females are better at biology is subjective and value-laden but may affect classroom interactions and processes. Similarly, the culturally-based beliefs a teacher holds may impact their expectations of the academic achievement of certain students, particularly those from an ethnic minority, economically disadvantaged backgrounds or those with special needs (Strand, 2010; Brady and Woolfson, 2008; Rubie-Davies *et al.,* 2006). It has also been found that with more experience, teachers beliefs can become ingrained as part of what they do (Levin *et al.,* 2013).

The association between a teacher's epistemic beliefs about science and their teaching practices is not straightforward (Apostolou and Koulaidis, 2010). If the teacher's epistemic beliefs treat knowledge as individual facts, learned and examined through tests, then students took a passive role in their learning. Although it was not always the case that holding more complex epistemic beliefs, led to the greater use of constructivist teaching strategies, in which students were co-constructors of their learning (Wallace, 2014; Gill *et al.,* 2004). However, school context, having to prepare for examinations and, a teacher's self-efficacy affected the use of constructivist practices even when a teacher held constructivist beliefs (Chen *et al.,* 2014). For the experienced and novice teacher alike, making strong links to personal learning and the classroom context were important factors in changing beliefs and practice (Fives and Buehl, 2016; Davis, 2003).

4.2.1.2 Self-Efficacy

A person's self-efficacy is related to beliefs about their ability to produce the expected levels of performance which influence and affect the events in their lives (Bandura, 1994). It was argued that the most effective means to develop a strong sense of efficacy is through mastering a task or skill, other factors include seeing the success of others similar oneself, being persuaded that one can succeed and adopting a positive emotional state (Bandura, 1994). Overall, self-efficacy influences the thoughts and emotions that enable individuals to undertake challenging actions that require effort and resilience (Tschannen-Moran and Hoy, 2007).

Self-efficacy is specific to a particular task (Tschannen-Moran and Hoy, 2001), therefore, where a teacher has made an effort to develop their science teaching and judges themselves capable of delivering engaging science lessons and improving the learning for students their sense of self-efficacy increases (Goddard et al., 2004). Furthermore, a teacher's beliefs are also related to their efficacy, where broadly, teacher efficacy refers to teachers' belief in their ability to influence valued student outcomes (Wheatley, 2005). Also understood to be a motivational construct, teacher efficacy is based on self-perception of competence rather than actual competence, meaning that a teacher enters the classroom with a certain level of expectation, a belief that they can teach science at the appropriate level (Tschannen-Moran and Hoy, 2007). However, inconsistency within the definitions of the constructs makes the evidence from research studies inconclusive (Vieluf *et al.*, 2012).

The suggestion of a causal relationship between teacher self-efficacy and student attainment is complex and difficult to establish due to the influence of other factors (Muijs and Reynolds, 2002). Teachers with high levels of self-efficacy operated well in any context and taught dynamically. They were more likely to do an effective job

preparing students for tests and providing them with opportunities to explore the multitude of "right" answers through scientific inquiry. Whereas, teachers with low levels of self-efficacy, took fewer risks in their teaching and viewed science as a static body of knowledge (Sandholtz and Ringstaff, 2014; Muijs and Reynolds, 2002). Furthermore, the link between teacher's epistemic beliefs and their self-efficacy did not always predict or reflect actual classroom practices (Chen *et al.*, 2014; Haney *et al.*, 2002). However, Miller *et al.*, (2017) concluded that individuals with a high level of self-efficacy were enthusiastic and expected success in their teaching, managed student behaviour and student learning effectively under challenging circumstances, and as a result, made a substantial contribution to student achievement.

4.2.1.3 Pedagogy and Pedagogy Content Knowledge

The seminal research literature of Bernstein (1975, 2000, 2004) theorised that educational knowledge, to a greater or lesser extent, defines our identities and this in itself is realised through three message systems: curriculum, pedagogy and evaluation (assessment). Behind these message systems lie two states, one an explicit transmission of knowledge and the other an implicit, hidden curriculum which reproduces structural power inequalities within pedagogic discourse (Hoadley, 2006). A narrower description sets pedagogy as "the activity of teaching or instructing and the methods used to instruct. It is the art or science of being a teacher" (Khader, 2012, p. 5). Simple statements like this mask the complexity of a term that is difficult to define (Zyngier, 2015). From the perspective of schooling, pedagogy is a combination of teachers' ideas, beliefs and attitudes, reflecting teacher's knowledge and understanding of their subject, the teaching and learning process and their students (Westbrook, 2013). This section relates pedagogy to teacher's science subject knowledge and classroom practices through developing an understanding of the role of Pedagogical Content Knowledge.

Alexander (2012) acknowledged that teacher beliefs are contextually based and that the definition of pedagogy should take into account social, cultural and political aspects. His definition stated that:

> Pedagogy is the observable act of teaching together with its attendant discourse of educational theories, values, evidence and justifications. It is what one needs to know, and the skills one needs to command, in order to make and justify the many different kinds of decisions of which teaching is constituted. (Alexander, 2012, p. 14)

Drawing on other constructions of pedagogy, Husbands and Pearce (2012, p. 3) synthesised a range of studies and presented nine key features of effective pedagogy. Nested among them is a statement outlining that effective pedagogies "… depend on behaviour (what teachers do), knowledge and understanding (what teachers know) and beliefs (why teachers act as they do)". The elements in this statement; what teachers do, understand and believe, are crucial and drive the other eight strategies, which include embedding assessment for learning, scaffolding and ensuring inclusivity. Pedagogy then displays multiple facets in that it enables teaching and transfer of knowledge through the educator's understanding of their subject, understanding of their students and personal beliefs. It is located and contextualised in the specific cultural framework (Jung and Pinar, 2015) which outlines what knowledge is of most worth via the visible and invisible curriculum.

Pedagogical Content Knowledge (PCK) can be defined as "the ways of representing and formulating a subject that make it comprehensible to others" (Shulman, 1986, p. 9). This conceptualisation of teacher knowledge includes an understanding of the teachability of the subject content; reflects on what makes particular subjects more challenging to teach than others and also recognises that students bring with them misconceptions and preconceptions which need to be addressed and overcome (Shulman, 1986). A clear operational definition for PCK has been a matter for debate

(Gess-Newsome *et al.,* 2017; Abell, 2008). However, I suggest that the manifestation of PCK is visible through the decisions made on the use of a particular mode of instruction to teach a particular topic, with a particular group of students. It depends on among other things, what the teacher believes are the goals of learning science. For example, a physics graduate teaching photosynthesis to a bottom set year 11 group on Friday period 5 may result in a very different lesson to that of teaching forces to an academically able year 9 group period 1 on a Monday morning. For science teachers, PCK represents the ability to employ a variety of representations to transform and adapt science content knowledge to meet the needs of students with different prior knowledge and cognitive abilities. Park *et al.* (2011) placed PCK at "the intersection of content and pedagogy" (p. 246). An understanding of this intersection is important because it adds to the suggestion that obtaining an undergraduate degree in a science discipline does not guarantee that an individual will become a "good" teacher (Kind, 2009a).

Over time, a teachers' orientations, knowledge and beliefs about the curriculum, assessment and instructional strategies changes and become threaded within the PCK of individual teachers (Brown *et al.,* 2013; Magnusson *et al.,* 2002). However, both novice and experienced teachers, with little PCK for specific content, are less able to structure and integrate learning particularly when they are teaching outside of their science degree specialism (Bartos and Lederman, 2014; Mulhall *et al.,* 2003). Additionally, teacher perspectives about particular teaching practices are often resistant to change even among trainee teachers (Brown *et al.,* 2013). So, while aspects of a teacher's PCK may be developed at any time in their career, for instance in response to changes in curriculum content, a teacher's familiarity with the subject matter may also blind them to an understanding of a student's difficulties (Chan and Yung, 2018).

A teacher's actions are the result of a complex thinking process reasoned through their PCK (Mulhall *et al.,* 2003), therefore, as a construct, it can be said that PCK emerges as tacit, "a hidden concept" (Kind, 2009a, p. 170). Furthermore, the range of models and elaborations create uncertainty around the claims of PCK which impact on its wider dissemination (Gess-Newsome *et al.,* 2017; Kind, 2009a; Abell, 2008). Hattie (2009) suggested that there was very little wide-ranging evidence to support the claim that improving teacher PCK improved student attainment as it was impossible to distinguish PCK from subject-matter knowledge and general pedagogical knowledge. Despite this, I have used PCK as a lens with which to draw together the literature around teacher's beliefs, orientations, subject knowledge, and pedagogical knowledge.

PCK represents "a unique knowledge domain within a teacher's mind" (van Dijk and Kattmann, 2007, p. 893). Consequently, it is subject to change, continually evolving as the teacher comes into contact with new contexts, ideas, students and reform. I argue that teachers' beliefs and practices cannot be taken out of the contextual and cultural framework in which they operate. Access to teacher development opportunities, curriculum change and school leadership are examples of the external, contextual factors which help to shape teacher practice. These factors are discussed in the next section of this chapter.

4.2.2 External Factors

Teachers bring their individual beliefs and experiences into the complex, socially, open education system in which they work. This section focuses on three external factors, which arguably create a structural boundary for teacher action. The first is engagement with professional development and the contribution made to changing teacher practice. Secondly, the impact that leadership and leadership style have on

teacher classroom actions within the bureaucratic context of the school is discussed. Finally, how reforms to the curriculum and its assessment impact teacher practice is examined. I contend that this third factor acts as a mechanism for change over which the individual teacher or school exercises little control.

4.2.2.1 Professional Development and Professional Learning

There are a number of different terms and conceptualisations associated with teachers' development and learning. While long-established terms such as In-Service Training (INSET) and Continued Professional Development (CPD) have been used to describe a range of professional learning activities, these have evolved as continued exploration has led to a more nuanced understanding of the constructs (Borko *et al.*, 2010; Desimone, 2009). Often used interchangeably (Jarvis and Doherty, 2016), there has been a shift towards using terms such as *teacher learning*, *teacher professional learning* and *teacher professional development* to describe teachers' on-going commitment to maintaining their professional expertise (Cumming, 2011; Borko *et al.*, 2010). However, a distinction can be made between professional development – as something that is done *to* teachers driven by external imperatives- in contrast to professional learning, as something that is done *by* teachers in response to their own issues or concerns (Nilsson, 2014).

Ideas about teacher learning reflect the shifts in ideas about student learning which lean towards more constructivist strategies, situated in classroom practice and often involving others in the formation of professional learning communities (Borko *et al.*, 2010). Avalos (2011) stated that "professional development is about teachers learning, learning how to learn, and transforming their knowledge into practice" (2011, p. 10). This complex process required both cognitive and emotional involvement of teachers as individuals and collectively. Broadening the concept

further, Korthagen (2016) put forward a framework representing teacher professional

learning, as an often unconscious, multi-dimensional and multi-level activity. The

model connects with the sources of teacher behaviour, which were cognitive,

affective, motivational, and embedded in the social context of the school. This

suggests that outcomes for teacher learning were unpredictable, therefore uniform

approaches to professional learning did not impact all teachers in the same way and

furthermore, professional learning occurred at any time through different

interactions and need not be experienced as a structured event.

Arguably, professional learning results from engagement with professional

development activities, whether labelled as a professional development programme,

CPD or INSET (Jarvis and Doherty, 2016). However, involvement may not always

develop a teacher's classroom practice (Korthagen, 2016; Borko *et al.*, 2010). Factors

that impacted the effectiveness of professional development programmes included

failure to take into account teachers' motivation for engaging in professional

development, misunderstanding the change process teachers underwent (Guskey,

2002), applying a deficit model to teachers or using programmes which consisted of

one-off interventions (Clarke and Hollingsworth, 2002). Pre-packaged, ready-made

CPD workshops presented to a passive audience of teachers failed to recognise that

teacher learning was an on-going process requiring continued support (Opfer and

Pedder, 2011). Furthermore, reasons for poor implementation of new ideas ranged

from the time and location of the training, the large teaching load of the participants

and the teacher's desire for the type of social constructivist delivery as described in

Murphy *et al.,* (2015) below (EL-Deghaidy *et al.,* 2015). The assumption that

teachers consciously translated theory into practice during classroom interactions

also recognised the limitations of applying professional learning to change teacher

actions (Korthagen, 2016). These conclusions reflect Shirley Simon's earlier remark that

> … unless teachers really want to change, or really value how a particular change can make their and their students' experience more worthwhile, they will not alter how they perceive themselves as science teachers or radically change their practice (Simon, 2012, p. 17)

Despite being time-consuming and challenging (Borko, 2004), the most effective development programmes are contextually situated, self-initiated and teacher-led; involve networking within supportive communities, incorporating peer observations and joint knowledge production (Vermunt *et al.,* 2019; Cheng and Ling, 2013). The types of professional activities that are thought to promote teacher learning and improve practice include experimenting, reflection and collecting new knowledge to keep up-to-date (Pyhältö *et al.,* 2015; Thoonen *et al.,* 2011). Furthermore, different types of collaborative work may instigate different types of teacher learning (Nilsson, 2014; Levine and Marcus, 2010).

Murphy et al. (2015) tracked 17 Irish primary school teachers in a 2-year professional development programme incorporating social constructivist approaches. An increase in the amount of science taught and more frequent use of student-led approaches in science teaching were evidence for the positive impact of the programme on teacher practice. The use of explicit modelling of the approaches expected in the classroom, hands-on reflective, inquiry-based activities; collaboration and the development of critical friendships made the programme successful. Still to be explored in a follow-up study, is the embeddedness of the new practices introduced to this small sample of teachers. A more extensive 2-year study, the *Getting Practical* programme*,* involved over 200 trainers, training over 2000 primary and secondary school teachers towards making improvements in the effectiveness of practical work in science (Abrahams *et al.,* 2014). A cascade model, combined with lesson observation, was

used to move teachers from a "hands-on" to a "minds-on" approach to practical

work to increase higher-order thinking. Although there was an increased awareness

among teachers about why practical work was used, the nature and extent of the

learning varied. When delivered appropriately, the cascade model was found to

change teacher practice in only one school but was not an effective transfer of

knowledge overall. This contrasts with work by Maass and Engeln (2018) whose

findings showed that the cascade model was effective in delivering a change to

teacher practice in the implementation of a large scale CPD program on modelling.

Change in teacher practice can be on-going in response to the day-to-day

interactions with staff, students and parents and it can be said that professional

development takes place every day through individual reflection, staffroom dialogue

(Akiba and Liang, 2016), and even lesson preparation (Weißenrieder *et al.,* 2015).

Improving teaching practices is a pivotal role of the headteacher, who strongly

influences the context in which teachers work (Sammons *et al.,* 2011; Leithwood *et

al.,* 2008). For this reason, it has been said that school leadership was second only to

classroom teaching as an influence on student learning (Leithwood *et al.,* 2008). The

next section looks at how the literature defines leadership in the school context and

examines the contribution that this makes to teacher practice.

4.2.2.2 School Leadership

Through a variety of leadership practices, a school leader, more usually the

headteacher, plays a pivotal role in shaping individual teacher practice and the

subsequent student learning outcomes. Practices which include setting the vision

and direction for the school; understanding and developing staff; creating productive

working conditions and managing the teaching and learning represent what successful

school leaders do (Pedder and Opfer, 2011; Leithwood *et al.,* 2008).

Theories that describe the different models of school leadership provide alternative interpretations of how vision, capacity building and goals are achieved (Bush and Glover, 2014). *Leadership for Learning* places the headteacher at the centre, driving the actions necessary to ensure teacher quality and support learning outcomes. This conceptualisation of leadership blends the instructional leadership, primarily concerned with teaching, with the features of transformational leadership to generate a school-wide focus on learning (Hallinger and Heck, 2010). Arguably, the transformational leader exists to change things for the better which is visible through their ability to increase the commitment, capacity and engagement of staff towards meeting agreed goals (Moolenaar *et al.,* 2010). This model stresses the importance of values but differs to that of moral leadership which places the focus on values, beliefs and ethics of the individual leader (Hoch *et al.,* 2018). While criticised as a means to control teachers, by requiring them to adopt centrally imposed policies and values, transformational leadership, when it works well is an effective means to achieve educational outcomes (Bush and Glover, 2014). Moreover, schools situated in an education context characterised by the marketisation and accountability structures of business organisations, require transformational leaders able to navigate this complexity (Anderson, 2017).

The headteacher also plays a part in managing and developing teacher practice through granting access to resources, building learning networks and fostering a collaborative culture of distributive leadership, research and innovation (Pedder and Opfer, 2011; Sammons *et al.,* 2011). Distributed leadership is claimed to have a significant influence on schools and students through its direct effect on staff working conditions and the indirect effect on student achievement. Although this leadership style might not fit comfortably within the bureaucratic organisational structure of a school (Hartley, 2010) and underplayed the hidden power dynamics

within the school (Lumby, 2013), distributed leadership recognises that leadership does not only reside in the role of the headteacher; any individual member of staff or group can influence others and what happens in the school (Bush and Glover, 2014).

Despite the headteacher retaining ultimate responsibility for the school, distributive leadership features particular behaviours. For example, acting as a role model and providing feedback to enhance a sense of self-efficacy by linking a teacher's needs to that of the school. This leadership style created a supportive environment for teaching and learning which allowed middle leaders and heads of department, for instance, to feel enthused, confident and able to communicate with their teams effectively (Diamond and Spillane, 2016; Leithwood, 2016). Teacher leaders in a formal leadership role, involving management and pedagogical responsibilities such as head of department or key stage coordinator, exhibited leadership practices that were undertaken collectively, becoming an effective driver of school improvement (Muijs and Harris, 2007). The role of the "good" subject leader was to share sound, up-to-date subject knowledge, display strong practitioner skills and maintain and develop resources. They developed the expertise and classroom practice in their team by introducing change in small incremental steps, building teacher confidence and providing the knowledge needed to implement changes (Leithwood, 2016). The teacher then becomes the leader in their own classroom, free to use their agency in decision making on what constitutes good practice (Boberg and Bourgeois, 2016).

Several studies have attempted to build an understanding of the relationship between school leadership, teaching and learning, and student outcomes. The findings from a 3-year mixed methods research study with 378 primary and 362 secondary schools indicated that the socio-economic context of the school impacts the overall levels of improvement. It was the leader's influence on teachers,

teaching quality, school culture and climate that acted as the conduit for delivering school academic improvement (Sammons *et al.,* 2011).   More recently, Boberg and Bourgeois (2016) put forward a model of leadership that advanced the work of Leithwood and Jantzi (2006). The study, involving 569 teachers in the south-central United States, supported previous understandings that leadership has an indirect impact on raising student attainment based on behaviours that promote cooperation, commitment and capacity. Boberg and Bourgeois (2016) further the discussion by associating leadership practices with its impact on teacher collective efficacy and conclude that "the more teachers felt supported in their collective capabilities the more optimistic they are about reaching their students" (2016, p. 369).

Ultimately, a subject leader, key stage coordinator or teacher is permitted the space for manoeuvre and opportunities to enact leadership behaviours and practices through the prevailing philosophy of the headteacher. The culture, structures, vision and trust, established by the headteacher, guides the work of the school and its staff. Though a range of leadership approaches enable headteachers to respond to changing or challenging circumstances (Bush and Glover, 2014), by influencing staff motivation, commitment and working conditions, school leaders can indirectly improve teaching and learning (Muijs *et al.,* 2014).

4.2.2.3 Policy Reform and the Teacher

Different authors have theorised the interplay between teacher practice and reform (Cuban, 2013; Pedder and Opfer, 2011) with arguments that explored the extent to which teacher practice changed over time and whether reform goals were met. This section explores the place of teachers in the context of reform, policy enactment and its ability to influence teacher classroom practice.

For more than a decade research literature (Winter, 2017; Swann *et al.,* 2010; Whitty, 2006; Day *et al.,* 2005) has suggested that from the UK perspective, educational reform played a significant role in shaping and changing teacher classroom practice in different ways. The highly regulated and controlled national policy framework used standardisation to create boundaries for teachers, students and schools (Pedder and Opfer, 2011). Subsequently, policies are issued and implemented on the ground by "policy actors… those involved in making meaning and constructing responses to policy through the processes of interpretation and translation" (Ball *et al.,* 2011a, p. 625). Policy actors can take an active rather than a passive role (Heineke *et al.,* 2015) and can be local authorities, senior leadership, pastoral leaders or teachers.

As key actors in the policy process teachers are both an agent and a subject of policy enactments through the contextualised interpretation, translation and decision-making of policy text (Braun *et al.,* 2011). Consequently, teachers, as policy actors, are thereby positioned, whether this is by virtue of their experience, subject department or the school phase in which they work (Braun *et al.,* 2011). Whilst not fixed or mutually exclusive, this positioning becomes relevant when exploring how teachers engage with teaching science in primary school or secondary school. This positioning also helps to explain the situated realities relative to the education policies on curriculum change and standardisation (Ball *et al.,* 2011a). Participation in the policy process has been found to vary amongst teachers. Newly Qualified Teachers (NQTs) displayed a form of policy dependency, using high levels of compliance as coping mechanisms, whereas subject leaders acted as a policy shield, interpreting policies for their team (Maguire *et al.,* 2015). Teacher response to curriculum and policy changes are therefore multi-faceted and varied (Ryder *et al.,* 2018). This goes back to the idea that "putting policy into practice is creative,

sophisticated and complex but also a constrained process" (Braun *et al.,* 2011, p. 586), especially when refracted through the complicated, multi-layered context of a school. The type of policy, mandated or recommended, plays an equally important role relative to the social and cultural interpretations of policy text and this contributes to an understanding of how the policy is adopted and by whom (Maguire *et al.,* 2015). The school context makes this point all the more relevant, as there are a wide range of practitioners, with differing roles, negotiating aspects of policy that apply to them in that role, formally and informally.

Different types of policy co-exist in the school environment, some more dominant than others. A look at any school website will reveal a list of policy documents covering many aspects of school life, on for example, homework, uniform or school trips; Braun *et al.,* (2010) reported finding over 177 different policies across their four case-study schools alone. Of the array of documentation introduced to schools, some policies were active, others were inactive and many ran concurrently, leading to potentially competing initiatives and challenges to resources. For the individual subject teacher, the top-down, externally mandated reforms such as those attributed to the national curriculum vie for attention with the internal, school or department policies (Ryder and Banner, 2013). The non-negotiable high-stakes policies such as child safeguarding or reforms to the curriculum and examination system, are upfront and visible even among the plethora of central government documentation and are more immune to the discretionary actions of staff (Maguire *et al.,* 2015).

The timescale for implementation can be a significant factor in curriculum policy change (Ryder *et al.,* 2018). Subject-specific change occurred in the context of other school constraints and interacted with differing policy objectives. It was

suggested that teachers may take 4-5 years to develop their classroom enactment in response to a new policy (Ryder and Banner, 2013). This period of stability ensured that teachers' own learning, understanding and practice was strengthened, and that policy enactment became routine and seen as a worthwhile investment of time and energy. Studies which looked at the introduction of new elements to the science curriculum such as scientific literacy (Ryder and Banner, 2013); the scientific inquiry strand (Sc1) (Jenkins, 2000) and the new subject, Earth Science, (King, 2001) concluded that changes to the curriculum at GCSE altered the practices of secondary science teachers in response. Furthermore, test-based accountability, albeit not the only top-down policy reform, frequently correlated with changes to teaching practice, the amount of science delivered and teacher satisfaction, although not always negatively (Anderson, 2012). Together, the studies pointed to the need for teachers to acquire new knowledge and develop associated pedagogies, some of which may take them away from their subject expertise (Ryder, 2015).

Ball described two conceptualisations of policy, "policy as text and policy as discourse" (1993, p. 10). Policy as text is "primarily a matter of language in speech and document and action in social events" (Ball, 2015, p. 307). A prime example of "policy as text" would be the National Curriculum, a key document structuring teacher's work. When speaking of "policy as discourse", Ball (2015) then draws upon Foucault's construction of discourse, as forming the objects of which we speak (Foucault, 1972). I interpret this understanding of discourse to include the way education is perceived, communicated and organised, as exemplified by publishing exam results and league tables; lavish open-evenings and brochures; learning walks and drop-in observations. These, among other things, form part of a teacher's working life. From a critical realist perspective, education policy acts as a mechanism to generate change and achieve government outcomes, therefore,

teachers' practice in the face of mandated policy, then becomes both mechanism and outcome. The government aims for education are clearly stated (DfE and Gove, 2013; DfE, 2010b, 2010a), the teacher is identified as the means through which the goals, such as raising standards, will be met. However, the intended outcome is dependent on how teachers internalise and integrate these aims into their work.

4.3 Summary

The factors that construct, shape and change teacher practice are multiple, varied and complex. Rooted in the individual as a function of their childhood educational experiences, their epistemological view on the nature of science, their self-efficacy and beliefs about themselves as science teachers, teaching practices are themselves not immutable. A teacher's classroom practice is subject to change as they respond to both internal and external factors influencing them as individuals. How teachers use their agency and discretion to act in ways that enhance student learning while simultaneously meeting the, often top-down demands of educational policy will continue to be an area of debate and further research. My research then sits within the literature in examining the changes to teacher practices in response to the most recent top-down mandated changes to the science curriculum, its assessment and accountability regime. The issues raised from the literature indicate that teachers' practice is contextual and individual, which makes it challenging to identify precisely what makes it "work". Teacher practice, teacher learning and the teaching environment remain at the heart of raising student engagement and attainment, but their effectiveness appears to rely on transformational leadership behaviours that encourage collective working practices.

The relationship between science teaching and "policy" is complicated and fluid. Different types of policy warrant different levels of engagement, interpretation and

integration. The impact of policy is therefore mediated through the quality of the school leadership and the extent to which teachers are given the time and space to engage as learners themselves. The dilemma arises when attempting to untangle how different elements of the reform agenda impact the complex classroom environment. In chapter 5, I draw upon the concept of agency and construct an identity for teachers as "Street-Level Bureaucrats" to support the theorisation on teachers' responses to education reform and the accountability regime.

Chapter 5: Theoretical Framework

Introduction

Chapter 2 outlined the context in terms of the marketisation, standardisation, and accountability discourses which frame England's education system. The literature review in chapter 4 summarised the ideas surrounding teacher classroom practice, what it is, and how it is shaped and changed. Bringing together the ideas explored in these two chapters generates an understanding of the dynamic relationship between internal and external factors that impact teacher practice and student attainment. The first part of this chapter introduces an argument that places teachers in the role of Street-Level Bureaucrats (Lipsky, 2010). The second section of the chapter discusses the concept of agency as defined by Emirbayer & Mische (1998). Through these theories, I propose a framework with which to explore teacher's perspectives on science education reforms, their autonomy and discretionary decision-making.

5.1 Teachers as Street-level Bureaucrats

Almost two decades ago, it was argued that schools possessed many of the organisational structures featured in a bureaucracy, including a hierarchy of authority, division of labour, objective standards, rules and regulations (Hoy, 2002). Cast as bureaucracies that work through hierarchical structures, schools "process" a large number of people with a degree of efficiency, often using "one size fits all" solutions imposed through the implementation of external policy (Hoy, 2002; Lieberman, 2000). This framing has remained the case even in the shift towards greater distributed leadership as discussed in chapter 4 (Lumby, 2013; Hartley, 2010).

Teachers are among a group of public sector workers referred to as "Street-level Bureaucrats", a term first introduced by Michael Lipsky (2010) in 1980. It describes

an analytical framework used to examine the similarities and differences in the collective behaviour of public sector professionals, such as police, social workers, nurses, and doctors. The commonality found among Street-Level Bureaucrats is in their ability to provide benefits and allocate sanctions through the stereotyping and mass processing of their service users (Gilson, 2015). Teachers teach whole-classes in established routines; devise strategies to overcome work-place limitations; adjust their practice and make judgments to cope with the uncertainties encompassed in their role. Thought to distort the service ideals and pure policy aims, these "unsanctioned coping mechanisms" (Lipsky, 2010, p. xv) helped to compensate for the ever-increasing demands found in the working environment. Ultimately, the actions and routines of Street-Level Bureaucrats then become the public policies carried out. Hence, teachers are no longer seen as mere education policy implementers or enactors but as policymakers, pivotal actors, moulding public policy as they delivered public services (Adami, 2014).

Discretion is about making decisions involving personal judgment in the assessment of a situation. It is a space in which individual choice is exercised in the interpretation of the rules and is used to meet the needs of competing demands (Loyens and Maesschalck, 2010). Lipsky (2010) used the concept of discretion as a unifying characteristic in his definition of the Street-Level Bureaucrat. While seeming not to preference one definition over another, Lipsky unpicked how public sector workers enacted discretion in the course of their work. Street-Level Bureaucrats as professionals were expected to use their discretion to respond to unforeseen incidents and make on the spot decisions for the benefit of both the service user and the organisation in which they work. Tummers and Bekkers (2014) further suggested that from the top-down perspective of policy implementation, the use of discretion at street-level led to the worker's pursuit of their own goals but the bottom-up perspective was seen as an evitable necessity if policy programmes are to be

effective. In essence, discretion is the mechanism through which Street-Level Bureaucrats applied the "rules" of a policy in different circumstances, to meet the needs of large numbers of service users, whilst satisfying the performance requirements of their organisations. Street-Level Bureaucrats mediated the relationship between the state and the citizen; the "face" of public policy in the daily reality of most people's experience (Gilson, 2015). Policy is no longer an abstract, faceless document but one that a person experiences through the decisions taken by the Street-Level Bureaucrat.

This position, which places teachers as the front-line face of education policy, was used in this research as an analytical tool to explore teachers' perspectives in response to reforms to science curriculum and its assessment. In this context, the students, as the non-voluntary clients or service users experienced reforms to education policy through the decisions made by their teachers. In contrast to Lipsky's suggestion, teachers did not distance themselves from their students, indicating that the framing of teachers as Street-Level Bureaucrats is nuanced and reflected the particular interactions found in schools. Exploring classroom practice shone a light on the extent to which "benign modes of mass processing" were used that permit teachers to deal with their students successfully (Lipsky, 2010, p. xiv). Street-Level Bureaucrat behaviour is complex and contextual. Specialisation among the variety of roles and expectations across the teaching profession mean that discretionary powers are not evenly distributed or easily attributable (Lipsky, 2010). Teachers in primary school settings, as non-subject specialists with weaker discretionary powers, faced different dilemmas to those of their secondary science counterparts. Consequently, the Head of Science, with stronger discretionary powers, played a significant role in deciding, for instance, which GCSE exam board the students will follow. Those Street-Level Bureaucrats with substantial discretion decided the criteria for decision-

making and made the final decisions, meanwhile those with weak discretion worked within the school and government policy "rules" (Gilson, 2015).

The original theorisation of teachers as Street-Level Bureaucrats came at a time before the Education Reform Act of 1988. The marketisation of education has since become one mechanism of the new public management through which successive governments have introduced policy (Taylor, 2007). As such, the notion that, as Street-Level Bureaucrats in Lipsky's 1980 original description, teachers used their discretion in the face of vague policy documentation holds less well due to the introduction of prescriptive curricular documents, standardised school organisation into key stages and standardised assessments which allow less scope for individual interpretation and decision making. The introduction of Ofsted inspections, teacher performance management and publication of performance tables further challenged the definition of teachers as being Street-Level Bureaucrats. In the context of these control mechanisms "discretion is a relative concept" (Lipsky, 2010, p. 15) and applying this fluidity gave scope to analyse the contextually driven behaviour of teachers as public sector workers.

Before the ERA, the only specific requirement for schools was to teach Religious Education, the curriculum for children aged 5 to 14 was largely decided by teachers in line with the available textbooks or teacher-selected public examination syllabi (House of Commons, 2009). With the introduction of the national curriculum and associated assessments, the previous high degree of autonomy and discretion were reduced. Subsequently, the regime of increased accountability, perceived as an attack on teachers' professional discretion, altered what teachers did in the classroom (Taylor, 2007). Although alternative forms of accountability ran in tandem with those driven by top-down policy (Hupe and Hill, 2007), alignment with the goals of the measurement system, put in place via the inspection and performance regime,

was expected from both teachers and leadership. However, being less able to exercise discretionary decision-making did not necessarily mean that teachers felt disempowered and the extent to which they exercised their autonomy pre-ERA may have been exaggerated (Taylor, 2007).

As discussed earlier, classrooms are complex often unpredictable spaces, and teaching cannot be reduced to a formulaic process. Arguably, teachers' personal views, pedagogic skills and priorities come into play as Street-Level Bureaucrats, seeking the best course of action for their students (Taylor, 2007). When teachers were able to use discretion to modify, adapt and integrate education policy in a way that benefited students, there was greater willingness to implement it, thereby raising the effectiveness of the policy (Tummers and Bekkers, 2014). With the introduction of the new national curriculum and the teaching of new specifications in science since 2016, it could be suggested that teachers have lost some of their valued decision-making power, possibly becoming more policy dependent. Teachers do not face the simple choice between either blindly following the rules to get children through exams and ready for the jobs market *or* developing the whole child through constructivist methods, both are equally important. But it has been found that the measurable evidence of success required by managers and policy-makers discouraged a more developmental approach to teaching (Taylor, 2007). Arguably, Street-Level Bureaucrat's practices are defined more by pragmatic improvisation than discretionary decision-making, relying on a delicate balance between following the prescribed directives and exercising agency, particularly when policy conflicts with teachers' beliefs and understanding of the purpose of education (Maynard-Moody and Musheno, 2012).

Lipsky's perspective offers what Evans and Harris (2006) suggested is a tentative framework for the exploration of how all street-level bureaucracies work which can

be used as a starting point for the analysis of complex public sector organisations, like schools. Introducing the concept of agency adds richness and meaning to the implications of street-level work and in support of the exploration of teacher's behaviour in the face policy implementation.

5.2 Teacher Agency

The two words "discretion" and "agency" are similar concepts with nuanced differences. 'Discretion' is an authority granted to front-line workers to make decisions within the rules of law and policy (Maynard-Moody and Musheno, 2012). A definition of agency is more difficult as it is an inherent quality of what it is to be human, rooted in the individual irrespective of the position that they hold. Everyone can exhibit a sense of agency, but not all can be described as a Street-Level Bureaucrat with discretionary power to make decisions that will benefit or sanction others. This is not to say that in making decisions about conforming to prescribed practice or creatively responding to an individual student's needs is not a feature of teacher agency but "their position, training and work shape(d) the nature and expression of their agency" (Maynard-Moody and Musheno, 2012, p. S19).

Commonly argued in opposition to structure, the Dictionary of Social Sciences, defined agency as "The capacity for autonomous social action... the ability of actors to operate independently of the determining constraints of social structure" (Calhoun, 2002). Whilst agency places the individual at the centre of analysis, structure, is aligned with an adherence to norms and values that constrain the individual and might only be visible through their conformity to cultural rules (Jones, 2003) or in terms of the context in which events acquired meaning (Hay, 2002). Through structures, one can expect order and routine to feature in social, political and economic events. Significantly this structure-agency debate has been

ongoing since the 1970s (Calhoun, 2002) therefore, as a lens through which to view the social world it is problematic (Biesta and Tedder, 2006; Hay, 2002; McAnulla, 2002). Social theorists introduced ways to understand the interaction as either a duality or a dualism. In Structuration Theory (Giddens, 1984), the role of agency is embedded within a duality with structure; the knowledgeable agent produces and reproduces the social world and can exert causal powers. Archer (1995, 2003, 2010) posited the interplay of structure and agency as an analytical dualism in which social theory has yet to reach an ontological consensus. Only a "slim agreement" existed which reduced the argument to a simple statement where structure is objective and agency entails subjectivity (2003, pp. 1–2). As divided as this may seem, Emirbayer and Mische (1998) suggested that attempts to define and theorise the concept of agency omit crucial aspects, which caused confusion and failed to distinguish the complexity of agency manifest in its own right.

Lasky (2005) in her study on teacher agency and identity suggested that agency started with a belief "that human beings have the ability to influence their lives and environment while they are also shaped by social and individual factors", (2005, p. 900). Nevertheless, there was a sense that teacher agency had not been well researched (Pyhältö *et al.,* 2015) with very little theory development in the field (Biesta *et al.,* 2015). Intertwined with identity (Buchanan, 2015), teacher agency situated individuals as active agents whose efforts, choices and intentional action made a difference and as pedagogic experts capable of managing learning (Toom *et al.,* 2015). However, an understanding of agency, as a social process which cannot be uncoupled from structural practices and norms, was not the only way to investigate the concept (Biesta and Tedder, 2006). For the purposes of this research, I have used the theorisation of agency by Emirbayer and Mische (1998) to explore teacher agency.

Emir and Mische define agency as

> … the temporally constructed engagement by actors of different structural environments—the temporal-relational contexts of action—which, through the interplay of habit, imagination, and judgement, both reproduces and transforms those structures in interactive response to the problems posed by changing historical situations. (Emirbayer and Mische, 1998, p. 970)

This definition encapsulates what teachers do in the classroom, how practice changes in different contexts and with different relationships over time. I used this work chiefly because of its analytical power to explain the agency of teachers in terms of the interplay between routine, purpose and judgement and; to bring about an understanding of agency in terms of problem-solving (Biesta and Tedder, 2006).

Referred to by the authors as a Chordal Triad, three different dimensions constitute human agency; iterational, projective and practical-evaluative. Each corresponds to the different temporal orientations, towards the past, the future and the present. Within each dimension is an internal structure which also orientates towards past, future and present reflecting that all forms of agency are embedded in the flow of time. Each dimension and its internal structure are briefly discussed next.

5.2.1 Iteration

 The iterational element refers to the:

> selective reactivation by actors of past patterns of thought and action, as routinely incorporated in practical activity, thereby giving stability and order to social universes and helping to sustain identities, interactions, and institutions over time. (Emirbayer and Mische, 1998, p. 971).

This definition locates teacher agency and classroom practices in terms of position, past experiences and beliefs (Biesta *et al.,* 2015). The internal structure of the iterational orientation enables elaboration of past action through the process of "selective recall from the past". Emirbayer and Mische distinguish three primary behaviours as *Selective Attention* – directing attention to single out an appropriate response; *Recognition of Types* – use simple models to identify patterns of previous experience and predict future occurrence and *Categorical Location* - locate typical

experiences in relation to others, contexts and events (1998, p979). Finally,

Expectations- orientated to the future and *Manoeuvre* which orientate improvised

actions in the present operate as secondary tones.

5.2.2 Projective

The projective element of the triad is described as

> the imaginative generation by actors of possible future trajectories of action, in which received structures of thought and action may be creatively reconfigured in relation to actors' hopes, fears, and desires for the future. (Emirbayer and Mische, 1998, p. 971).

This definition reflects teachers' possible future actions in response to decisions

made or imposed upon them. The projective orientation links to the aspirations that

a teacher holds for both themselves and their students (Priestley *et al.*, 2013). The

internal structure of the projective orientation, like that of the iterative, maintains a

connection with the other time frames in its secondary tones of *Identification*- a

retrospective process and *Experimentation* – tentative responses to presently

emerging situations (Emirbayer and Mische, 1998, p. 988). Conceived as overlapping,

working synergistically with each other within the social context the dominant tones

of the future-orientated dimension are *Narrative construction* – identification of

typical trajectories located in causal or temporal sequences; *Symbolic recomposition*

– reimagining a variety of possibilities based on different alternative scenarios and

Hypothetical resolution – proposing responses to concerns arising from lived

conflicts.

5.2.3 Practical Evaluative

The final element of the triad is concerned with practical evaluation, it entails

> the capacity of actors to make practical and normative judgments among alternative possible trajectories of action, in response to the emerging demands, dilemmas, and ambiguities of presently evolving situations (ibid. p.971).

This element is orientated towards the present and day-to-day decisions and actions teachers make using their discretion, professional judgement and reflexivity. As an internal structure, the practical-evaluative element takes as its primary tones, *Problematisation* - a recognition that the particular situation at hand is somehow ambiguous, unsettled, or unresolved. *Decision* – making choices, resolution and the move to concrete action and *Execution* –Doing what was planned in the way that it was intended. Each contextualises what is projected and what has become habit within the concrete situation of the present. The secondary tone which relates past experience to the present is *Characterisation* – defining the problem based on schemas and types from past experience, the relationship to the future is *Deliberation* – weighing plausible choices in the light of practical perception and understanding.

As analytical distinctions, the constitutive elements of agency are found in varying degrees in any concrete empirical instance of action (Emirbayer and Mische, 1998) and as such I believe that it is the appropriate tool to explore the data from my research as it enables the temporal orientation of teacher's actions.

5.3 Illustrative Theoretical Model

Combining the framework derived from the work of Lipsky (2010), which positions teachers as Street-Level Bureaucrats making decisions in a rule-driven context, with the concept of agency developed by Emirbayer and Mische (1998) results in the model illustrated below (Figure 2). The model was used as an analytical device to explore the dynamic interplay between the past, future and present actions of teachers, relative to others and across different structural contexts. As each temporal location plays a minor role in the other, the integration recognised that when we are in the present, we use experiences from the past and aspirations for the future to guide our actions.

The iterational dimension is crucial to this analysis of agency, reflecting the idea that an individual's projective (future-orientated) and practical-evaluative (present judgement) actions are deeply grounded in their day-to-day, taken for granted habits, routines and experience built over time. Habit and routine are typical of social spaces such as schools, as it allows for the mass processing of large numbers of students (Lipsky, 2010). Habits and routines can change as a result of conscious response and interaction to changing contexts and expectations; for example, a science teacher might previously have written the letter grade (A*-G) on student's marked work, changes to school policy dictate that 'comment only' marking is now used, however the teacher might still record the letter grade in their mark-book.

When attempting to analyse the projective domain, Emirbayer and Mische argued that a crucial factor in human action is an imaginative engagement with the future. Teachers' forward-facing orientation takes them on a timeline, in which they "construct images of where they think they are going, where they want to go, and how they can get there from where they are at the present" (1998, p. 984). At a time of curriculum change, for instance, this manifests in the hopes for their students' future exams success, planning for teaching a new topic the second time or intentions for meeting performance management targets.

The third dimension of teacher agency focusses on response to the demands of the present and how teachers make situationally based judgments, using discretion in the face of ambiguity, uncertainty or conflict (Emirbayer and Mische, 1998). Context plays a primary role in the practical-evaluative dimension by determining what actions are taken, if a problematic situation is encountered, a teacher can select and apply their practical wisdom or discretion, to arrive at a normative solution. Deliberating over the choice of solution becomes part of the day-to-day

actions of teachers, responding to student's questions and deciding which pedagogic practices to use to instigate learning in a given timeframe.

All three aspects of the triad exist at any given time and although one may feature more strongly than the others depending on the context, this comes to shape how actors understand and talk about their orientations (Biesta and Tedder, 2006). The diagram (Figure 2) represents the different elements of the research, with double-headed arrows to show the interactions between the internal and external factors that may affect teachers' practice.

Figure 2: Theoretical Framework illustrating the association between teacher agency, discretion and practice and the structural factors in education
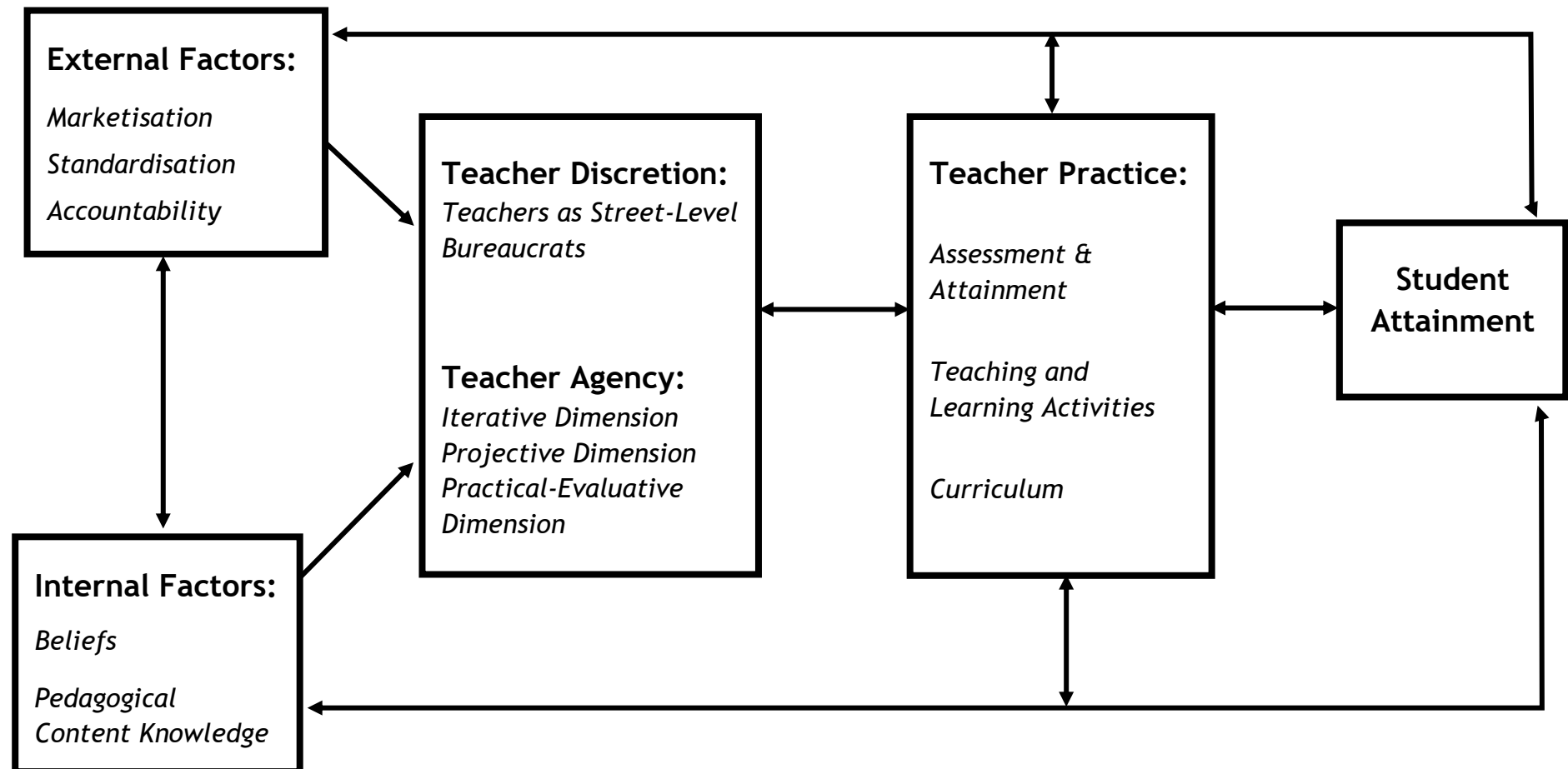
In the top left box are the external factors of marketisation, standardisation and accountability. These mechanisms, as discussed in chapter 2, frame the education context and determine what is taught and how performance is measured in maintained schools. These external factors, to some extent, limit teachers' work but do not guide everything they do in the classroom. The internal factors, in the bottom left box, play a part in teacher classroom practice as discussed in chapter 4. Beliefs about their capabilities, about the nature of science and different types of students, are personal constructs and determine how an individual teacher approaches their class. In addition to this, subject knowledge and the ability to transform this into lessons that promote learning for different students in changing circumstances is reflected in a teacher's pedagogical content knowledge. The model, in Figure 2, suggests that the external and internal factors are directly linked to teacher discretion and agency and to student attainment which in turn impacts teacher classroom practice.

In developing the framework, I drew upon the notion that teachers' work is multi-faceted. It is collaborative yet idiographic, whether undertaking joint planning in a team or marking books at home. When a teacher closes their classroom door, they are alone with their students; they are the leader of learning, making micro-decisions, often on a minute-by-minute basis. They are reflexive, drawing on their own experiences, pedagogic subject knowledge and adherence to school "rules" in order to manage their work effectively.

5.4 Summary

I assert that Lipsky's articulation of teachers as Street-Level Bureaucrats alone is insufficient to capture this sensibility. Likewise, the sociological understanding of agency may not fully reflect the realistic, often pragmatic decisions that teachers

are involved in every day. In combination, the theories complement each other allowing for a contextually and temporally nuanced level of analysis. Taylor (2007) suggested that overall, teacher's as Street-Level Bureaucrats in the sense of Lipsky's policy-making through their professional discretion has been "severely compromised by education reform" (2007, p. 569). My research aims to revisit and re-evaluate this area of inquiry in the light of the recent changes to the national curriculum and its assessment. Through seeking the answers to the research questions below, I will bring new knowledge to the field.

5.5 Research Questions

The model representing the theoretical framework outlined the different components of my research study. In contributing to the framework, chapter 2 has discussed the external contextual factors and chapter 4 has explored the literature associated with the internal factors that may impact teacher practice. Two secondary research questions address the different aspects of the study concerned with teacher practice and student attainment. An overarching research question then brings this together by asking

"How have recent reforms to the science curriculum and its assessment affected student attainment and science teacher classroom practice?"

Secondary questions:

Quantitative:

"What do historical trends in student's end of key stage science attainment since 2008 show, and how does this reflect policies to raise attainment?"

Qualitative:

"What are teachers' perspectives on the 2014 science education reforms and how has this affected their practice?"

Next, the methodology chapter explains and justifies how the data was collected in order to address the research questions raised.

Chapter 6 Methodology

Introduction

This chapter outlines my philosophical viewpoint, ontological and epistemological assumptions derived from personal experiences and education. These are a property of who I am as a researcher, not easily changed, they influenced the actions taken in theorising and conducting this research (Maxwell and Mittapalli, 2010). Exploring these fundamental truths established my position and served to explain and justify the critical realist perspective taken throughout the study. Taking a particular philosophical stance also provided the rationale for using a mixed methods design, directed the choice of theory, and framed the analysis.

The chapter also explains how the secondary data on student attainment were obtained from the National Pupil Database (NPD), including details on how the variables were selected and used to generate appropriate statistics. A description of the development of the teacher questionnaire and interview tools follows, after which, the strategies used to analyse the survey and qualitative data are discussed. In the penultimate section of this methodology chapter, I describe the ethical considerations involved when conducting research in schools and working with individual pupil data. I close by explaining my positionality and how this influenced the nature of the knowledge produced.

6.1 Research Design

6.1.1 Ontological and Epistemological Position

Conducting research is an endeavour to seek the truth about a phenomenon and gain a better understanding of the world. There are several approaches which can be taken in pursuit of this aim, these rely on the different conceptions of social reality and the different ways of interpreting it (Cohen *et al.,* 2017). Therefore, as an

inquirer I am guided by Lincoln *et al.,* (2018) to make explicit my view of the world and that in which the research is situated and by doing so, to address three questions that frame the parameters of the study. The question of ontology, which deals with the nature of being, is followed by the issues of epistemology, which address the ideas around how we understand what constitutes knowledge; finally, the methods, explain how the knowledge will be gathered.

Ontology asks, is there a real world "out there", independent of our knowledge of it or is it a product of our consciousness? (Cohen *et al.,* 2017). Realist perspectives about the nature of reality take the position that the world is real and knowable as it is, whereas, constructivists apprehend reality from multiple, intangible, mental constructions (Lincoln *et al.,* 2018). For example, in response to a question about gender differences; a realist would assert that there is a real difference between boys and girls foundational to our being. Meanwhile a constructivist might argue that the differences between boys and girls are socially constructed. A researcher's ontological position outlines their view about the nature of the world and is fundamental to their epistemology position. The epistemological viewpoint postulates ideas about what can be known about the world and how we can know it. Taking a realist position means that knowledge can be uncovered using objectivist means, conversely, seeing knowledge as individually or socially constructed means that an understanding of it can be captured through the subjectivities of those researched.   Therefore, a realist may research the differences between boys and girls by measuring the differences in test scores; whilst a constructivist might explore what the test scores meant to each group.

The different perspectives or paradigms (Lincoln *et al.,* 2018) specify the researcher's assumptions about reality that determine not only the nature of

knowledge and how knowledge is accumulated but also the values, ethics and positionality of the researcher. Whilst new ways of conceptualising research have emerged, e.g. Complexity Theory (Cohen *et al.,* 2017), each paradigm, Positivism, Post-positivism, Critical Theory, and Constructivism, for instance, has inherent strengths and weaknesses (see Lincoln *et al.,* 2018, p. 110; Tikly, 2015; Niglas, 2010). Research is often positioned as a dichotomy, either quantitative or qualitative, the answer to which can determine the choice of methods used. This dichotomous boundary, however, places unnecessary restrictions on how a question can be legitimately investigated, clustering research design decisions together, and driving researchers down a particular path (Biesta, 2010a).

6.1.2 Philosophical Viewpoint

My experience of research originates from the physical sciences where I routinely used universal laws to generate hypotheses that could be empirically tested. For instance, teaching about rates of reaction uses a combination of different concepts and theories which when tested by students in the school laboratories rarely yielded the expected outcomes. The students interpreted the results in their way and theorised about the results according to their understanding of scientific knowledge. In keeping with my understanding of the physical sciences, the world is a separate entity, outside of my ability to completely know it, and what we can know about the real world is imperfect and subject to our personal interpretation (Maxwell, 2018). It was not until I embarked on my PhD journey that I could put the name "Critical Realism", to this way of thinking about and viewing the sociological world.

Stemming from the work of Bhaskar (2008) the critical realist philosophy of science was first postulated to explore the natural sciences but has since been developed as a broad logic of inquiry and as an analytical tool by scholars across the range of

social sciences (Maxwell and Mittapalli, 2010; Mingers, 2006; Pawson, 2006). The position is founded on a realist ontology, which states that there is a real world existing independent of perceptions, theories and constructions. This is bounded by a constructivist epistemology, where any understanding of the world is determined from the individual's perspective and is therefore a social construction (Maxwell and Mittapalli, 2010). This epistemological and ontological viewpoint is linked to and determines the way that research is written, the language and the way it is used to construct the objects of research (Dunne *et al.,* 2005). At first glance this combination of realist ontology and constructionist epistemology may seem at odds with the paradigms mentioned above which guide qualitative and quantitative research methods (Gorard, 2010). Principally this is because qualitative and quantitative research have fundamental differences in their ontological and epistemological positions and in the way that they approach what is important and inherent in a phenomenon (Lincoln *et al.,* 2018).

6.1.3 Critical Realism

As a lens for understanding the world, Critical Realism supports the beliefs that an individual's social and physical context is real and has a real causal influence on how they experience their world (Maxwell and Mittapalli, 2010). A realist stance accepts that all things, whether physically manifested or derived from meaning such as concepts, beliefs, intentions, and values make up reality. Separating ontology from epistemology in this way enables research that can explore causation within qualitative research in a way that constructivist viewpoints cannot. The term 'causation' described in this way is not synonymous with that in the natural sciences but acknowledges the reality that mental phenomena are causally connected to physical ones through mechanisms and processes that result in particular outcomes (Maxwell, 2018). Having said this, the belief that an independent reality exists does

not assume that absolute knowledge of how reality works is ever possible, and any such attempts are always fallible (Scott, 2005).

In social science, from a realist perspective, objects, actors and social structures have generative causal powers (Pawson, 2006). In this empirical study, I conceive that the teachers have causal powers in their ability to support student learning through their experience, beliefs, and qualifications. Education policy has casual powers in the ability to structure the national curriculum, dictate what is taught, what is examined, and what is deemed important to measure. This does not mean to say that all teachers will be effective in the classroom to bring about cognitive gains in their students, or that newly introduced education policy will deliver an increased percentage of students achieving a higher grade in science GCSE. However, causal powers (and their opposite, causal liabilities) do not imply cause and effect, but can be taken as ways of acting or "mechanisms" (Sayer, 2010, p. 105). These causal powers are not fixed and exist by virtue of the object, structure or individual. For instance, an experienced biology teacher may decide to retrain through a Subject Knowledge Enhancement course in order to confidently teach A-level physics. However, opportunities to teach physics A-level in their current school may not exist as too few students have signed up for the course, meaning that the teacher is unable to use their causal powers in this context. This illustrates how casual powers may go unused and unexercised. They are not simply inherent within an individual but are manifest in the structures and relations of the social world; they may be overt, obvious and explicitly used to negotiate the world and guide action.

Staying with the illustration above, retraining in a particular curriculum subject will not always bring the expected teaching role in that subject. Causal powers and their effects are contingent upon context and conditions in which they operate (Sayer,

2010) and although this research is not a formal evaluation study and is limited in scale, it borrows from the components of realist causal explanations devised by Pawson and Tilley (1997) to explore the Mechanism + Context = Outcome dynamic. Teachers have a variety of ways of working that are tied into their causal powers; they work in a variety of contexts and are responsible for numerous outcomes. Their agency and discretion interact with the structured social world of the school environment and with the students that they teach. All of this takes place within a complex socio-political context based on measurement and accountability. Despite Hammersley's (2014) suggestion that critical realism forwards value judgements, the realist perspective affords me the analytical power to explore these interactions and the issues surrounding them.

Overall, a critical realist stance emphasises the importance of context. In a study such as this, which is to some extent theory-based, there are no postulated general laws that will govern the outcomes. The work is situated both temporally and contextually in a specific situation and aims to understand the causal interactions specific to that particular situation. Through this context, the terms quantitative and qualitative describe the types of data that can be derived. Therefore, the design aims to draw evidence and data from different sources, using different methods to develop understanding. The research design was structured to maximise the possibility of generating sufficient knowledge to respond to the research questions (Gorard, 2010), without being restricted by the need to cling to methodological traditions. A mixed methods design will ensure that I do more than simply work deductively to collect the facts to confirm a hypothesis, it allows me to accommodate the existence of a subjective way of knowing.

6.2 Methods

6.2.1 Justification for mixed methods

To overcome the "contentious philosophical issues involved in the paradigm wars over qualitative and quantitative research", Maxwell and Mittapalli (2010, p. 146) suggest taking a realist stance through a mixed methods research design. Analysing quantitative data requires subjective and interpretative decision making in the selection of variables and defining attributes of a concept; whilst qualitative research requires decision making and interpretation involving the use of "numbers" in terms of cases and specifics (Biesta, 2010a).

In my study, the quantitative data, as derived from student attainment at KS2 and KS4 is factual and treated as an entity in its own right. However, as an outcome of education policy what it measures and what it represents are socially constructed and interpretive outcomes, not value-free. A cohort of students have certain facts generated about them and their learning, this is interpreted by different actors, in different ways and leads to different responses from different quarters. The students, parents, the school, the local authority, and the education department view the results from their understanding of reality. The student attainment data is also conceived as having causal powers that can impact the introduction of a new phase of policy initiatives and on teacher behaviour. The frequencies, trends, and associations found in the student attainment data cannot be uncoupled from the impact that they have at classroom, school and (inter)national level. From this positioning, I ascribe the quantitative data with the ability to generate meaning through the analysis of its impact on the decisions made by both policy-makers and teachers alike.

The qualitative data and the understanding that is interpreted from the teacher interviews is also treated as an outcome of education policy and interrelated student attainment data. The teachers, as entities with causal and decision-making powers, act in ways that may be constrained or unconstrained by the context. Through dialogue with the teacher participants, the interview seeks elaboration and an enhanced understanding of their world. I draw upon the complementary strengths of both qualitative and quantitative research doctrines to reduce unintended bias and to come to conclusions not possible through using either method alone (Maxwell and Mittapalli, 2010). The social and policy context of the interconnected realities associated with student attainment data and teacher practice frame the study and in doing so, further justifies the mixed methods design.

6.2.2 Mixed method typologies

The definition of the term "mixed methods research" has been contested, Johnson *et al.,* (2007) identified 19 different definitions. Put simply, mixed-methods is as an approach to knowledge, theory and practice which takes into account multiple perspectives (Creswell, 2010; Johnson *et al.,* 2007). Seminal work by Greene *et al.,* (1989), introduced many different arguments for pursuing a mixed methods research design including:

- Triangulation - to corroborate data and increase validity
- Complementarity - to elaborate and increase meaningfulness
- Development – to use the results from one method to increase the validity of constructs
- Initiation - contradiction in the research and increases breadth and depth.
- Expansion - to extend the range of enquiry using multiple inquiry components   (p. 259)

Whilst many arguments exist that determine the purpose and strategy of a mixed methods study (Bryman, 2006, 2016), an increasing number of mixed method typologies have emerged (Creswell, 2010, 2018). These differ in how the research tools and data collection are sequenced, weighted and theorised. The data mixing, thought to be the most difficult aspect of mixed methods research, can also be achieved in several different ways. The data can be merged or kept separate; mixed during data collection, during data analysis or at the interpretation stage (Creswell, 2018). Specifying precisely how each approach and the data collected achieve the purpose of the mixed method inquiry contributes to the validity of the overall research.

6.2.3 Applying a mixed methods typology

Following Creswell's (2018) typologies, I originally planned to request and analyse the pre-existing quantitative data first in a sequential, explanatory design. With the intention that this secondary quantitative data would uncover the trends and patterns in student attainment around which to focus the interview questions. However, the complexity of requesting, receiving and processing the whole cohort National Pupil Database (NPD) information proved greater than expected, a detailed explanation of what this entailed is discussed later in this chapter. The delay in receiving the NPD data necessitated a change to the mixed methods design to avoid causing a delay to the participant recruitment stage. As a result, the framework for the study was reconceptualised, shifting the focus away from the quantitative and qualitative having equal power and status in their ability to produce the knowledge and understanding of this issue; towards affording the qualitative data the primacy in bringing meaning to the individual actors at the centre of the research. Therefore, a concurrent, embedded design was undertaken, whereby the quantitative data was collected and analysed at the same time as the qualitative data.

The secondary quantitative data from the NPD took a supporting role in the study and addressed a different but interlinked question that incorporated a different unit of analyses to that of the qualitative study. This is not to say that this process was unproblematic or open to critique (Johnson and Gray, 2010; Woolley, 2009) not least in part due to the differences in measurement and concept definition found in the qualitative and quantitative stages (Goertz and Mahoney, 2012). The research took a bottom-up approach, meaning that it was driven by the research questions (Johnson *et al.,* 2007); the mixed methods design provided the platform so these different questions could be answered (Bryman, 2006). Figure 3 illustrates the research design showing from where the data was drawn and, how it was analysed and mixed. The remainder of the chapter details the methods and limitations of the quantitative and qualitative data collection approaches.

```
┌─────────────────────┐    +    ┌─────────────────────┐
│    QUALITATIVE      │         │    quantitative     │
└─────────────────────┘         └─────────────────────┘

┌─────────────────────┐         ┌─────────────────────┐
│ Data Collection:    │         │ Data Collection:    │
│                     │         │                     │
│ Teacher             │         │ Secondary data on   │
│ Questionnaire.      │         │ student attainment  │
│ Semi-structured     │         │ from NPD            │
│ interviews.         │         │                     │
└─────────────────────┘         └─────────────────────┘
          │                               │
          ▼                               ▼
┌─────────────────────┐         ┌─────────────────────┐
│ Data Analysis:      │         │ Data Analysis:      │
│                     │         │                     │
│ Thematic and        │         │ Descriptive and     │
│ Template Analysis.  │         │ Statistical         │
│                     │         │                     │
│ Explores the        │         │ Explores the        │
│ perceptions         │         │ association         │
│ surrounding         │         │ between attainment  │
│ attainment data,    │         │ data and reform.    │
│ teacher practice    │         │                     │
│ and reform.         │         │                     │
└─────────────────────┘         └─────────────────────┘
          │                               │
           ╲                             ╱
            ╲                           ╱
             ▼                         ▼
        ┌───────────────────────────────────┐
        │ Interpretation of entire analysis │
        └───────────────────────────────────┘
```

Figure 3: The Components of the Research Design

6.3 Data Collection: Student Attainment

This section of the methodology explains the purpose and objectives for the use of secondary data as derived from the National Pupil Database (NPD). Decisions surrounding the selection of appropriate attainment indicators, their analysis, and limitations are addressed. Additionally, the contribution the analysis of this dataset made to answering my research question and to the generation of new knowledge is argued.

Firstly, the definitions for the key terms used in reporting and analysing student learning are given contributing to an understanding of how the quantitative data set was established. Secondly, the use of secondary data sources in general, their usefulness, advantages and disadvantages is briefly discussed whilst explaining the specific nature of the NPD. This describes the interpretative processes at play in deciding the range and scope of the quantitative variables selected for use as indicators of student attainment. Also included is an explanation of how the data was prepared and analysed, the statistical methods used and the limitations of the dataset.

6.3.1 Key Terms

Words such as achievement, attainment, progress, and outcomes require definition because they are used across different contexts and as different measures of performance. The term "outcomes" appears to be used as a generic descriptor that takes in a much broader range of measures surrounding the expectations for students that may encompass, for instance, aspects of physical health and mental well-being, engagement, involvement, motivation and dropout rates (OECD, 2018).

Achievement is concerned with the progress students make over time, meaning that a student may have achieved well given their starting point but not have reached the standard as described by the performance criteria. Achievement measures provide

information about what students have done to be awarded their final grade. The progress of a pupil measures how far they have come against their individual starting points as opposed to making a comparison of their attainment against children of the same age (Black-Hawkins et al., 2017). The attainment targets of the national curriculum (DfE, 2014) layout the expected level of cognitive understanding that children are expected to achieve as a result of good quality teaching. The attainment and achievement of students can therefore be tied to their performance in standardised and non-standardised assessment scenarios. Attainment measures provide information about a student's final grades, the highest level of education that an individual has completed (Baum *et al.,* 2015), it is this metric that will be used in the quantitative analysis of student data.

6.3.2 The National Pupil Database (NPD)

The National Pupil Database (NPD) holds a wide range of information about students who attend schools and colleges in England. The NPD is an amalgamation of a number of different datasets, including key stage attainment data and Schools Census data formerly known as Pupil Level Annual Schools Census (PLASC). The data are linked using a unique identifier to the examination results of pupils with information on pupil and school characteristics (UK Data Service, 2019). Covering the school student population from 2002 to 2016, the data set are made available for research purposes by making an NPD Data Request through the ONS Secure Research Service (DfE, 2018f). Although the NPD is updated annually and contains a rich range of student-level variables, data on students characteristics or attainment can be missing (Gorard, 2016; Strand *et al.,* 2015) as I also found during my analysis. The quantitative strand of this study used the secondary data provided by the NPD as one of the key sources of data on student attainment. Secondary data are data that has been collected by someone else for a particular purpose, whilst primary data are

collected by the individual researcher or research team with a particular analysis in mind (Boslaugh, 2010). Secondary data analysis uses existing data by either applying a different analytical, theoretical or statistical framework or for exploring new research questions (Smith, 2008a). Access to secondary data sources opens opportunities to extend research on existing data which then helps to set the context for the researcher's primary data (Gorard, 2002). It is for this purpose that the National Pupil Database (NPD) was used in this research.

Possible disadvantages of using secondary data are that:

- The data may contain errors that are difficult to account for, such as missing cases or data missing from individuals or particular groups.

- Differences in the way concepts are operationalised can result in measures that do not fully meet the needs of the research question.

- The data and analysis may end up being removed from the context of its original collection (Boslaugh, 2010; Smith, 2008b).

However, it is argued that the advantages can outweigh these disadvantages. For example, large, well-resourced teams of researchers employed in the generation of official statistics help to minimise possible errors (Smith, 2008b). Additionally speed, cost and the longitudinal nature of some secondary datasets can also make this type of analysis fruitful for researchers, particularly where the source of the data is one of authority (Gorard, 2002).

6.3.3 Selecting the Variables

6.3.3.1 Student Cohorts and Characteristics

This study analysed the National Pupil Database (NPD) data matched to students on the Spring Term School Census (January) between 2008 and 2016. The analysis explored the trends in the attainment of students with different characteristics and any associations with reform. This level of fine-grained analysis over and above the

aggregated data was required to avoid suppressing anomalies and to compare trends *within* the population and *between* populations of students. The following student characteristics variables were used as independent variables: -

- Gender: male or female,

- Eligible for Free School Meals: Non-FSM or FSM,

- Special Educational Need: Non-SEN or SEN,

- Ethnicity: Major Groups White, Other or Missing,

- English as an Additional Language: Non-EAL or EAL.

Table 10 shows the aggregate student population in KS2 and KS4 by their characteristics for the academic years analysed in this study. Analysing the attainment of students by gender was a useful way to examine the extent to which government policy on closing the attainment gaps between boys and girls has been successful. The variable, eligibility for free school meals (FSM), was used as a proxy measure of social deprivation or disadvantage (Macleod *et al.*, 2015; Sammons *et al.*, 2014).   Alternative measures such as Pupil Premium, FSMEver6 (eligible for FSM in the last 6 years) (DfE, 2015c) or the income deprivation affecting children index (IDACI) at local area level (Thomson and Plaister, 2019), capture information about student's deprivation but none of these measures were available for the entire period between 2008 to 2016. As students eligible for FSM represented approximately 15.4% of the population (DfE, 2019d), analysing their attainment was a useful way to explore the extent to which government policy on closing the attainment gaps between disadvantaged students and their peers has been successful. In 2019, students with special educational needs constituted approximately 14.9% of the school population, with 3.1% of the total student population having a statement or Education, Health and Care plan (DfE, 2019g). Historically, students with SEN have the largest attainment gap when compared to non-SEN students (DfE, 2019b) making this an important and necessary group to study.

Although a student's ethnicity intersects with other variables such as socio-economic group (Gorard, 2016; Strand, 2014b; Goodman and Burton, 2012), there are differences in attainment between students from different ethnic backgrounds (GOV.UK, 2019; Gayle *et al.,* 2016; Strand, 2015). Furthermore, the percentage of students of minority ethnic origin has increased over the past few years, now making up over 26% of the school population overall (DfE, 2019d) (see also Tables 8 and 9, chapter 3). Therefore, trends in student attainment by ethnic group was undertaken because these students make up a significant proportion of the student population and it was important to explore the association between reform and the persistent attainment gaps. The NPD and government statistical tables (DfE, 2018g) present ethnicity as both major and minor groupings, this study used the major ethnic groupings as the basis of the descriptive analysis: White, Mixed, Asian, Black, Chinese and Other as this provided sufficient differentiation between groups, maintaining robust sample sizes. Analysing attainment data at this level helped to give a detailed picture of the impact of reform on particular groups of students from different ethnic backgrounds. However, 4.5% of primary students and 10.7% of secondary students have no ethnicity information entered into the database and it is not possible to ascertain whether this was due parental refusal to provide the information or other unspecified errors in the data collection. No data was missing for the other student characteristic variables therefore, I decided to include students with missing ethnicity data in the descriptive and statistical analysis, as applicable, to maintain consistency of the sampled population. For the analysis of mean attainment, the data was aggregated and the students grouped into three categories "White", "Other" and "Missing". For the regression analysis, dummy variables were created using White students, as the reference group, and students of Other Ethnicities as the test variable. Where no ethnicity data was entered in the database, the cases were excluded from the analysis. A student is recorded to have EAL if they speak a

language at home other than English. As a heterogeneous group, the descriptor does not indicate a student's fluency in the English language. Identifying students with EAL and analysing their attainment, gave a picture of students who were possibly at risk of underachievement through difficulties accessing the curriculum due to language barriers (Strand *et al.,* 2015).

Table 10: Frequency and Percentage of students by key stage and student characteristics in England

Variable	KS2 2007/08 to 2014/15		KS4 2010/11 to 2015/16	
	Number	%	Number	%
Gender				
Male	1,739,025	51.1	1,935,693	51.3
Female	1,665,510	48.9	1,837,531	48.7
Eligible for free school meals (FSM)	501,302	14.7	492,792	13.1
SEN statement/school action plus (SEN)	266,351	7.8	292,652	7.8
English as an Additional Language (EAL)	455,658	13.4	449,128	11.9
Ethnic Groups				
White	2,599,784	76.4	2,727,359	72.3
Other	651,667	19.1	641,349	17.0
Missing	153,084	4.5	404,516	10.7
Total	3,404,535		3,773,224	

Note: KS2 attainment and student characteristic data is missing from years 2011/2012 and 2012/2013. Source: National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

The data were prepared for analysis by removing cases for which there was no KS4 highest point score in science (309,072) and no KS2 Science TA level (98,216), leaving a population sample of 3,464,152 for KS4 and 3,306,319 for KS2 respectively. The KS4 attainment data analysed was that of matriculating students in year 11 only, thereby avoiding counting the GCSEs from year 10 student who may have sat their GCSE examinations earlier. Matched files were created by merging the GCSE and KS2 data files for the appropriate academic year with the school census database. This

resulted in one large combined data file that was converted to an SPSS (Version 11) data file covering the years from 2007/2008 to 2015/2016 featured in the research.

6.3.3.2 Attainment Measures

At KS2, the teacher assessment data in the form of national curriculum levels in science was used as the dependent variable to make year on year comparisons, as this measure has remained relatively stable since its introduction in 1998 (Whetton, 2009).

The following KS2 attainment data were extracted from the NPD

- Key Stage 2 National Curriculum science teacher assessment (KS2 Science TA level): Level 1 to 6 from 2007/08 to 2014/15.

Figure 4 shows the distribution of the KS2 science TA levels which approximates a normal curve for the population of students in this study.
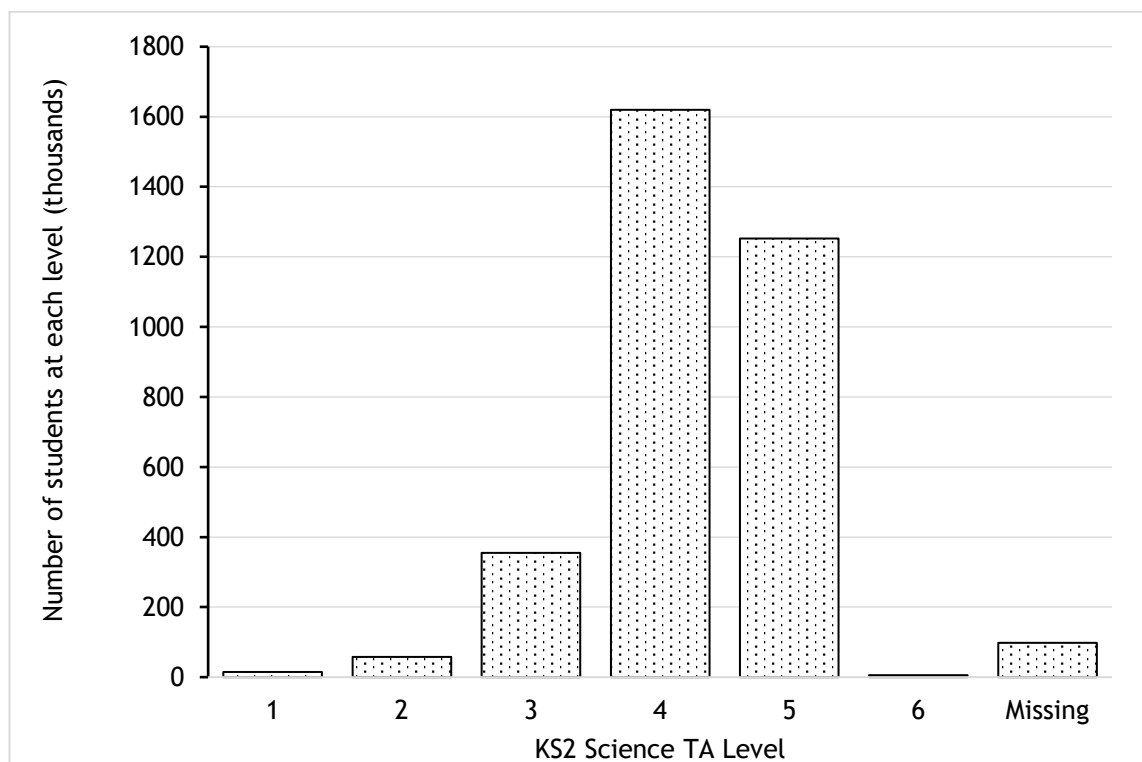
Figure 4: Distribution of KS2 Science TA Levels 2007/08 to 2014/15

At key stage 4, three dependent variables were selected to represent the range of student attainment in science at GCSE from 2010/11 to 2015/16. These measures were:

- Achieved two good science GCSEs (C grade and above): Dichotomous variable coded Yes=1 or No= 0
- Entered Biology, Physics, Chemistry GCSEs and achieved equivalent of A*-B GCSE in Physics and Chemistry GCSEs: Dichotomous variable coded Yes=1 or No= 0
- Highest points score in science GCSE: range from 0 to 58 and for vocational range from qualification 0 to 55

The NPD data gives the highest grade achieved in any GCSE science as expressed as a point score, with 40 representing a C grade and 58 points an A*. The vocational qualifications are given intermediate point scores, with a Distinction at Level 2 given 55 points (see Table 11).

Table 11: GCSE and equivalent points scores

GCSE Grade	Point Score
A*	58
Level 2 Distinction	55
A	52
Level 2 Merit	49
B	46
Level 2 Pass	43
C	40
Level 1 Distinction	37
D	34
Level 1 Merit	31
E	28
Level 1 Pass	25
F	22
	19
G	16
Ungraded	0

The highest point score was useful because it is independent of the science course taken and captures the attainment in GCSE or vocational qualifications equivalent to one GCSE. The standard deviation in each case is relatively large compared to the mean, which may indicate that the mean is poor fit for the data. However, as Figure 5 shows, the grade distribution approximates a normal curve, although a significant number of students entered (434,384 12.5%) did not achieve a grade in any GCSE science qualification.
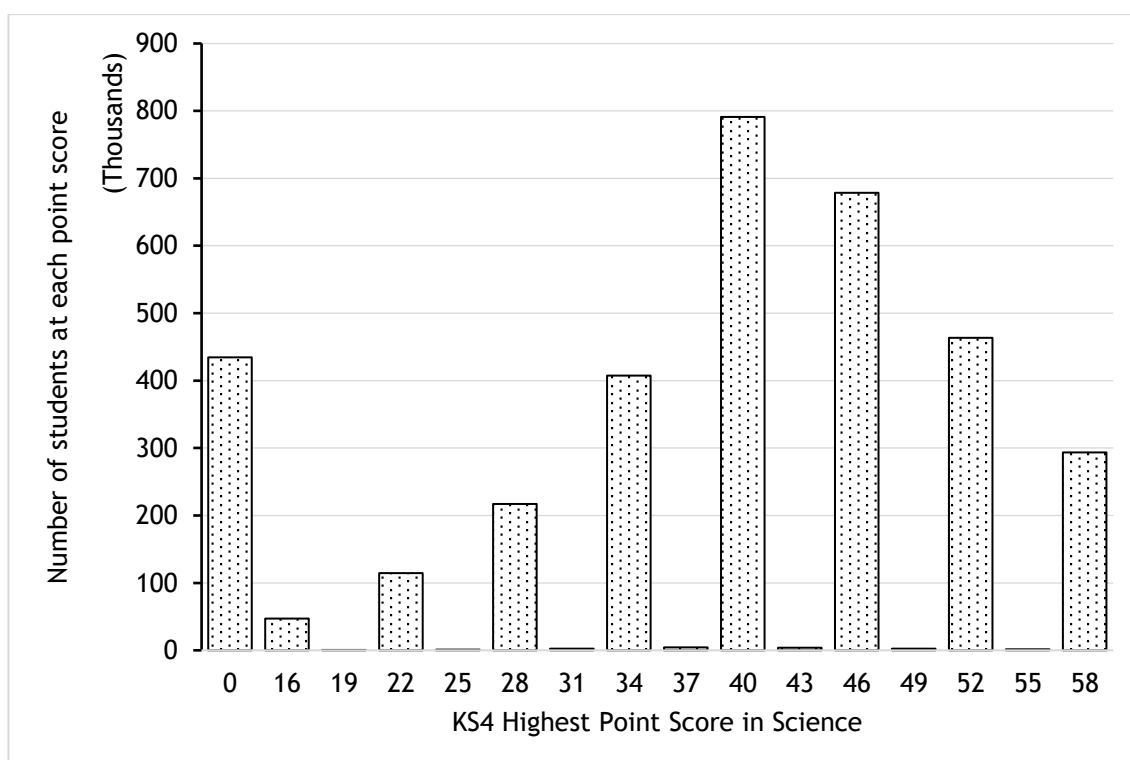
Figure 5 Grade distribution for the highest point GCSE point score in Science 2010 to 2016

6.3.3.3 Policy Changes to Assessment

The context chapters commented on the changes to education policy over the past ten years that have directly affected what is taught in science and how science is assessed particularly at KS2 and KS4. This section briefly revisits the changes and outlines the assessment regime in place for each year of data analysed in this study (see Appendix Table E, p325). For example, a student who completed end of KS2

examinations in English, maths and science in 2008 would have also been given Teacher Assessment levels at the end of KS2 and have possibly completed a linear Core Science GCSE in 2012 and the Additional Science GCSE in 2013 both of which had a controlled assessment component to examine practical science skills. Equally, this student may have chosen to follow a BTEC qualification, completed the Core Science GCSE over two years of key stage 4 or followed a triple science route.

In a later cohort, a student at the end of key stage 2 in 2014, is likely to have taken national curriculum tests in Reading, Spelling, Punctuation and Grammar and maths but may not have been selected as part of the sample sitting the key stage 2 science assessments. Instead the student would have been assessed by their teacher and a judgment made on whether they had made expected progress in science-based on national curriculum criteria. The student would not have received a national curriculum level at the end of KS3 as these were no longer used in 2016 and they would have been part of the first cohort to follow the new, more challenging GCSE science national curriculum. On this course, the assessment of practical skills is no longer carried out via coursework or controlled assessments but via the examination of the required practicals outlined in the specifications. The course is linear with terminal examinations after a minimum of two years and students' grades are numbered 9 to 1 as opposed to A* to C. The opportunity for the student to follow an alternative curriculum or vocational course has been limited due to the influence of the EBacc accountability measure. The outline of the changes and the examples explain some of the limitations to year-on-year comparisons.

The reform variables of interest (Table 12) were coded based on the dates (pre- and post-) when the reform came into effect. Two of the variables of interest related to the changes at KS4; the first is the change from modular to linear examinations (Baird et al., 2019) and the second the change in allowing GCSE equivalent

qualifications (Burgess and Thomson, 2019). The third variable of interest related to reforms at KS2 which saw the end of KS2 tests.

Table 12: Reform variables of interest, student populations and their mean attainment

	No. of Students	Mean Point score	Std. Dev.
Dependent Variables			
GCSE Highest point score	3,464,152	36.91	16.64
KS2 Science TA Level	3,306,319	4.23	.744
Independent Variables of Interest			
Modular exams pre-2012 (ref: Modular)	1,269,270	33.41	19.80
Linear exams post-2012	2,194,882	38.94	14.12
GCSE Equivalents allowed (ref: Equiv.)	1,912,409	33.22	19.78
GCSE Equivalents not allowed	1,551,743	41.46	9.94
KS2 Test Year (ref: KS2 TA pre-2009)	1,083,745	4.22	0.755
KS2 TA post-2009	2,222,574	4.23	0.739

Notes:
   a. Mean GCSE Highest point score equates to the sum of all the highest point score achieved by a student in any science GCSE between 2010-2016 divided by the total number of students.
   b. Mean KS2 TA Level equates to the sum of the science TA level achieved by a student between 2008 to 2015 divided by the total number of students.

6.3.4 Statistical Analysis

The analysis followed a four-stage process. Firstly, descriptive statistics were used to summarise the attainment data and present the percentage of students achieving the benchmarks of (a) Level 4+ at KS2, (b) 2 good GCSEs in science and (c) were entered for triple science and achieving an A or B in Chemistry or Physics (this variable in the NPD does not including Biology) at KS4 by student characteristic. The second part of the analysis explored the attainment gaps between students by calculating the mean science TA level at KS2 and the mean highest point score in science at KS4 for each of the different student characteristics selected. Thirdly, the mean GCSE highest

point score in science at KS4 and mean KS2 science TA level, and their standard deviations were calculated by student characteristic for each of the reform periods in the main variables of interest. The differences between the populations of students experiencing different assessment regimes and the effect sizes were also calculated to quantify the size of the difference over and above an indication of statistical significance (Coe, 2002). Cohen's d was used to calculate the effect size, where Cohen's $d = (M_2 - M_1) / SD_{pooled}$, where $SD_{pooled} = \int (SD_1^2 + SD_2^2) / 2)$ (Cohen, 1988, p. 44). M_1 represented the mean attainment pre-reform with standard deviation SD_1 and M_2 represented the mean attainment post-reform with standard deviation SD_2. Using Cohen's d, an effect size can lie between 0 to greater than 1, the guidelines for determining whether the effect is strong when are

0 – 0.20	= weak effect
0.21 - 0.50	= modest effect
0.51 – 1.00	= moderate effect
> 1.00	= strong effect (Cohen *et al.,* 2017)

Finally, ordinary least squares (OLS) regression was used in order to determine the association between reform and student attainment, controlled for student characteristics. Two baseline models were created one for KS2 and the other for KS4, these were coded for dummy variables (dichotomous Yes=1 or No= 0) representing each student characteristic as determinants of attainment. Six reform models were subsequently generated and compared to the baseline models for each key stage.

Key Stage 2

    Model 1: KS2 Baseline student characteristics model
    Model 2: KS2 Test Year variable
    Model 3: KS2 Baseline student characteristics model + KS2 Test Year variable

Key Stage 4

    Model 1: KS4 Baseline student characteristics model
    Model 2: Modular vs. Linear variable

Model 3: KS4 Baseline student characteristics model + Modular vs. Linear variable

Model 4: GCSE Equiv. vs. No Equiv.

Model 5: KS4 Baseline student characteristics model + GCSE Equiv. vs. No Equiv.

In each case the variables were entered to determine the difference in attainment of reform and the unstandardised coefficients are reported in the tables. In my research, the quantitative student attainment data is derived from the whole population and is therefore considered to be parametric in nature (Cohen *et al.,* 2017; Muijs, 2011). Comparisons made in the findings were tested for statistical significance to ensure that the differences are larger than might be expected due to sampling variation. The statistical significance of the differences between estimates is at the 0.05 level.

6.3.5 Limitations

Descriptive and inferential statistics were used to explore trends, patterns and relationships in the data that relate education policy changes to student outcomes, there are benefits, drawbacks and limitations to the various approaches.

Using the NPD gave access to the attainment and personal characteristics data for whole populations of students. However, for a number of reasons, it was not possible to include the entire population of KS2 or KS4 students in state education during time period explored in this study. For instance, students may have been excluded from school (Smith, 2008a), have been home educated or off-rolled prior to the examinations (Bradbury, 2019). As previously discussed, where data relating to key variables such as academic year or attainment level were missing or entered incorrectly a student was automatically excluded from the study. Therefore, the resulting sample, although very large and representative of the selected population

cannot claim to be the entire population. The population contained sub-populations of students who had experienced different reform regimes and although the raw differences between assessment routes could be affected by selection bias due to entry patterns (Baird et al., 2019) comparisons could be made between independent groups and statistical significance testing carried out appropriately (Rubin, 1985).

Applied to large samples, statistical significance testing can uncover rare associations (Khalilzadeh and Tasci, 2017) but statistical significance on its own should be used cautiously as the cut-off point is comparatively arbitrary and dependent on the sample size and coefficients (Cohen *et al., 2017*). The large sample sizes in this study, combined with small coefficients automatically generated statistically significant correlations at the 0.05 or 0.01 level hence effect size was used to give a standardised measure of the size of the differences between two groups. The different measures used to identify effect size have differing cut-off points which determine what constitutes a weak or strong effect (Khalilzadeh and Tasci, 2017), Cohen's d was used in the study, despite the relatively large difference in the standard deviations between the GCSE Equiv. and No GCSE Equiv. groups (Coe, 2002).

Both the KS4 highest point score and KS2 science TA levels are ordinal variables and not interval variables as required for regression analysis, however, the method is robust enough to withstand bending the assumptions for the statistic (Field, 2000). However, statistical analysis using any teacher assessment data, should take in account possible inconsistencies across teachers and whilst moderation is said to take place, there may be discrepancies between TA and external examination attainment levels (Perry, 2016); which may impact on the reliability of the student data and subsequent analysis.

6.4 Data Collection: Teacher Practice

Both qualitative and quantitative data were collected from the teachers in this study; using a pre-interview questionnaire and one-to-one, semi-structured interviews with participants in their settings. This approach quantified and explored teacher responses to build an understanding of how their teaching practice may have changed as a result of changes in government policy. This approach to data collection opens up the possibility of generating knowledge about what happens in classrooms, how common the practices are and how this links to the biographical profile of the teachers.

This section begins with a description of the population sample and an explanation of the reasoning for the sampling strategy and recruitment decisions. Next, the development of the pre-interview questionnaire and related interview schedule is described which indicates the degree of structure built into the research study. I then explain how the data was collected and brought together before detailing the methods of analysis used to interpret the data.

6.4.1 National and Regional Teacher Population

In 2016, there were 503,900 contracted teachers in state-funded schools in England, 222,400 teachers in nursery/primary phase and 208,200 in secondary, of which, 36,600 were teachers of science, the remainder in special or alternative provision institutions. Of the entire population of teachers 25% were male and 75% female, with the majority of teachers between the ages of 30 and 50 (see Appendix Table C and Appendix Table D, p324). Generally, all primary school teachers teach most curriculum subjects including science during their teaching week while in secondary schools, science is generally taught by a specialist teacher of science who may teach across the three different key stages in that phase. The total population of classroom

teachers in the South East region was 69,848 as of November 2016 (National

Statistics, 2017) and the composition was similar to the national statistics where 25%

of teachers in the South East were male and approximately 19% aged over 50. As the

most recently available reporting of teacher numbers, this represented the

population from which I drew my sample, therefore it was likely that the profile of

my sample was representative of the teaching population as a whole (De Vaus, 2014).

6.4.2 The Sample and Sampling Strategy

The unit of study for this part of the research was the individual secondary science

teacher or primary science coordinator and was drawn from the population of

classroom teachers in the South East of England. This relatively homogeneous group

share commonality in that teaching science plays a significant role in their practice.

These teachers represented and symbolised key elements for the research such that

valid data can be generated through them (Ritchie et al., 2013). However, within the

sample there were points of difference that brought diversity to the research and

ensured that the full range of factors associated with the research focus were

explored.

As this target population was located within different organisational units (schools) a

multi-stage sampling design was used (Ritchie et al., 2013). By reviewing three local

authority websites, the School sample frame (De Vaus, 2014) was developed

generating a list of primary and secondary schools in neighbouring counties. This

comprised of 193 primary schools and 75 secondary schools representing the

population of schools across the location and arranged in order of the school's

relative proximity to the university. All schools in the sample frame were invited to

take part. However, due to limited researcher resources, the schools closest to the

university, with the shortest travel time, were placed at the top of the list and contacted first.

From the range of available sampling strategies, I elected to use a non-probability, convenience sampling strategy at the start of the research (Cohen *et al.,* 2017; Daniel, 2011). This method is commonly used in small-scale research and is not intended to be statistically representative as would be the case with a random probability sample (Ritchie et al., 2013). The sample represents only itself, and as such the findings from the research cannot be used to make generalisations about the wider teacher population. However, it may be a case that the findings are generalised on the assumption of the "cultural consistency" (Schreier, 2018, p. 86) shown in education settings, for instance with teachers in a neighbouring local authority in England. Non-probability sampling was easy to set up and had little or no monetary cost (Cohen *et al.,* 2017). Cases were selected according to availability, that while convenient, meant that there was a lack of known probabilities for the inclusion of different cases in the sample. A randomised selection process based on the inclusion probabilities would have generated a more representative, less arbitrary, sample (Vehovar *et al.,* 2016). In non-probability sampling it is difficult to target specific elements of the population, giving an over representation of members of the population who are most accessible and an underestimation of variability (Daniel, 2011). However, non-probability sampling does allow a wide range of different recruitment methods to be employed (Daniel, 2011) which when applied systematically can approximate the reach of probability sampling (Vehovar *et al.,* 2016).

During the fieldwork, snowball sampling (Cohen *et al.,* 2017) was also used, to increase the number of teachers recruited. After interviewing a participant, my details were subsequently passed on to their colleagues. Potential participants were

primed to expect an approach from me, which made it easier to secure further interviews, particularly with primary science coordinators. The recruitment process was carried out concurrently alongside the fieldwork interviews until a sufficient sample size was obtained (n=26) and data saturation was achieved (O'Reilly and Parker, 2013); this number falls within the conventions for exploratory research of between 20 to 150 participants (Daniel, 2011). The sample is large enough to generate "thick descriptions" and with this homogeneous population, a smaller sample size is acceptable (Cohen *et al.,* 2017). Furthermore, there is a law of diminishing returns where increasing the sample size no longer contributes new knowledge to the evidence (Ritchie et al., 2013). In this qualitative exploratory research, where statistical analysis and generalisability is not the final outcome the "rules" which apply to quantitative research regarding sample size do not strictly apply (De Vaus, 2014).

6.4.3 Recruitment

The recruitment letter along with the supporting consent form and participant information documents (Appendix A3, pp330-333) were trialled with peers and teacher educators before dissemination. The redrafted documents were subsequently emailed to schools in the sample frame and the responses were received, collated and managed electronically. Although this made it easier to monitor the recruitment process, it was not always possible to ensure that the recruitment email arrived in the headteacher's inbox directly. I also encountered similar problems to Morrison (2006) in recruiting participants, where the headteacher had given consent but had not consulted the staff. Updating the sample frame throughout the research ensured that those schools and teachers who provided a definitive no response, were no longer contacted.

The data collection period ran from October 2017 to April 2018 to avoid the examination period in the spring and summer terms. Once recruited and consent forms received, the science teachers were then sent an electronic link to the pre-interview questionnaire. The next section of the methodology chapter describes how the teacher questionnaire and interview schedules were developed and how they will be integrated and analysed.

6.4.4 Data Collection Tools

As mentioned in the introduction to this chapter, the qualitative data collection in this study was structured around a teacher self-administered questionnaire, followed by a semi-structured face-to-face interview. These methods created boundaries around this phase of the study, narrowing its focus on the description and exploration of the aspects of teacher practice affected by education reform. Specifying in advance the topics to be investigated may have limited the exploration of unanticipated issues, however, it served to keep the main objectives of the interview in focus (Ritchie et al., 2013). Using a mixed methods approach to collect data from one source to study some aspect of human behaviour avoids the reliance on one method which may bias the outcomes (Cohen *et al.,* 2017). Therefore, the data from the interview was used to elaborate, enhance and illustrate the results from the questionnaire (Greene *et al.,* 1989).

As a data collection tool, questionnaires are one of the most familiar and frequently used methods (Cohen *et al.,* 2017). Self-administered questionnaires need to be clear and simple and their design needs to capture the essence of the concepts and theory being researched (De Vaus, 2014). Serving as an interview schedule, my questionnaire (Appendix A3, p338), laid out, *a priori,* the topics for discussion. This gave the participants time to reflect, recall and engage with the elements of their

practice in private and without researcher intrusion. But also meant that teachers might not have given a naturalistic and genuine response in the subsequent interview. Pre-knowledge of the interview content could have introduced a form of confirmation bias or limited teacher's responses to those that they believed represented appropriate or mandated teacher behaviour (Cohen *et al.,* 2017).

From the range of different types of questionnaire: structured, semi-structured and unstructured, the structured format was adopted. With a narrow focus on the research themes, this type of questionnaire limited responses but constructed appropriately required little or no clarification from the researcher. The advantages of using an on-line electronic questionnaire were that costs, without the need for postage, were negligible and chance of losing data through participants skipping questions was reduced (Cohen *et al.,* 2017). Participants were able to complete the questionnaire through any suitable electronic device at a convenient time and once submitted the data was received immediately, easily recorded and stored. Though the respondent's names were visible to me, data security and confidentiality were preserved as access to the system remained password protected. Disadvantages of conducting questionnaires electronically are that participants may become frustrated and opt-out entirely. To encourage responses, the electronic link was personalised to each respondent and reminders were issued prior to the interview.

To address the topics of the research and to gather data on behaviour, knowledge, attitudes and attributes different types of question item were used (De Vaus, 2014). Two sections of the questionnaire contained items taken from international surveys on teaching practice (IEA, 2014; OECD, 2012). This provided an analytical basis for the evidence collected which had a high level of validity. Firstly, the attribute questions collected information on participant characteristics using closed multiple-

choice, categorical items for interval and numerical data input. The second section concerned classroom practice and pedagogy. These questions used a 6-point Likert scale ranging from *Every or Almost every lesson* to *N/A*. I added a second question, which was not part of the original scale, that asked teachers to indicate if their practice had changed in the last 2 years. The third section of the questionnaire asked about teacher professional development and collaborative working and used closed questions along with a 5-point Likert scale ranging from *Weekly* to *N/A*. Section four explored teacher understanding of the new curriculum, progress and accountability measures and their confidence in using them. This item used a sliding scale from 0 to 100, where 0 represents zero confidence and 100 is maximum, fully confident. This allowed for a wide variation in response. Section 5 explored teacher's attitudes to the reforms as opposed to their use of the reforms, here, a 6-point Likert scale is again used (see Table 13).

Table 13: Questionnaire item and response type

Question Type	Data Type	Response Type
Attribute	Biographical e.g. Age range, gender	Closed multiple-choice and an ordinal or numerical data entry
Behaviour	Record of teacher classroom practices.	A 6-point Likert rating scale.
	Record of teacher professional development and collaborative working.	Closed multiple-choice and 5-point Likert rating scale.
Knowledge	Measure of the confidence of the new reform	Ratio data on a scale out of 100
Attitudes	Measure of the impact of the new reform	A 6-point Likert rating scale.

Rating scales can be easy for participants to respond to and can be used to present a range of linked constructs which do not constrain how each is answered (De Vaus, 2014). Rating scales also required a single response to an item and although there is

no assumption of equal intervals, these rating scales allowed for the ordering of responses from high to low in analysis. However, ordinal data from the Likert scales, does not indicate if the respondents are telling the truth or if there should be other categories or items included (Cohen *et al.,* 2017). For this reason, a number of the questions also contained free text, "Not Applicable" and "Other" options to ensure that respondents were given a sufficient range of alternatives and not forced to give a false choice (De Vaus, 2014).

The questionnaire was piloted with practicing secondary and primary teachers and subsequently, amendments were made to particular questions. Changes improved wording, reduced ambiguity and opt-out statements (De Vaus, 2014).   The overall questionnaire totalled 23 items with a projected 20 minutes completion time. Instructions for completing the questionnaire were incorporated in the email inviting teachers to take part in the study and were fully accessible through the online software, increasing the chances of a high response rate.

The collection of participant's' biographical information allowed a range of different analysis to be undertaken so that differences between teachers working in different key stages or with different levels of experience were illustrated. The questionnaire began with non-threatening questions that were easy to respond to, followed by the more behavioural closed questions and finally the personalised attitude questions. This layout helped to maintain interest, assisted with the flow and enabled participants to feel comfortable and build confidence as they progressed (Cohen *et al.,* 2017).

6.4.4.1 Development of Questionnaire Items

The following explains how the questionnaire was developed and discusses the validity of particular items to measure the constructs under investigation. The review of the literature on classroom practice indicated that several international surveys exploring this field had already been developed (IEA, 2014; OECD, 2012). TIMSS (Trends in International Mathematics and Science Study) is conducted every four years and began surveying students and their teachers in 1995. Although TIMSS distributes its survey to teachers, the unit of analysis is the students in the sample groups taught by these teachers. TIMSS collates information on teacher's biographical characteristics, experience, professional development, and the exploration of classroom practices within the context of scientific inquiry and conducting investigations (Mullis and Martin, 2013). TALIS (OECD Teaching and Learning International Survey) differed in that it focused on the teacher and their professional activities, and how their classroom teaching practice and their participation in professional learning communities contributed to improving the learning conditions of students (Vieluf *et al.*, 2012).

Despite the complexity in investigating the interrelation between classroom practice and student learning (OECD, 2013a; Dillon and Manning, 2010) international survey research questionnaires aim to capture the essence of teacher classroom practice through a limited number of survey instruments. The more that is known about what makes good teacher practice, the better the ability to design an instrument that has high content validity. Here, content validity refers to extent to which the items on a questionnaire appropriately capture the concept they are designed to measure (Muijs, 2011). From the review of the literature (Muijs and Reynolds, 2017; Furtak *et al.*, 2012; Hattie, 2009) I identified seven significant domains in classroom practice (Appendix Table F, p336). These were used to develop a simple framework which

sketched out the aspects of classroom practice that warranted exploration namely: Whole class teaching; Differentiation; Rote Learning; Extending Learning; Formative Assessment; Summative Assessment and Practical Work. Borrowing from both the TIMSS and TALIS teacher's questionnaires, the teaching and learning actions were mapped onto each domain, where the action represented something that was carried out either by the teacher or by the student.

On reflection, I concluded that the planned questionnaire was too lengthy and its validity was difficult to verify in this small-scale study. The questionnaire schedule was revised, this reduced the number of items and shortened the overall length of the questionnaire to limit response burden. The resultant question items included the full use of the teacher practice questions (item 42- *"How often do each of the following activities happen in this class?")* from the TALIS teacher questionnaire (OECD, 2012, p. 22). Table 14 demonstrates how the 13 items have been grouped into three dimensions representing a simplified model of the constructs related to the teaching practices: Structuring, Student Orientation and Enhanced Activities, which encapsulate a measure teacher practice (Vieluf *et al.,* 2012, p. 52).

Crucially, for this research, adhering to the items used in the TALIS teacher questionnaire meant that the prerequisite validation processes had already been carried out. The question items have undergone large-scale, statistical factor analysis and reliability testing across different national contexts in a manner not possible on a small-scale study such as this and therefore no further testing of the item content validity was required. From the OECD analysis each dimension achieves relatively high in Cronbach's Alpha score of reliability; structuring with an $\alpha = 0.73$, student orientation, $\alpha = 0.70$, and enhanced activities achieving $\alpha = 0.72$ (Vieluf *et al.,* 2012, p. 54).

Table 14: Teaching Practice Dimensions and Questionnaire Items

Pedagogy	Teaching and Learning Actions
Structuring	• I explicitly state learning goals • I review with the students the homework they have prepared. • At the beginning of the lesson, I present a short summary of the previous lesson. • I check my students' exercise books. • I check, by asking questions, whether or not the subject matter has been understood.
Student Orientation	• Students work in small groups to come up with a joint solution to a problem or task. • I give different work to the students that have difficulties learning and/or to those who can advance faster. • I ask my students to suggest or to help plan classroom activities or topics. • Students work in groups based upon their abilities.
Enhanced Activities	• Students work on projects that require at least one week to complete. • Students make a product that will be used by someone else. • I ask my students to write an essay in which they are expected to explain their thinking or reasoning at some length. • Students hold a debate and argue for a particular point of view which may not be their own.

The questions were therefore appropriate to use in this study as they represented a valid method to measure what they were designed to measure and would generate reliable data on teacher classroom practice. The TALIS questionnaire, however, measured the frequency of teacher activity on a response scale that included measure such as *"In about one-quarter of lessons", "In about one-half of lessons"* and *"In about three-quarters of lessons".* Having trialled my questionnaire, these

measures did not express an appropriate meaning. The final response scale used in this study was edited to read *"often", "sometimes"* and *"rarely"*.

The starting point for the development of the question items on homework and teacher professional development was the TIMSS survey of 2007 (IEA, 2007). Six questionnaire items were selected and incorporated into my research instrument. Four items *("Do you assign science homework?", "How often do you usually assign science homework?", "How often do you assign the following kinds of science homework?"* and *"How often do you do the following with the science homework assignments")* provided the scales for exploring homework. Two other questions *("How often do you have the following types of interactions with other teachers?"* and *"In the past two years, have you participated in professional development in any of the following?")* were used as the foundation to explore teachers access to and use of professional development and collaborative working. These measurement scales had also undergone a range of testing and development to ensure their reliability and validity (Martin *et al.,* 2015; Olson *et al.,* 2008). However, to explore teacher behaviour around collaboration and professional development in the context of this study, it was necessary to incorporate additional aspects into the questionnaire items. The TIMSS teacher questionnaire of 2015 (IEA, 2014) gave a wider range of responses to the question *"how often do you have the following interactions with other teachers?"* than the 2007 edition, increasing the number of options from four to seven. This incorporated asking teachers about how often they *"Work together to try out new ideas"* and *"Work as a group on implementing the curriculum"*. One further option asking teachers to indicate whether they carried out joint marking and assessment was added. This was intended to capture data on whether teachers continued to corroborate their marking and moderation of students work. The Cronbach's Alpha coefficient for

England stated in the original 2011 analysis was α = 0.82 (Martin *et al.,* 2011), this is considered high enough to indicate that the question measures constructs of *"collaborative working to improve teaching"* well. I surmised that adding my additional response item would not alter the validity measures in any significant way as the question was sufficiently robust and the methodology had been used consistently in the TIMSS survey structure for some years (Martin *et al.,* 2015; IEA, 2011).

Finally, the last two questions list the key curriculum, assessment, progress and accountability measures featured in the research. It was decided not to develop a separate questionnaire for secondary and primary colleagues in order to highlight the potential similarities and differences in their engagement with these reforms. By focussing first on a confidence measure in the scale of 0 to 100, the item drew out how well informed the participants were in the prevailing education reforms overall, regardless of the key stage they taught in. The option of a question item not being applicable could generate useful data in itself. Similarly, the role of the final attitudinal question was two-fold. Firstly, as a means for teachers to reflect on their teaching practice and draw together elements of the earlier question items to consider these in the context of reform and its impact. Secondly, it acted as an engaging prompt during the interview enabling deeper discussion with the teachers and the impact that education reform had on their teaching and students.

6.4.4.2 Interview Design

> "If you want to know how people understand their world and their lives, why not talk to them?" (Kvale, 2007, p. 1)

Conducting interviews with teachers in this research provided the substantive qualitative data upon which the findings and conclusions were based. As a key research method, qualitative interviewing used conversation with human actors as

the means to generate knowledge of the lived experience. As a tool for data collection, the face-to-face interview gives flexibility in the capture of verbal and non-verbal cues inherent in a conversation (Cohen *et al.,* 2017). Generating data through participant interview can be undertaken through alternative routes; the post-modern, constructionist or decolonising conceptions (Roulston, 2010), or in-depth interviews in phenomenological or oral history research (Guest *et al.,* 2012). The literature (Cohen *et al.,* 2017; Borer and Fontana, 2012; Kvale, 2007) outlines the practicality of conducting interviews and theorises its associated epistemological facets.

At one end of the spectrum are structured interviews. These are closed in style, employ a structure with established a priori categories and pre-established questions, the answers to which can be categorised, codified and to some extent generalised (Borer and Fontana, 2012). Through this method, the control and purpose of the interview conversation was determined by me, the researcher. Understandably, a certain level of control must be held if the research is to meet its aims (Kvale, 2007). However, Cohen et al. (2017) argue that this rigour dehumanises the interviewee making their responses impersonal and mechanistic, limiting the meaning and nuance that can be derived from the data. The closed interview style is also said to limit the interactions and responses available to the interviewer, who is unable to react, prompt and explore interesting and unexpected avenues of inquiry. At the other end of the spectrum, interviewing using an open, informal conversation style, where there is no pre-determined topic schedule, provides naturalised, rich and wide-ranging data. During the course of the "conversation" salient points can be raised and the interviewer can build on and develop the interviewee's comments (Cohen *et al.,* 2017). This type of interview stems from an epistemologically subjective interpretative perspective, where the need is to delve into the complexity

of the situation to capture the unique responses of individuals on a factual level or on the level of meaning and feelings (Kvale, 2007). Importantly, however, with open-ended-interviews, the outcomes may be different for different participants making a systematic analysis more difficult. In addition, key topics of interest might not arise and as such the data might not be useful in addressing the research questions (Cohen *et al.,* 2017).

Taking this into account, I decided to conduct semi-structured interviews (Cohen *et al.,* 2017; Morse, 2012) in synergy with the critical realist standpoint of the overall research. This is manifest as a real world in which teachers as social actors exists, but that our knowing of it is subjective (Maxwell, 2018). It is possible then that the inclusion of open-ended probing questions (Legard *et al.,* 2003) incorporated with the systematic closed questions of a semi-structured interview will provide data which can be used to explore causal mechanisms, context and outcomes related to the individuals. In a sense, this interview strategy leans heavily on the method of Problem Centred Interviewing (Witzel and Reiter, 2012) as the research has a clear focus and is geared towards a socially relevant problem to which the participants can articulate their practical knowledge.

A starting base of a clearly defined topic guide places the interviewer in a neutral role and helps to minimise bias (Roulston, 2010). However, it was important to acknowledge the possibility of my subjectivities and beliefs unintentionally creeping into the interview and therefore, I aimed to refrain from participating in the data generation other than through posing the questions. All the participants were asked the same questions to ensure comparability and ease of analysis, this also ensured that the interview and data collected covered all of the topic areas sufficiently to address the research questions (Legard et al., 2003). Furthermore, by leaving room

to probe and respond to interviewees naturalistically, the standardised semi-structured interview let me collect rich qualitative data over and above the key topics of interest. Criticism of this type of approach to interviewing and interview data suggests that research participants do not necessarily do what they say they do or are not always being truthful. These critiques can be levelled at other epistemological perspectives (Cohen *et al.,* 2017; Borer and Fontana, 2012). Notwithstanding, the semi-structured interview approach is often seen in mixed methods studies as a means through which quality data can be generated which produces valid findings (Roulston, 2010).

6.4.4.3 Preparing and Conducting the Interviews

Having completed the consent form, the interviews were arranged and took place at a time suggested by the participants and in a safe, convenient location. Creating a rapport between me as the researcher and the teacher as interviewee involved building and maintaining a sense of trust and mutual understanding (Roulston and Choi, 2017; Legard *et al.,* 2003). From the outset and throughout the interview, participants were reassured about confidentiality and reminded of their right to withdraw consent. A digital device was used to record the interviews, which were then saved and renamed with an assigned teacher alias. Experience from the early interviews led to adjustments in the practical conduct of subsequent interviews including slowing the speed of questioning to give interviewees more thinking time and being sensitive to interviewee tone of voice and using this as opportunities for deeper exploration (Roulston and Choi, 2017; Legard *et al.,* 2003). Following the interviews, participants were promised a copy of the executive summary of the research and thanked in recognition of their generosity in giving up their often-precious non-contact time.

6.4.6 Processing and Analysing the Data

6.4.6.1 Processing the Questionnaire Data

The questionnaire was sent to participants via a secure link to Qualtrics, an external, secure site provided by the university. Once completed, the raw data was stored electronically as a pdf document and the individual responses exported, printed and used as the basis for the interview schedule with each teacher. With a sample of this size, it was important to ensure that no teacher could be identified from their data, therefore, it was coded using an alias in place of teachers' real names. The same teacher alias was used to tag the interview recordings and consequent transcripts. The codes used referred to the participant's school phase, gender, key stage and order in which they were interviewed, for example, the alias for my third interview with a male biology teacher in secondary school was smb3, whereas pfks25, represented the fifth interview with a female teacher working in key stage 2. Using an electronic, computer-based questionnaire as opposed to a paper-based one enabled faster data processing as all information provided by participants was collated centrally; frequency counts, cross-tabulations and visualisations of the data were very quickly produced.

Due to the small sample size, it was not possible to carry out an in-depth quantitative analysis involving the use inferential statistics. The descriptive statistics, means and frequencies were, however, useful in determining the commonality of certain teacher practices across year groups. The outcomes from the questionnaire played a role in investigating how wide-spread the changes to teaching practices were among the sample and as such complemented the knowledge revealed through the interview.

6.4.6.2 Processing the Interview Data

The teacher interviews generated approximately 23 hours of audio recordings which were transcribed and converted from audio to text. The abstraction process transforms the interview "conversation" from that of social interaction between two people removing the associated gestures and oral language signifiers, losing the tone of voice and intonation giving rise to a "decontextualized" written account (Kvale, 2007). Decisions on whether to produce verbatim accounts were embedded in the theoretical assumptions of the research; this dictated what was constituted as a legitimate form of knowledge to reflect the social reality of the interviewee (Poland, 2003).

The interpretation of a transcribed interview can be altered through any number of researcher analytical decisions such as where to insert full stops and commas or how to interpret a misheard word. Making judgements about when sentences start and end, the use of idioms, inclusion of pauses and hesitations all need to be made explicit by the researcher at the outset of the transcription process (Poland, 2003). Frontloading the decisions on how the research conceptualised the interview data and how this data would be analysed and reported, reduced any transcription dilemmas, narrowed the possible options and brought greater consistency to the process. The analysis of the interviews did not rely on linguistic style, social interactionism or conversation analysis therefore, there was no requirement for in-depth detailed transcription notes (Kvale, 2007). However, to limit the risk of losing nuanced data through reductive summarising of the audio recording, the interviews were transcribed verbatim. For expediency, I used voice recognition software which converted the audio recording into text as I "re-said" the participant's words whilst re-listening to the recordings. This helped to re-imagine the interview and rebuild the picture, tone and meaning of the participant's responses (Brooks, 2010).

During the transcription, I decided to correct spoken grammatical errors in the written text and to omit muffled words rather than guess at what the interviewee had said. This subjectivity was part of my position as a researcher, recognising that my role in interpreting the data and ensuring accurate transcription was not value-neutral (Allen, 2017). The pauses between the speaker's utterances and run on sentences were indicated by an ellipsis (…) as illustrated by the extract below.

> sfb9: Then I feel negative about that, definitely… maths and English have got 10 lessons a fortnight and science has got 9… if that takes into account that they are all doing double then definitely… it's awful…
>
> I: Overall, with all the things that have been going is there anything in particular that you can say… about changes to your teaching?
>
> sfb9: … certainly compulsory double award has made a big difference, it's made teaching bottom set year 11 soul destroying… that's for me, I thing that's the biggest thing.

Once completed, the transcribed interview texts were uploaded to NVivo (version 11), a software programme used to support qualitative data analysis; each transcript was saved using the ascribed alias.

6.4.6.3 Analysing the Interview Data

As stated above, analysis and interpretation of the interview data, began during the transcription phase, in fact, keywords and themes had already emerged from the first two conversations. However, I was also anxious to avoid shutting down openness to the emergence of alternative or contradictory issues. While the semi-structured interview was guided by the pre-interview questionnaire, the data derived from the conversations added meaning to the quantitative responses already collected. This interpretative phase was a key part of the research because it asked "…what does this mean…? (Willig, 2014, p. 137). Although the analysis began with the data handling and management through the transcription; bringing order to the

data to summarise and look for patterns was an on-going analytical process
(Bathmaker, 2010).

The diversity of research designs and qualitative interview strategies invite a
corresponding diversity in strategies for conducting qualitative data analysis; framing
these are the researcher's theoretical assumptions and representational strategies.
Different approaches, for example, hermeneutics, ethnographic, and narrative
methods influence how interview data is analysed (Roulston, 2014). Generally, the
type of analysis and interpretation carried out depends largely on the researcher's
ontological and epistemology positions (Willig, 2014; Braun and Clarke, 2006) and
involves the use of theory and theorising (Bathmaker, 2010). Each analytical
approach moves from description to explanation and possible theory generation
(Cohen *et al.,* 2017) and at each turn research subjectivity is inescapable.

Research analysed deductively begins with a certain system of rules and decides if
the phenomena observed obeys that rule. Working inductively moves from the
specific observation of a particular phenomenon towards developing a theory.
Alternatively, abduction takes its starting point from the empirical data. The
analysis goes on to explore the data to uncover new knowledge, this new knowledge
is then tested in different contexts (Reichertz, 2014). Whereas retroduction involves
making inferences from a description of a problem or phenomena, leading to an
understanding of the casual properties producing it (Sayer, 2010). Each strategy has
its strength and limitations (Cohen *et al.,* 2017; Braun and Clarke, 2006). In reality, I
sense that all of these approaches to data analysis occur at some point before, during
and after the research process, what is important is that as the researcher, I make
explicit how and why a particular path was taken.

The analysis of the qualitative data was conducted using a combination of thematic analysis and framework analysis. This maintained an open dialogue between me and the data through the search for meaning and alternative explanations. Thematic analysis can be flexibly applied within any theoretical position and works as a method to reflect reality and unpick the surface of reality (Braun and Clarke, 2006). Although, critique of this method suggests that it is merely the first step in analysis before the real decisions about representation and interpretation are made (Willig, 2014), thematic analysis avoids the pitfalls of content analysis, where meaning and context are lost through extensive reduction and codification of the data (Vaismoradi et al., 2013).

A theme "captures something important about the data in relation to the research and represents some level of patterned response or meaning" (Braun and Clarke, 2006, p. 82). Themes can be abstract concepts formulated through the words, expressions and images that the interviewee reveals (Ryan and Bernard, 2003). During the analysis categories and themes were identified that not only described the issues but looked behind the text to uncover the latent meaning (Vaismoradi *et al.,* 2013). The judgement about what constituted a theme was made after repeated reading of the entire interview data set and was a continuous, iterative process based on several factors. These included, representation within the theoretical framework, relevance to the research questions, prevalence within each data item and the entire data set, and the expressions and metaphors used by the interviewees (Ryan and Bernard, 2003) .

In my research, the data collection, transcribing and coding of each data interview took place concurrently throughout the fieldwork. Immersion in the data began whilst transcribing, with the stop-start, back and forth replay of the interviews. This

generated initial ideas for the structure of the coding, themes and categories. On first reading of the early interviews frequently used words and phrases were coded as keywords and labelled as nodes in NVivo, the entire interview data set was coded with this initial set, with similar key words grouped into categories. During the remaining data collection process new codes and themes emerged, these gave unexpected insights. Subsequently, once all interviews were completed, each transcript was re-read and recoded with the additional themes to ensure that, as much as possible, all of the features had been captured from the data. The semantic nodes (Table 15) were categorised to bring together data that reflected what teachers did in the classroom, their teaching and learning activities, whole school issues and specifics about the science curriculum.

Table 15: Data analysis: Descriptive Semantic Codes

Key nodes	Key Categories
Rigour	Whole school
Target grades	Teaching and Learning Activities
Practicals	Curriculum
Teaching hours	Changes
Time Pressures	KS3
Fun	KS4
Flight path	KS5
	Primary
	CPD

Grouping by key stage was included as a node to help investigate areas where teaching strategies had changed for different students. The search tool option Matrix Intersection was used to construct tables of coding across groups to facilitate this (Woolley, 2009).

The concept-driven coding stage sought out data that explored the emergent latent themes and the *a priori* themes derived from the theoretical framework. The latent

themes included those related to ideas around fairness, the nature of science for all
and the teacher's sense of professionalism. The *a priori* themes focussed on
identifying where teachers had expressed their thoughts on the reforms relative to
their past experience, future projections and present challenges to their decision-
making (see Table 16).

Table 16: Data analysis: Latent Codes

Actions/interactions	Concept – Agency
Decision making within rules - Discretion	Past: Iterative
	Future: Projective
Skills & Experience - Professionalism	Present: Practical-evaluative
Sense of Fairness - Justice	Beliefs
Suitability for students – Justice	Self-Efficacy

The latent themes were interpreted through the theory and literature to examine the
participant's conceptualisations of the current context, this brought an
understanding of how teachers used their agency and discretion in difficult
circumstances.

6.5 Ethics

Ethics refers to questions of values, a question of duty, of procedure, virtues or
consequences, (Cohen *et al.,* 2017) with the core values of respect, justice,
beneficence, non-maleficence, fidelity and academic freedom (Julnes, 2018;
Farrimond, 2017) being built into all stages of the research process. The Ethical
Guidelines for Educational Research (BERA, 2018) reminds all researchers of the
overall need to protect the integrity and reputation of the field by adhering to the
highest standards of academic integrity. Ethical issues arising in this context can be
found in the analysis of both quantitative and qualitative data. Areas of misconduct
involve the fraudulent practices of the suppression or falsification of findings

(Creswell, 2018). Farrimond (2017) also warned against "cherry-picking" data, deliberate or accidental manipulation to meet some predetermined conclusion.

The application for ethical approval from the university Cross-Schools Research Ethics Committee (C-REC), was considered to be low-risk and the application was approved without revision. Subsequently, all of the research was conducted in accordance with C-REC and BERA guidelines. Prior to commencing fieldwork further permissions were required to obtain the secondary data required for the quantitative phase of the study.

6.5.1 NPD

Working with secondary data such as the NPD involved additional ethical and security considerations, and throughout, I was bound by the Data Protection Act 1998. The holder of this annually released data is the Department of Education, who make it clear to children and families what information is held on them, why, and what uses the data is put to (DfE, 2015c). The law allows this information on pupils to be shared by the DfE with third parties; which is accepted as informed consent. Anonymity was guaranteed through unique pupil numbers that can be matched across years but does not enable the identification of an individual child. Following the completion of a data request form, confirming the appropriate credentials and the use of the information, the data arrived in an encrypted zip file. Access to the raw data is limited to the researchers named on the original application form and can be held for a period of three years. As an institution, the university is registered with the Information Commissioner's Office and as such, I am authorised to process the data using statistical software. Secure disposal of the data must take place on completion of the research (DfE, 2015c).

It was important to maintain the strict guidelines set out for ethical research across every aspect of my study, not just as a researcher but with respect to my positionality. This applied when working with the sensitive quantitative data linked to student characteristics or when visiting schools, working with participants from a relatively small catchment area.

6.5.2 Participants

In working with teachers within their school settings, I was guided by the University of Sussex and the British Educational Research Association Ethical Guidelines (BERA, 2018). I had a responsibility to the interviewees who gave up their time and throughout, their confidentiality and anonymity was assured. Continuous informed consent was achieved at all times during the data collection process. All participants were treated fairly, severing their needs first.

Access to schools can be an issue (Farrimond, 2017), and gaining the trust of the headteacher (Harvey, 2011), letting them know who I am and the nature of my research, was a prerequisite to gaining permission and approval prior to data collection. Confirming how the outcomes of the teacher interviews would be shared was important in maintaining a level of trust between myself and all those involved, particularly where teaching staff responded to questions with sensitive information or gave responses that led to emotional distress. Cohen et al (2017) referred to this as non-maleficence, the intent not to harm the participants in any way; physically emotionally, professionally or psychologically. To empower the participants, I outlined at all times their right to withdraw from the study without obligation. Empathising with the situation some practitioners found themselves in, I strived to investigate their experiences through building a rapport, without prejudice or bias.

As the gatekeeper, allowing access to staff, it was important to first engage the headteacher to gain their consent for the research to take place in their school (Cohen *et al.,* 2017). An invitation letter, outlining the purpose of the research and level of involvement required, was sent to schools. Even though the headteacher did not take part in the interviews making the invitation personal avoided alienating these important stakeholders. Informed consent was obtained from each headteacher and all teacher participants. At the end of the interviews, I debriefed the participants and confirmed how their contribution would be analysed and used as part of the wider research document. At all times I remained aware of the power dynamic within the relationship between me as the researcher and them as the researched. To protect their privacy and allow participants to speak freely, their anonymity and confidentiality of the data were assured at the outset (Farrimond, 2017).

An alias was used to disguise the identity of the teachers and references to their school name, locality or federations were also removed from interview transcripts and completed questionnaires. This also ensured that information shared with me was less likely to harm others through inadvertent disclosure of personal information. Furthermore, often the most insightful and revealing comments would come at the end of the interview, at those unguarded moments when the participant was at their most relaxed. Preserving the integrity and authenticity of their voice required ethical decisions about participant representation throughout the analysis process (Creswell, 2018; Cohen *et al.,* 2017).   A commitment was made to share the findings with the participants in an executive summary and to disseminate the research more widely. This promise was kept and reflected one of my responsibilities as a researcher (BERA, 2018).

In line with the data protection laws, all data was stored on an encrypted hard drive and securely held. Limited amounts of anonymised personal data from the participants were held in line with the new GDPR laws (European Commission, 2018) as this may be required for future research or by any participant who may request it (BERA, 2018).

6.6 Positionality

This chapter charted my journey as a researcher, and served to explain and justify my philosophical viewpoint, the choice of methods, data collection, analysis and subsequent conclusions. The chapter created space for my voice in each stage of the research and provided an opportunity to reflect on my actions, which in turn has had a transformative impact on my identity as a researcher (Dunne *et al.,* 2005).

An understanding of one's positionality as a researcher is crucial to understanding one's subjectivities (Louis and Barton, 2002). My chosen research trajectory stems directly from my experience as a secondary science teacher and assistant principal in a large secondary school. Where, in addition to teaching science to 11 to 18-year-olds, I was responsible for the whole school raising attainment policies. Consequently, the research is coloured by my perceptions and interactions with policy and how it affected my teaching practice. Over my time teaching science from key stage 3 to A-level, I was involved in the chalk-face enactment of many education policies, both curricula-based and whole-school focused. As each new policy was introduced, whole-school, departmental and personal "work" was carried out in order to deliver, what was deemed to be the essence of the reform (Ball *et al.,* 2011b). To some extent, I agree with Noyes (2010), in that research is autobiographical and whilst acknowledging the origins of my interest in the field which placed me on this research journey, I believe that bringing too much of oneself into the research risks increasing the bias in the study (Cohen *et al.,* 2017).

However, whether a quantitative or qualitative approach, the researcher cannot take themselves out of the research process. At times, this may be difficult to achieve, but limiting the impact of personal biographical subjectivities, particularly during the interview and analysis stages, can allow unexpected and unintended knowledge to come through (Suter, 2012). Growing into a new identity as a researcher necessitates owning up to these tensions and constructing a coherent narrative through the research methodology (Dunne *et al.,* 2005). Over the course of the research, and in the process of grappling with my philosophical position, I came to a better understanding of the extent to which so much of what is known is a social construction. So, my research was located in my interpretation of the social world, and therefore, validity in the research centred on the extent to which my decision-making processes were transparent, justified and ethical (Hammersley, 2011). "Research derives from the social interaction of the researcher with the researched" (Dunne *et al.,* 2005, p. 5) and as such positionality influenced my relationships with the organisations and participants in the study (Mason-Bish, 2019). When considering my positionality, I reflected on the effect that being a former-science teacher brought to my position as an insider/outsider relative to the social world of the teachers that I interviewed. Insider-outsider positionality acts along a continuum which impacts on the relationships between researcher, the participants and knowledge production, and is therefore important in the research design, epistemologically and methodologically (Mason-Bish, 2019). I placed myself as an outsider with insider knowledge, which recognised that the participants have a tacit understanding of their situation, which was independent of my interpretation of it.

Whilst acknowledging a degree of commonality with the secondary science teachers in the study, this was less so with the primary science coordinators, who as non-science specialist focussed more on the child's experience of the whole curriculum.

There were differences, not only in terms of gender, age and length of service for instance, but also the differences in power, intention and expectations for the study (Mason-Bish, 2019). The possible power issues for me as a researcher, lay in the fact that I analysed attainment data and its relationship to classroom practice. These two dimensions were often used by those in "authority" to make judgments about teacher performance. Taking a collaborative stance, avoided coercion (Cohen *et al.*, 2017) and helped to remove participant's perception that the research was an outside, top-down intervention, "another thing being done to them"; thereby establishing the legitimacy of the outcomes. Over the course of the research, I interviewed participants who were known to me, this brought continual awareness of my positionality and how it changes with each interview participant (Mason-Bish, 2019). Thus, every interaction with each school and the participants within each organisation was approached with the understanding that the knowledge uncovered was shaped by the inherent, unchangeable properties of me as a human being, by my past experience and relationships with participants and, also by the perceptions and intentions that the participants had of me, and of the study.

In this chapter, I began by stating my philosophical viewpoint; laid out the purposes and structure of the mixed methods design; illustrated the development of valid data collection instruments and brought transparency to the data analysis process. Furthermore, by highlighting and mitigating for the range of ethical issues which were involved in this research and revealing any potential bias as a result of my positionality, I have set out a detailed methodology. In answering the research questions, the study makes a contribution to new knowledge and the understanding of the ways that student attainment, teacher practice and education policy are interrelated.

6.7 Presentation of Findings

The next three chapters share the findings of the research study. Each chapter uses the data to address the issues raised in the research questions. The first, chapter 7, describes the quantitative analysis of the student attainment data derived from the NPD through the results of end of key stage national examinations. This includes descriptive statistics and regression analysis, which explore the trends in student attainment and the possible interaction between reforms in science education and attainment in science.

Chapter 8 describes the results generated from the pre-interview questionnaire completed by teachers. The chapter reflects back on the literature discussed in chapter 4 and systematically summarises the participants teaching practices, CPD and understanding of reform. The summary generates a picture of my participant's science lessons and how the teachers have engaged with the recent changes to the science curriculum and its assessment. The third findings chapter uses the qualitative interview data to build upon chapter 8 by offering a lens through which to flesh-out and interpret the numerical data. Chapter 9 uses voices of the participants to add meaning and context to the analysis and captures the reasoning behind the questionnaire responses. The theoretical framework (chapter 5) is used as the analytical tool with which to explore actions and practices of teachers; how they bring their experiences to bear in decision making; how they use their discretion in challenging circumstances on a day-to-day basis; what aspirations they hold for the future for themselves and their students. This qualitative approach arrives at an understanding of the impact of change in education policy from the perspective of those experiencing it. The discussion chapter that follows brings the entire analysis together, interpreting the actions of teachers against the backdrop of raising student attainment.

Chapter 7 Analysis of Student Attainment Data in Science

Introduction

This chapter addresses the research question, *"What do historical trends in student's end of key stage science attainment since 2008 show, and how does this reflect policies to raise attainment?"* What this does is look at the evidence used to drive the current reforms in science education.

In Chapter 2, I put forward an argument suggesting that government shapes education outcomes through three key mechanisms namely, marketisation, standardisation and accountability. In order to build on this, outcome data in the form of student end of key stage attainment was used to explore the basis for changes in education policy. This chapter reports the findings from the analysis of data extracted from the NPD for KS2 years 2007/08 to 2014/15, and KS4 2010/11 to 2015/16. Starting with descriptive statistics, which were used to explore attainment and highlight any attainment gaps between students with different characteristics (i.e. gender, FSM, ethnicity, SEN, and EAL), OLS regression modelling was then carried out to explore the impact of reform on student attainment, controlled by their characteristics. The three reforms explored were:

- The removal of the end of key stage 2 examinations in 2009,

- The change from modular to linear examinations at GCSE in 2012

- The change in the acceptance of qualifications which were deemed as GCSE Equivalent from 2014.

7.1 Attainment by Student Characteristic

Data from 2008 to 2015 was aggregated (see Table 17) to show the percentage of students attaining the government benchmark in science of level 4+ at KS2. Female students outperformed male students in the percentage achieving level 4+ at the end of KS2, 88.4% versus 85.8% respectively. On average, fewer students who are eligible for free school meals (FSM) achieved the expected standard of level 4+ (74.9%) compared to those who are not FSM (89.3%). Similarly, of the statemented students with additional educational needs, approximately 51.1% achieved a level 4+, against those students without (90.2%). The percentage of EAL students who achieved level 4+ in science (80.8%) is below the non-EAL students average of 88.2%. There is variation in attainment among students from different ethnic backgrounds. Of the Chinese students, 91.4 % achieved level 4+, compared to their peers from other ethnic groups the percentage meeting this benchmark was no greater than 88%.

Table 17: Percentage of Students by group achieving level 4 and above in science 2007/08 to 2014/15

Student characteristics	% of students achieving level 4 or above in science between 2007/08 to 2014/15*
All	87.2
Gender	
Males	85.8
Females	88.4
Free School Meals	
No	89.3
Yes	74.9
Ethnic Group	
White	87.9
Mixed	87.6
Asian	84.3
Black	82.5
Chinese	91.4
Other	80.1
Missing	87.9
SEN Statement or Action Plus	
No	90.2
Yes	51.1
English as an Additional Language	
No	88.2
Yes	80.8

Source: Drawn and aggregated from the NPD N=3,306,319. Detailed tables can be found in the appendices (Appendix Table M, p364).

*Note: As discussed in the methodology (chapter 6) end of KS2 Teacher assessment data for the academic years 2011/12 and 2012/13 is missing from the NPD analysis.

Government guidelines suggest that students are expected to obtain two "good" GCSEs at grade C or above in science and this measure is included in the accountability framework (DfE, 2019a). The two "good" science GCSEs could consist of a combination of Core and Additional Science or be made of two of the three

separate sciences (Zanotti and DfE, 2011). For KS4, as Table 18 shows, the percentage of students meeting the expected attainment. At GCSE females outperformed males in their science attainment, almost 51.6% of females achieved two good science passes (A*-C) against 46.4% of males. The gender gap between the percentage of students who entered biology, chemistry and physics GCSEs (triple science) and achieved an A*-B in Chemistry or Physics was negligible (15.1% males and 15.2% females).

The data also show the gaps between the outcomes of all students in both of these measures for each student characteristic, FSM or SEN or EAL aggregated for 2011 to 2016 inclusive. On average only 26.3% of FSM students achieved two good GCSEs in science against 52.3% on non-FSM students; similarly, far fewer FSM students entered and achieved in triple science than their non-FSM peers (4.7% versus 16.7%). Among SEN statemented students, 11.6% attained two good GCSE in science against 52.0% non-SEN; only 2.2% entered triple science and achieved A*-B in Chemistry or Physics versus 16.2% of non-SEN students. Although the overall number of students from Chinese heritage is relatively small (about 0.4%) of the secondary population (DfE, 2019d) they outperformed all other students across the different ethnic backgrounds. With over 75.8% of Chinese students having achieved two good GCSEs and 37.4% entered and achieved triple science.

Table 18: Mean GCSE Attainment in two sciences and in Triple science NPD matched data 2010/11 to 2015/16

Student Characteristics	% Students 2 Good Science GCSE	% Students entered for triple science and achieved A*-B in Chemistry or Physics
All	48.8	15.1
Gender		
Males	46.4	15.1
Females	51.6	15.2
Free School Meals		
No	52.3	16.7
Yes	26.3	4.7
Ethnic Group		
White	49.3	14.2
Mixed	50.1	15.2
Asian	55.6	19.0
Black	45.0	10.4
Chinese	75.8	37.4
Other	49.8	15.2
Missing	41.7	20.1
SEN Statement or Action Plus		
No	52.0	16.2
Yes	11.6	2.2
English as an Additional Language		
No	48.9	15.2
Yes	49.0	14.6

Source: Drawn and aggregated from the NPD N= 3,464,152 (see Appendix Table O, p366 and Appendix Table P, p367)

Students with English as an Additional Language performed well in their GCSEs, with 49.0% who achieved two good science GCSE against approximately 48.9% of students whose first language is English; whilst 14.6% of EAL students entered and achieved triple science compared to non-EAL students (15.2 %).

7.2 Attainment Gaps by Student Characteristic

The previous section gave the percentage of students by gender, FSM, ethnicity, SEN and EAL who met government benchmarks at KS2 and KS4. Further analysis of NPD was carried out to explore the attainment gaps between students with different characteristics in KS2 Science TA level and KS4, the highest grade achieved in GCSE.

The mean attainment levels at KS2 between the academic years 2007/08 to 2014/15 inclusive and highest point score in GCSE science and between the years 2010/11 and 2015/16 inclusive are shown in Table 19. At both KS2 and KS4 the gaps in attainment between different groups of the population are evident. Students in receipt of Free School Meals (FSM) have lower levels of attainment than their non-FSM peers in both KS2 (3.88 versus 4.28, respectively) and KS4 (28.07 versus 38.18, respectively). At KS2, the difference in mean science TA level for students in the population who have Special Educational Needs (SEN) compared to their Non-SEN peers is almost equivalent to a whole level (3.44 versus 4.31, respectively). At KS4, the gap in points core between non-SEN and SEN students is considerable at 19.38 (38.26 versus 18.88) equating to over 3 GCSE grades. There is no discernible numerical difference between the mean attainment levels by gender, in absolute terms the difference of 0.03 is less than a sub level at KS2; males 4.21 versus females 4.24. At KS4 the difference is marginal at 1.51 points which equates to about one quarter of a GCSE grade; with mean highest point score for males at 36.17 and for females 37.68. The gap between students with EAL is relatively small and remains so as students move from KS2 to KS4. At KS2, the mean science TA level was 4.07, a difference of 0.18 against non-EAL students at 4.25. For students at KS4 the difference is 0.43 (36.96 Non-EAL against 36.53 for EAL students) which would have no impact on the final grade awarded. Whilst EAL students do not constitute a homogenous group, it has been found that they have lower attainment than their non-EAL peers during primary

school but, by the end of secondary school, this gap has disappeared (Andrews *et al.,* 2017, p. 35). Finally, although the ethnicity data was missing for a proportion of the dataset it was important to include these students, in order to provide a full reflection of the variation between students from different ethnic backgrounds. At KS2, the science TA level for students with no ethnicity data was higher on average than their White peers or those from Other ethnic backgrounds (4.32 versus 4.24 and 4.15, respectively). The picture at KS4 changed slightly, the gap between the group with missing ethnicity data and other students still remains; this is 0.64 for Other ethnic groups and 1.61 White students. However, students from Other ethnic backgrounds attained marginally higher GCSE point scores (37.51 against 36.62, respectively) than their White peers.

The next section presents data which analyses the impact of the additional layer introduced through educational reform, on the attainment of students with different characteristics.

Table 19: Means and Standard deviation for end of key stage 2 and key stage 4 attainment by student characteristics

Variable	KS2 Science TA level 2007/08 to 2014/15		KS4 Highest Point score in science 2010/11 to 2015/16	
	Mean $\bar{x}$	SD	Mean $\bar{x}$	SD
Total	**4.23**	**0.74**	**36.91**	**16.64**
Gender				
Male	4.21	0.77	36.17	16.85
Female	4.24	0.72	37.68	16.39
Free School Meals				
No	4.28	0.72	38.18	16.08
Yes	3.88	0.81	28.07	17.80
Major Ethnic Group				
Ref: White	4.24	0.73	36.62	16.20
Other	4.15	0.77	37.51	16.27
Missing	4.32	0.80	38.23	20.56
SEN Statement or Action Plus				
Ref: No	4.29	0.68	38.26	15.73
Yes	3.42	0.98	18.88	18.07
English as an Additional Language				
Ref: No	4.25	0.73	36.96	16.63
Yes	4.07	0.83	36.53	16.78

Source: Drawn and aggregated from the National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

7.3 Student Attainment and the Impact of Reform

This section uses the NPD to explore the association between student attainment, student characteristics and education reform. The independent variables associated with the three reforms were

- KS2 Exam Reform: Test Years pre-2009 and No Test Years post-2009

- GCSE exam reform: Modular exams pre-2012 and Linear exams post-2012

- GCSE exam reform: GCSE Equivalents allowed pre-2014 and post-2014 No GCSE Equivalents.

The same measures were used in this analysis as above. In this case, the end of key stage teacher assessment level (science TA level) was the dependent variable at KS2. The dependent variable in the analysis of KS4 attainment was the highest grade achieved in science at GCSE as expressed as a point score. In addition to calculating the differences between the mean attainment pre and post reform, the effect size was calculated using Cohen's d where Cohen's $d = (M_2 - M_1) \diagup SD_{pooled.}$

7.3.1 Descriptive Data Analysis

The data shows that different groups of students were affected by the reforms in different ways with no uniform changes in attainment across all students. Beginning with KS2 (Table 20), overall, there was little numerical difference between pre-2009 and post-2009 science TA levels overall (4.21 versus 4.23 respectively). The effect size was 0.03 suggesting that the impact of reform was very weak. Across each student characteristic variable, the differences between pre-2009 and post 2009 attainment levels was negligible, with exception of SEN and EAL students. SEN student attainment levels marginally decreased after the reform, post-2009 (3.48 to 3.37) and attainment for EAL students marginally increased (3.98 to 4.10). But weak effect sizes (0.11 and 0.14, respectively) indicate that the degree of difference between reform period groups is negligible.

Table 20: Means, Standard Deviations, difference and effect sizes for student attainment data by student characteristics for the KS2 reform variable of interest 2008-2015

	KS2 TA pre-2009 with KS2 Exams		KS2 TA post-2009 no KS2 Exams		Difference	Cohen's *d*
	Mean(M_1)	SD_1	Mean(M_2)	SD_2	$M_2 - M_1$	Effect size
All	4.21	0.76	4.23	0.74	0.02	0.03
Gender						
Male	4.20	0.78	4.21	0.76	0.01	0.01
Female	4.23	0.74	4.25	0.71	0.02	0.03
Free School Meals						
No	4.28	0.72	4.29	0.71	0.01	0.01
Yes	3.85	0.83	3.90	0.80	0.05	0.06
Major Ethnic Group						
White	4.24	0.74	4.24	0.73	0.00	0.00
Other	4.09	0.80	4.17	0.76	0.08	0.10
Missing	4.33	0.79	4.32	0.80	-0.01	0.01
SEN Statement or Action Plus						
No	4.29	0.69	4.29	0.69	0.00	0.00
Yes	3.48	0.97	3.37	0.98	-0.11	0.11
English as an Additional Language						
No	4.24	0.74	4.25	0.72	0.01	0.01
Yes	3.98	0.86	4.10	0.81	0.12	0.14

Source: Drawn and aggregated from the National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

N= 3,306,319, 0 – 0.20 = weak effect, 0.21 - 0.50 = modest effect, 0.51 – 1.00 = moderate effect, > 1.00 = strong effect (Cohen *et al.,* 2017)

The difference in KS4 attainment (Table 21) in the mean of the highest point score, pre-2012 (modular 33.14) and post-2012 (linear 38.94) is 5.80. The gap between one grade and the next is 6 points (Table 11, chapter 6), therefore, the difference between two styles of examination (modular versus linear) equates to almost one GCSE grade. The effect size of 0.34 suggested that this was a modest effect in terms of the impact of reform. Across each student characteristic variable, the mean highest point score for all students at KS4 increased after the introduction of linear exams with increases varying between 3.55 to 8.79, with modest effect sizes.

Students in the Missing ethnic group showed the lowest increase in point score of 3.55 and weak effect size of 0.17. Students in the White and Other ethnic groups improved by 5.65 and 6.46 respectively with modest effect sizes of 0.34 and 0.38 respectively. Compared to non-FSM students who increased their point score by 4.96 (35.03 to 39.99), FSM students showed the largest increase in point score 8.79 points equating to more than one GCSE grade, with an effect size of 0.50 suggesting that there was a modest effect. Students with SEN have a larger difference in point score than non-SEN students (6.18 vs. 4.46 respectively with modest effect sizes of (0.35 and 0.27). Similarly, EAL students increased their points core in the linear GCSEs more than their non-EAL peers (6.89 and 5.37, respectively).   Of the student characteristics, the differences across gender were the least marked. Girls increased their highest point score by 5.65 (from 34.08 to 39.73) and boys increased by 5.39 (from 32.77 to 38.16), with modest effect sizes of 0.33 and 0.31 respectively.

Table 21: Means, Standard Deviations, differences and effect sizes for student attainment data by student characteristics for the KS4 reform Modular GCSE vs. Linear GCSE 2010-2016

Variable		Modular GCSE		Linear GCSE		Difference	Cohen's d
		Mean(M_1)	SD_1	Mean(M_2)	SD_2	M_2 - M_1	Effect Size
All		33.14	19.80	38.94	14.12	5.80	0.34
Gender							
	Male	32.77	19.91	38.16	14.40	5.39	0.31
	Female	34.08	19.65	39.73	13.79	5.65	0.33
Free School Meals							
	No	35.03	19.31	39.99	13.56	4.96	0.30
	Yes	22.63	19.61	31.42	15.68	8.79	0.50
Major Ethnic Group							
	White	33.02	19.34	38.67	13.68	5.65	0.34
	Other	33.20	19.61	39.66	13.82	6.46	0.38
	Missing	36.31	22.64	39.86	18.44	3.55	0.17
SEN Statement or Action Plus							
	No	35.38	19.02	39.84	13.33	4.46	0.27
	Yes	16.01	17.98	22.19	17.60	6.18	0.35
English as an Additional Language							
	No	33.59	19.77	38.96	14.08	5.37	0.31
	Yes	31.92	19.99	38.81	14.40	6.89	0.40

Source: Drawn and aggregated from the National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

N= 3,464,152, 0 – 0.20 = weak effect, 0.21 - 0.50 = modest effect, 0.51 – 1.00 = moderate effect, > 1.00 = strong effect (Cohen *et al.,* 2017)

A similar pattern was shown when the highest point scores were compared between pre-2014 and post-2014 reforms to GCSE equivalent qualifications (Table 22). The data from pre-2014 where GCSE equivalents were allowed, the mean highest point score was 33.22, where equivalents were not allowed, the score was 41.46. The difference of 8.24 roughly equates to over one GCSE grade, the effect size was 0.53 which suggested that this was a moderate effect in terms of the impact of reform. Across each student characteristic variable, the mean highest point score for all students at KS4 increased after the reform to GCSE Equivalents with increases varying between 6.61 to 15.90 with modest to strong effect sizes. Although it should be noted that scale for GCSE points differs from that of the vocational qualifications (see Table 11, chapter 6).

Non-SEN students showed the lowest increase in points score compared to SEN students with the highest (6.61 vs. 15.90 respectively) with a modest effect size of 0.44 for non-SEN students and a strong effect (1.04) for SEN students. There was a large increase in point score for FSM students compared to the increase for their non-FSM peers post reform (13.56 vs 7.29 respectively) with differing effect sizes (0.86 vs. 0.48). Across the different ethnic group variables, students in the Missing ethnic group showed the largest increase in point score compared to their White and Other peers (9.50, 8.16 and 8.73, respectively) with very similar moderate effect sizes of around 0.53. EAL students increased their points core after the reform to GCSEs more than their non-EAL peers (9.54 and 8.07, respectively). The differences across gender were surprising, girls increased their highest point score less than boys (8.04 vs. 8.39 respectively) with similar modest effect sizes of 0.53 and 0.52 respectively.

Table 22: Means, Standard Deviations, differences and effect sizes for student attainment data by student characteristics for the KS4 reform GCSE Equivalents Allowed vs. GCSE Equivalents not allowed 2010-2016

Variable	GCSE Equivalent allowed		GCSE Equivalent not allowed		Difference	Cohen's d
	Mean(M_1)	SD_1	Mean(M_2)	SD_2	$M_2 - M_1$	Effect Size
All	33.22	19.78	41.46	9.94	8.24	0.53
Gender						
Male	32.45	19.93	40.84	10.12	8.39	0.53
Female	34.04	19.59	42.08	9.72	8.04	0.52
Free School Meals						
No	34.88	19.27	42.17	9.63	7.29	0.48
Yes	22.43	19.66	35.99	10.57	13.56	0.86
Major Ethnic Group						
White	32.90	19.27	41.06	9.80	8.16	0.53
Other	33.24	19.59	41.97	10.05	8.73	0.56
Missing	35.34	23.02	44.84	10.69	9.50	0.53
SEN Statement or Action Plus						
No	35.18	18.99	41.79	9.69	6.61	0.44
Yes	15.54	17.96	31.44	11.87	15.90	1.04
English as an Additional Language						
No	33.40	19.74	41.47	9.86	8.07	0.52
Yes	31.84	20.09	41.38	10.47	9.54	0.60

Source: Drawn and aggregated from the National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

N= 3,464,152, 0 – 0.20 = weak effect, 0.21 - 0.50 = modest effect, 0.51 – 1.00 = moderate effect, > 1.00 = strong effect (Cohen *et al.*, 2017)

Although the differences at KS2 were on the whole very small, across both GCSE reforms discussed here, all students improved their level of attainment as measured by the mean highest GCSE point score, post-reform compared to the pre-reform. The increases for students with SEN, FSM, EAL were much greater than for the non-SEN, non-FSM or non-EAL peers, while the differences across ethnic group and gender was more mixed. It should be noted that the highest point score reported for GCSEs (max = 58) differs from that reported for GCSE equivalents (max = 55), this may contribute to an overall difference in the mean over time as more students follow the GCSE route after 2014.

7.3.1 Regression Analysis

In this next section, OLS regression analysis was used to explore the association between student attainment at KS2 and KS4 and reform while controlling for student characteristics. The variables, KS2 Science TA level and GCSE highest point score in science, both conform to the assumptions necessary to generate valid regression analysis - normality, homoscedasticity and linear association, making this type of statistical testing valid.

For the statistical analysis six models were specified that estimated attainment based on reform and student characteristics. Table 23 contains the unstandardised β coefficients for models for the KS2 analysis, the data shows that the β coefficients for each model were significant at the $p<0.05$ level. For model 1 (KS2 Test Years), there was a positive coefficient ($\beta = 0.012$, $p <0.05$) which suggested that students did better post reform although the effect is negligible. Taking student characteristics into account, the β coefficients for model 2 suggested that students performed worse ($\beta = -0.001$, $p <0.05$) albeit with almost no difference. The coefficients for the baseline and model 2 are identical for student characteristics, SEN ($\beta = -0.824$, $p <0.05$), being in receipt of free school meals ($\beta = -0.313$, $p <0.05$)

and EAL (β = -0.170, p <0.05) all show negative β values and are clear determinants of student attainment. The β coefficient for gender is also negative but on a smaller scale than (β =-0.008 p <0.05) the previously mentioned, meaning that girls did very slight worse than boys post reform. With ethnicity there is a very small positive β value, indicating that non-white students perform slight better than their white peers post reform (β =0.002 p <0.05). Despite the statistical significance of the coefficients, as the effect sizes discussed above show, at KS2, the impact of reform on attainment was weak.

Table 23: OLS Coefficients for KS2 Teacher Assessment level in science by pre- and post- KS2 exam reform and selected student characteristics 2007/08 through to 2014/15

Variables	Baseline Model ß	SE	Model 1 ß	SE	Model 2 ß	SE
KS2 Test Years vs. No Test (ref. pre-2009 Exam)						
post-2009 No Test			0.012	0.001**	-0.001	0.001**
Student characteristics						
Gender (ref: male)						
Female	-0.008	0.001**			-0.008	0.001**
Free school meals (ref: No)						
Yes	-0.313	0.001**			-0.313	0.001**
Major Ethnic Group (ref: white)						
Other	0.002	0.000**			0.002	0.001**
Special Educational Need (ref: No)						
Yes	-0.824	0.002**			-0.824	0.002
English as an additional language (ref: No)						
Yes	-0.170	0.001**			-0.170	0.001**
Constant	4.356	0.001**	4.217	0.001	4.357	0.001**
r square	0.122		0.000064		0.122	
n	3,306,319		3,306,319		3,306,319	

**p=> 0.05

Source: Drawn and aggregated from the National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

Table 24 contains the unstandardised β coefficients for models for the KS4 analysis, the data shows that the β coefficients for each model were significant at the $p < 0.05$ level. For model 3 (Modular vs. Linear), there was a positive coefficient ($\beta = 5.529$, $p < 0.05$) which suggested that students did better post reform. Taking student characteristics into account, the β coefficients for model 4 also suggested that girls performed better than boys post-reform, albeit with a very, very small difference against the baseline ($\beta = 0.769$ against $\beta = 0.767$, respectively), which was reflected in the effected sizes reported above.   The β values for SEN and FSM students were negative in model 4 but slightly less so in comparison to the baseline (SEN: $\beta = -16.917$ vs. $\beta = -17.766$, $p < 0.05$ and FSM: $\beta = -8.237$ vs. $\beta = -8.276$ $p < 0.05$). The β coefficients for ethnicity (non-white students) and EAL show small positive values for the post reform variable. When compared to the baseline the β coefficients for non-white students increased slightly from $\beta = 0.002$ to $\beta = 0.007$ while for EAL students the β coefficients decreased by almost half from $\beta = 0.304$ to $\beta = 0.167$.

Models 5 and 6 (Table 24) specify the β coefficients for the second KS4 reform considered in the study, where GCSE equivalent qualifications were no longer allowed. For model 5 (GCSE Equiv. allowed vs. GCSE Equiv. not allowed), there was a positive coefficient ($\beta = 8.231$, $p < 0.05$) which suggested that students did better post reform. Taking student characteristics into account, the β coefficients for model 6 also suggested that girls performed better than boys post-reform, with a very small difference against the baseline ($\beta = 0.771$ against $\beta = 0.767$, respectively).

Table 24: OLS Coefficients for highest KS4 point score in science by KS4 reform and selected student characteristics 2010/11 through to 2015/16

Variables	Baseline Model ß	SE	Model 3 ß	SE	Model 4 ß	SE	Model 5 ß	SE	Model 6 ß	SE
Modular vs. Linear Reform (ref. modular)										
Linear			5.529	0.018**	4.618	0.017**				
GCSE Equiv. vs. No Equiv. (ref. GCSE Equiv.)										
No Equiv. accepted							8.231	0.017**	7.092	0.017**
Student characteristics										
Gender (ref: male)										
Female	0.767	0.017**			0.769	0.017**			0.771	0.016**
Free school meals (ref: No)										
Yes	-8.276	0.026**			-8.237	0.026**			-8.094	0.025**
Major Ethnic Group (ref: white)										
Other	0.002	0.000**			0.007	0.000**			0.013	0.000**
Special Educational Need (ref: No)										
Yes	-17.766	0.033**			-16.917	0.033**			-15.951	0.033**
English as an additional language (ref: No)										
Yes	0.304	0.026**			0.167	0.026**			0.048	0.026**
Constant	38.754	0.013**	33.408	0.015**	35.736	0.017**	33.224	0.12**	35.215	0.015**
r square	0.115		0.026		0.132		0.060		0.158	
n	3,464,152		3,464,152		3,464,152		3,464,152		3,464,152	

The β coefficients for SEN and FSM students, while still negative in model 6, were slightly higher in comparison to the baseline (SEN: $\beta = -15.951$ vs. $\beta = -17.766$, $p < 0.05$ and FSM: $\beta = -8.094$ vs. $\beta = -8.276$, $p < 0.05$). The β coefficients for students with EAL showed a relatively large difference compared to the baseline, $\beta = 0.048$ and $\beta = 0.304$ ($p < 0.05$) respectively, an indication that EAL whilst still significant, contributed less to the attainment post reform. When compared to the baseline, the β coefficient for students from Other ethnic backgrounds increased from $\beta = 0.002$ to $\beta = 0.013$ ($p < 0.05$), an indication that being a non-white student made a greater contribution to final attainment after the reform.

7.4 Summary

The analysis focussed on the attainment of students at two key points in their science education and explored the trends in attainment over time, by student characteristic and in response to reform.   The data shows that student attainment at KS2 and KS4 is more closely related to student characteristics i.e. gender, FSM, ethnicity, SEN or EAL. At each point in the analysis it was found that particular students did not attain in line with the average, irrespective of reform, the attainment gaps still persisted.

There is more convincing evidence surrounding the impact of reform on attainment at KS4 than at KS2. Despite the increase in mean Science TA level at KS2 for all student groups the low effect sizes show that the difference between the pre- and post-reform attainment was relatively small. Therefore, the impact of the reform which removed KS2 examinations had limited effect on student attainment overall. The decrease in the percentage of students reaching expected progress in the years following the introduction of the new national curriculum might be read as evidence

for the impact of reform. Although the statistical analysis indicated that the removal of the KS2 exam made no difference to the teacher assessment grades awarded after 2009, it is evident that the KS2 teacher assessment grades from 2016 onwards have declined (see Table 7, chapter 3). Explanation for the decline could be related to a number of factors, for instance the increased cognitive demand of the new curriculum or teacher's lack of certainty in awarding the new grade if they feel less confident that the students have reached the required level of proficiency.

Both KS4 reforms analysed in this study generated improved mean attainment scores for all students across all characteristics, giving rise to a conclusion that the impetus for reform was justified and the mechanism of change has delivered the desired outcome. Coe (2002) suggested that the standard deviation in core subject GCSE grades was between 1.5 – 1.8 grades, and an improvement of one GCSE grade represented an effect size of 0.5 – 0.7. Therefore, even a weak effect size, around 0.2, would make a considerable difference particularly if applied across all curriculum subjects (2002, p. 7). In this study, the statistical testing indicated that reform to GCSEs contributed to an increase in attainment for all students but the effect sizes calculated for the difference in GCSE attainment pre- and post-reforms ranged from weak to moderate (see Table 21 and Table 22) and supported the conclusion that the changes to GCSE assessment and accountability regime impact student attainment but this varied by student characteristic. However, despite the overall increases, several other factors might also be associated with the changes in attainment. A reduction in the number of GCSEs entered, a shift from Applied Science to Core and Additional Science GCSE (Burgess and Thomson, 2019) alongside the changes to the examination of practical science make it difficult to precisely unpick direct associations between variables in the data.

In summary, the quantitative strategy, though a relatively small part of this research, contributed to the field by positioning the attainment of students in science as a generative mechanism with the causal power to effect change. The trends and patterns in the data can be associated with and used to justify the most recent reforms to science education and science teaching. The next chapter analyses the outcomes from the Teacher Questionnaire and uses this to describe and explore what teachers do in the classroom.

Chapter 8 Analysis of Teacher Data – Descriptive

8.1 Participants and their practice

This chapter reports on the responses to the self-administered pre-interview questionnaire completed online, via a secure university link, by the 26 teachers who consented to take part in the study. As part of the mixed method design, the structured questions were linked to the interview schedule and formed the basis for the exploration of teacher's experiences of reform and how this influences their practice.

The five-section questionnaire collected biographical information in addition to responses on the teacher's classroom practice, confidence and understanding of recent reform measures (see Appendix A3, p338). Threaded throughout this chapter, alongside the questionnaire responses, are extracts selected from the interview transcripts. These short commentaries add greater detail to the quantitative reporting of the questionnaire and illustrate key points in the participant's classroom practice. Selection of the illustrative extracts corresponds to the semantic codes derived from the interview data as described in chapters 6 and 9, and include categories, for example, "KS3", "Teaching and Learning Activities" and "Curriculum" and nodes such as "Time Pressures", "Practicals" and "Teaching hours".

In two parts, this chapter presents the findings through tables and figures illustrating the associated frequencies and percentages where appropriate. Part one reports on the data surrounding teacher classroom practice and part two reports on the information generated around teacher's collaborative work and, their understanding and confidence in using the new curriculum, progress and attainment measures.

8.1.1 The Participants

The sample consisted of 12 females and 14 male teachers, of which 16 taught in secondary schools and 10 were primary specialists. The mean length of teaching experience was 14 years and 11 months, with the longest-serving teacher working for over 30 years and the least experienced teacher had been teaching for two years (Table 25).

The sample consisted of participants working in a variety of positions, in addition to main-scale teachers, including KS2, KS3 or KS4 Coordinators, Head of School or other senior leadership roles. Secondary science teachers had a range of subject specialisms, but biology was the most frequently cited specialist qualification (7 out of 16, 43.8%). Although this question may not have been as appropriate for primary participants, one primary teacher had a subject specialism in biology, two in maths and 8 of the primary participants had non-science backgrounds, but identified themselves as primary specialists.

Table 25: Anonymised biographical profiles of the participants

Alias		Gender	Specialism	Age	Years Teaching	Key Stages	Role
Amber	sfb1	F	Biology	30-39	8	3 & 4	Head of Science
Auburn	sfc2	F	Chemistry	40-49	18	3 & 4	Head of Science
Blue	sfb3	F	Biology	30-39	10	3, 4 & 5	Head of Science
Briar	smb4	M	Biology	30-39	8	3 & 4	KS4 Coordinator
Blossom	sfc5	F	Chemistry	22-29	8	3 & 4	Head of Subject
Brown	smp6	M	Physics	50-59	31	3, 4 & 5	Head of Science
Citrine	smp7	M	Physics	30-39	3	3 & 4	Main scale
Chestnut	sfc8	F	Chemistry	40-49	20	3 & 4	Head of Science
Cerise	sfb9	F	Biology	40-49	18	3, 4 & 5	KS3 Coordinator
Cerulean	smc10	M	Chem/Phys	50-59	28	3, 4 & 5	Head of Science
Ecru	smc11	M	Chemistry	50-59	33	4	Senior Leadership
Ebony	smb12	M	Biology	50-59	29	3 & 4	KS3 Coordinator
Coral	smp13	M	Physics	40-49	23	3 & 4	Head of Science
Emerald	*smp14i*	*M*	*Physics*	*30-39*	*10*	*3, 4 & 5*	*Head of Dept*
Garnet	sfc15	F	Chemistry	40-49	9	3 & 4	Head of Science
Green	smb16	M	Biology	40-49	22	3 & 4	Head of Science
Jade	pmks21	M	Maths	30-39	2	2	Science Co-ord.
Jet	*pfks22i*	*F*	*Biology*	*50-59*	*22*	*2*	*Science Co-ord.*
Mustard	pmks23	M	Primary	40-49	7	2	Science Co-ord.
Khaki	pfeyfs14	F	none	22-29	2	EYFS	Main Scale
Lavender	pmks25	M	Primary	30-39	4	2	Science Co-ord.
Mocha	pfks26	F	Primary	40-49	12	2	Science Co-ord.
Olive	pfks17	F	Primary	40-49	20	1	Science Co-ord.
Lime	pmks28	M	Maths	30-39	6	2	Science Co-ord.
Lilac	pfks19	F	Biology	50-59	19	1	Head of School
Lemon	pmks210	M	English	40-49	15	1 & 2	Science Co-ord.

i = independent school

In the analysis that follows the response count may differ from the sample size (n=26) as teachers work across more than one key stage, as shown in Figure 6. However, this breakdown by key stage enabled an exploration of how teachers adapted their teaching for different year groups.
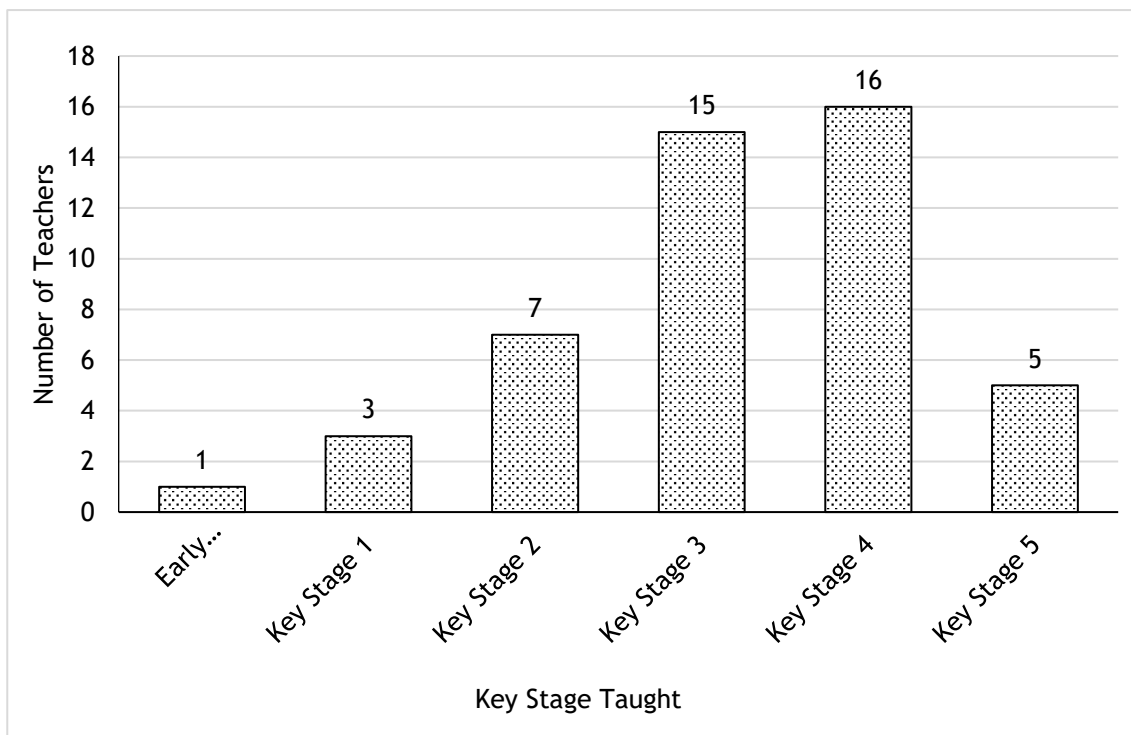
Figure 6: The number of participants teaching at each key stage

8.1.2 Teacher Classroom Practice

The second section of the questionnaire listed a range of teacher-led and student-led classroom practices and asked respondents to indicate how often these were usually carried out. Due to word limitations, this analysis primarily focuses on the first question which asked, *"How often do you carry out the following activities in your classroom*?"* and *"Has this changed?"* This question used 14 closed ordinal responses listing a range of teaching activities common to many classrooms and representing classroom practices derived from the TALIS survey (OECD, 2013b). Responses ranged from *"Every Lesson"* or *Almost Every Lesson*" to *"Never"*. Included in the scale was an option to select *"Not Applicable"* (N/A); this ensured that every question was completed (see Table 26).

Of the teaching practices listed, *checking learning by asking questions* was most frequently reported as being carried out *often* or *in every or almost every lesson*, by 25 of the 26 respondents. *Explicitly stating learning goals* was also a common practice, 22 of the respondents claimed to do this *often* or more, alongside *posing open-ended questions* (23 teachers of 26), *checking exercise books* (21 teachers of 26) and engaging in whole-class discussions (24 of 26 teachers).

Table 26: Frequency of teaching approaches used in the classroom

Activity	Every or almost	Often	Sometimes	Rarely	Never
I present new topics to the class (lecture-style presentation).	4	10	9	1	2
I explicitly state learning goals	16	6	3	1	-
I review with the students the homework they have prepared.	3	11	2	5	3
I give different work to the students that have difficulties learning and/or to those who can advance faster.	3	9	12	2	-
I ask my students to suggest or to help plan classroom activities or topics.	1	2	9	9	5
I ask my students to remember every step in a procedure.	2	7	7	6	2
I check, by asking questions, whether or not the subject matter has been understood.	19	6	1	-	-
At the beginning of the lesson I present a short summary of the previous lesson.	7	10	8	-	1
I check my students' exercise books.	11	10	5	-	-
I work 1-to-1 with individual students	2	11	6	4	2
Engage the whole class in discussions	13	11	2	-	-
Pose open-ended questions	14	9	3	-	-
I administer a test or quiz to assess student learning.	3	14	4	2	3
Organise practical hands-on/laboratory science activities or investigations	9	15	2	-	-

Looking more closely at the break down by key stage, there was little difference between key stages in the use of questioning to check for understanding. Figure 7 below depicts the number of teachers and the key stages in which the practice was carried out.
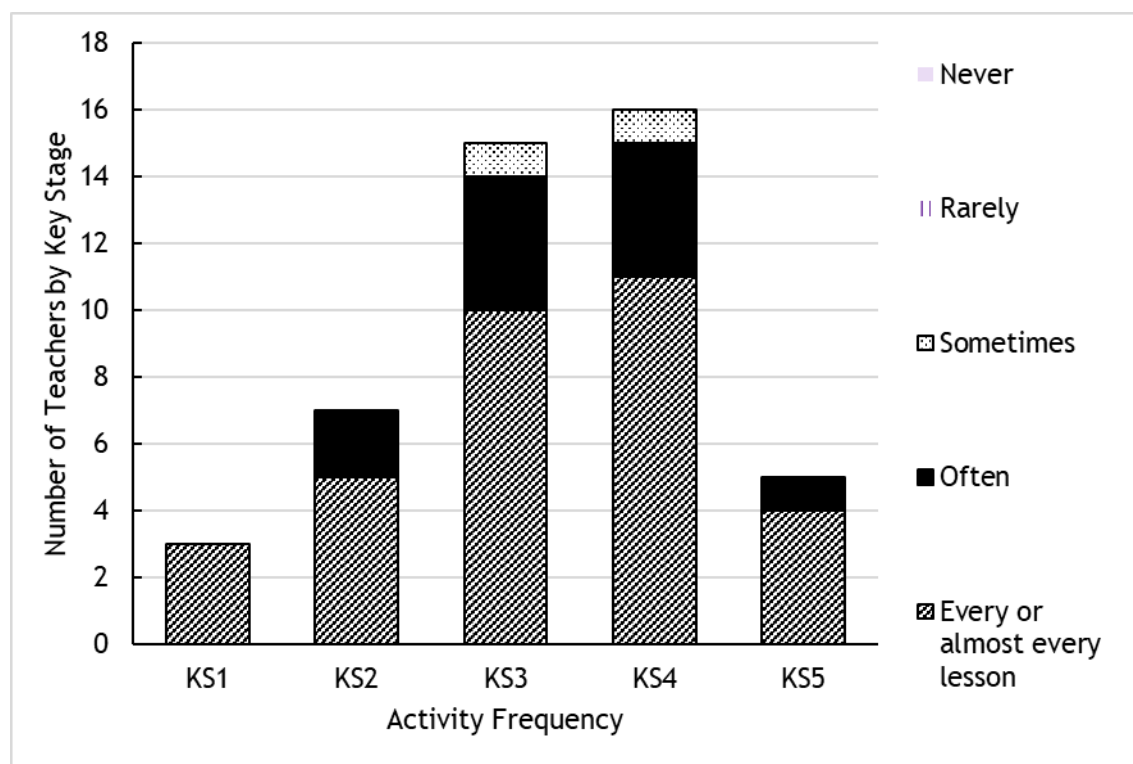
Figure 7: Frequency of teachers use questioning to check understanding at each key stage n=26

Almost all of the participants (25 of 26) indicated that *questioning for understanding* was part of their routine classroom practice, occurring often or in almost every lesson. For all of the participants, this aspect of their practice had not changed over the past two years. However, whilst continuing to include questioning in their practice, one participant remarked that their questioning style was slowly changing in response to the new examinations. At KS4, however, questioning appeared to centre around the use of exam questions and quizzes, for knowledge rather than understanding. Nine of the secondary schools subscribed to electronic learning resources such as Exampro, Testbase or Kerboodle, this indicated a move toward setting more exam-style questions both in class and for homework. Through these

electronic systems student's work is marked and given feedback with the benefit of reduced teacher workload.

Of the 14 activities listed in Table 26 above, there were nine teaching activities where teachers identified a change in practice over the past two years. Along with '*explicitly stating learning goals*', the most frequently cited were, *'posing open-ended questions' 'administering tests and quizzes'* and *'introducing topics lecture-style'*. The figures below show the frequency and key stages in which these activities took place.
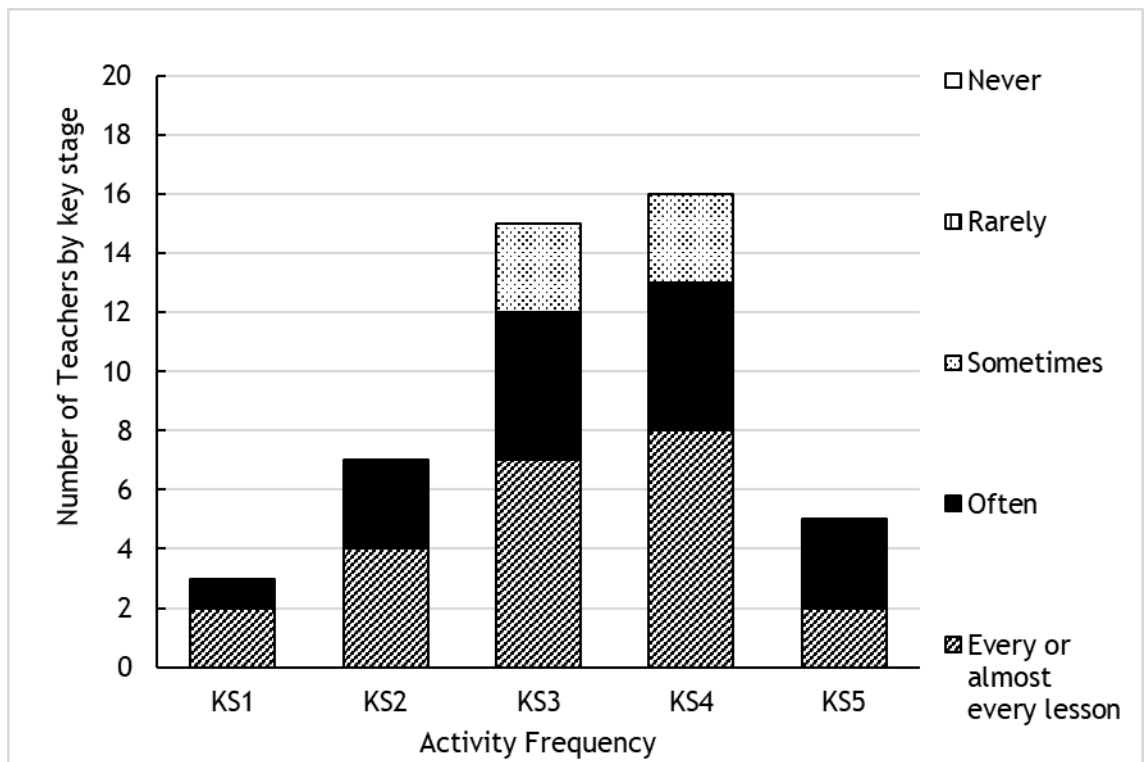
Figure 8: Frequency of teachers posing open-ended questions at each key stage n=26

'Asking open-ended questions' was regularly used by all of the participants as a teaching strategy irrespective of key stage, with 23 responses in total indicating that this occurred in every lesson (see Figure 8). Coral stated that asking open-ended questions is "where learning takes place" and specifically made time for it in the

lessons to explore student misconceptions and misguided thoughts. Teachers having the time to pursue the discussions in more detail arose as a significant issue. Blossom commented that open-ended questions were previously used to stretch the more able learners in lessons, but now, with the new GCSE curriculum, the use of this style of question is not always necessary as the course is inherently more challenging.
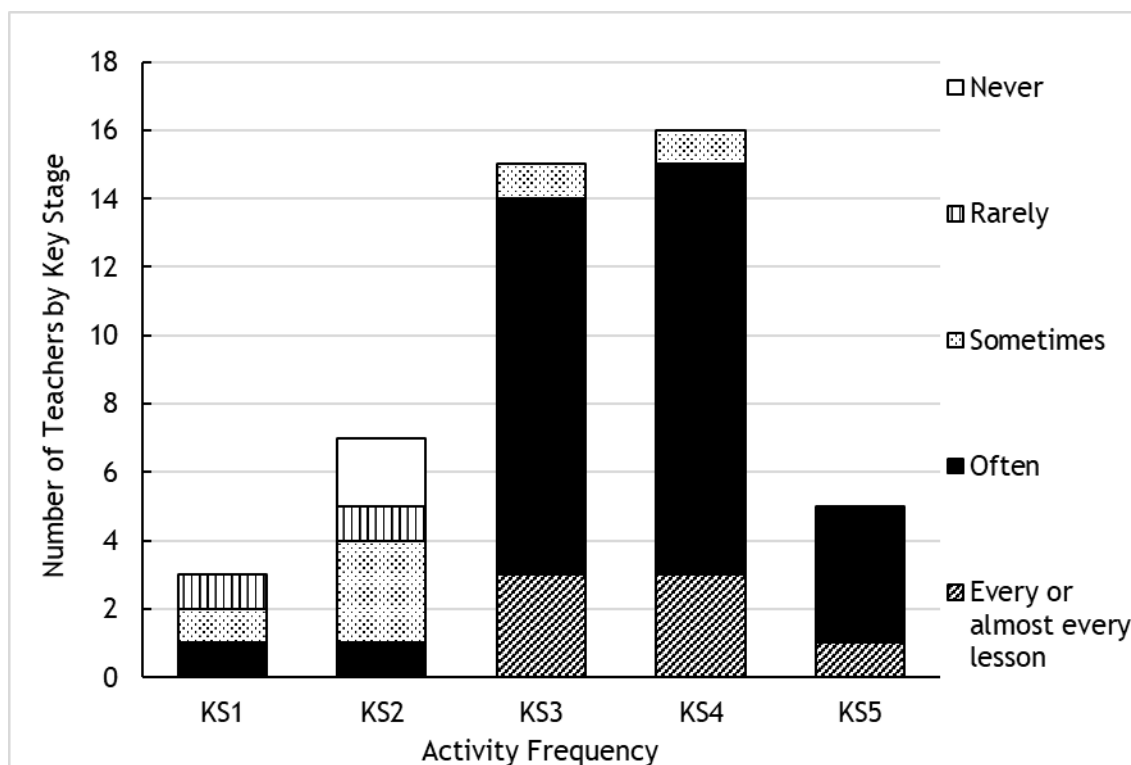
Figure 9: Frequency of teachers using tests and quizzes at each key stage n=26

The testing of knowledge recall emerged as a dominant theme for secondary teachers; Emerald talked of the increasing use of short, low stakes test and quizzes as a means to support students' memory retrieval. From Figure 9, *'Giving quizzes and tests'* to students was carried out *often*, by 11 teachers in KS3 classrooms and 12 who teach at KS4. Five of the secondary teachers remarked that they started every lesson with a mini quiz, recapping work from the previous lesson. The focus was often around learning the key equations, practicing the new maths skills and remembering the steps in the required practicals. Amber admitted that the science

department now tested the students more often in both KS3 and KS4, and felt that it was no longer about testing students' ability to apply and use the science but merely their ability to remember facts. Mini-tests were used in primary lessons but not in every lesson.

'Explicitly stating learning goals' was another key strategy of AfL handbook (DfES, 2004) and played an important part in participants' lessons (see Figure 10). Only one teacher stated that they did this *rarely*, whilst 16 of the participants reported stating the learning goals in every lesson, irrespective of the key stage overall and nine teachers reported they did this *often or more*.
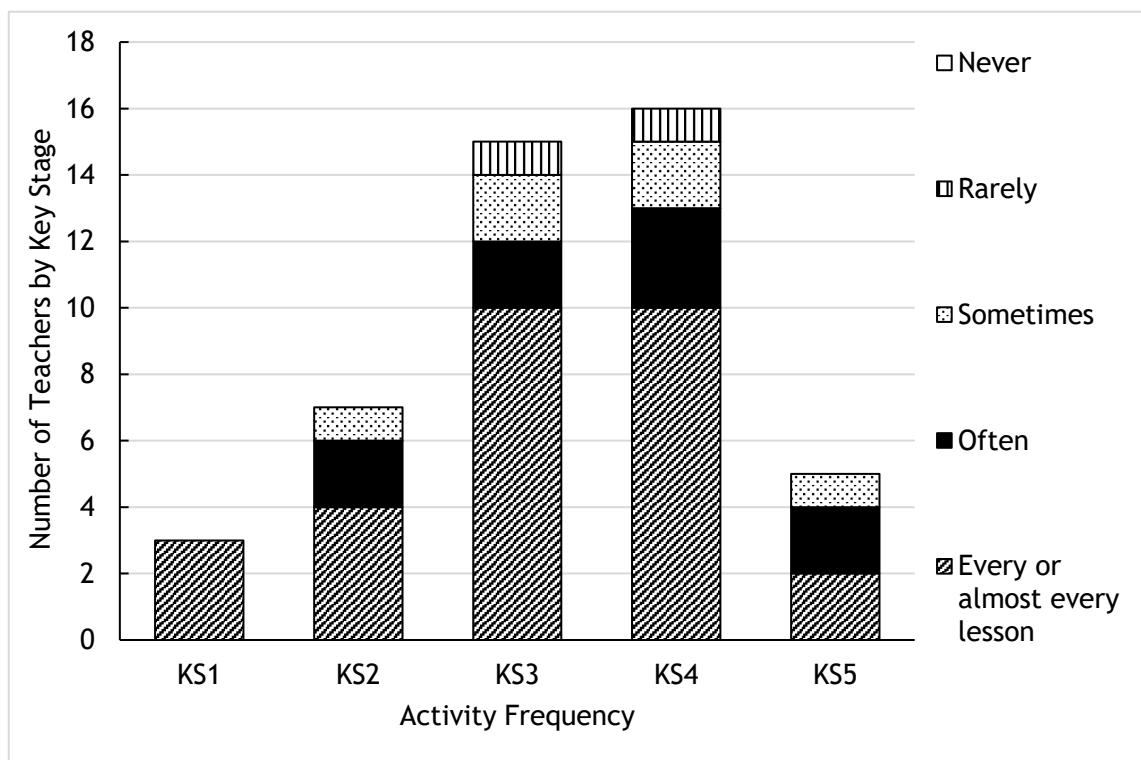
Figure 10: Frequency of teachers explicitly stating learning goals at each key stage n=26

Briar stated that as part of the science department policy, the lesson objectives were written into the student's books, and acted as future learning records and

revision prompts. Contrastingly, at the independent secondary school, Emerald

pointed out that there was no specific teaching and learning policy that required

teachers to adhere to particular strategies or templates regarding lesson objectives

or learning intentions. Emerald found it "refreshing" to be "largely left to get on

with it", and were happy with the level of autonomy in the classroom.

Figure 11 portrays the extent to which teachers use a lecture-style delivery to

introduce topics.
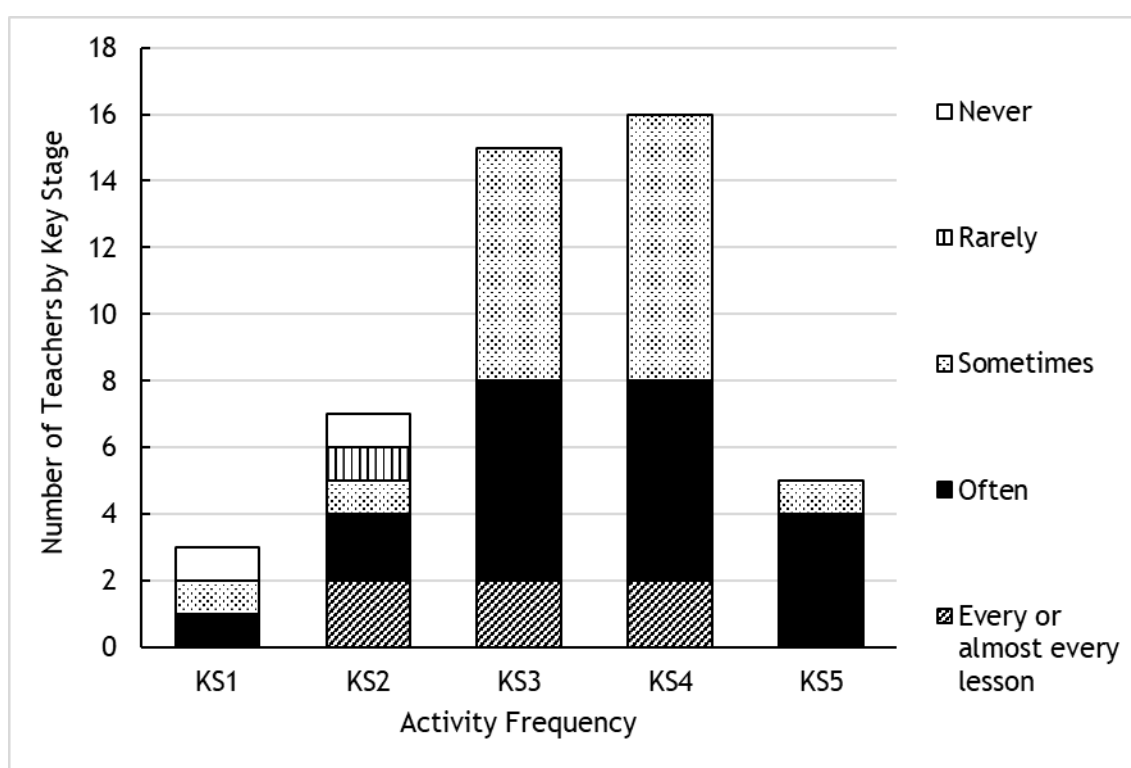
Figure 11: Frequency of teachers introducing topics lecture-style at each key stage
n=26

Some KS1/KS2 teachers claimed to *never or rarely* use lecture-style teaching to

introduce a topic but three primary participants stated that they did. For example,

Lavender, combined using lecture-style teaching during teacher demos, before the

students took part in practical activities. Figure 11 suggests that secondary teachers

at KS3, 4 and 5 used lecture-style *often* when introducing new topics but not in every

lesson. Emerald had moved away from lecture-style in favour of group-work and

active learning, Auburn, however, saw an increase in the use of lecture-style teaching to the detriment of more student-led lessons and added that this was due to the time constraints bought on by the increased subject content at GCSE. There was agreement amongst the participants that there was far more lecture-style teaching occurring at KS4 following the increased content and demands in maths skills; particularly with Triple Science groups, where it was felt that there was much less curriculum time to complete the course.

Whole class discussions remained a feature of classroom activity, 13 participants reported doing this in every lesson. Figure 12 suggests this was still a widely used method of instruction with every teacher claiming to hold discussions with their classes across each of the key stages. This form of interactive dialogue sits alongside open-ended questioning and was reflected in how teachers structure class debate, how they started a new topic or how they handled feedback arising from homework corrections.
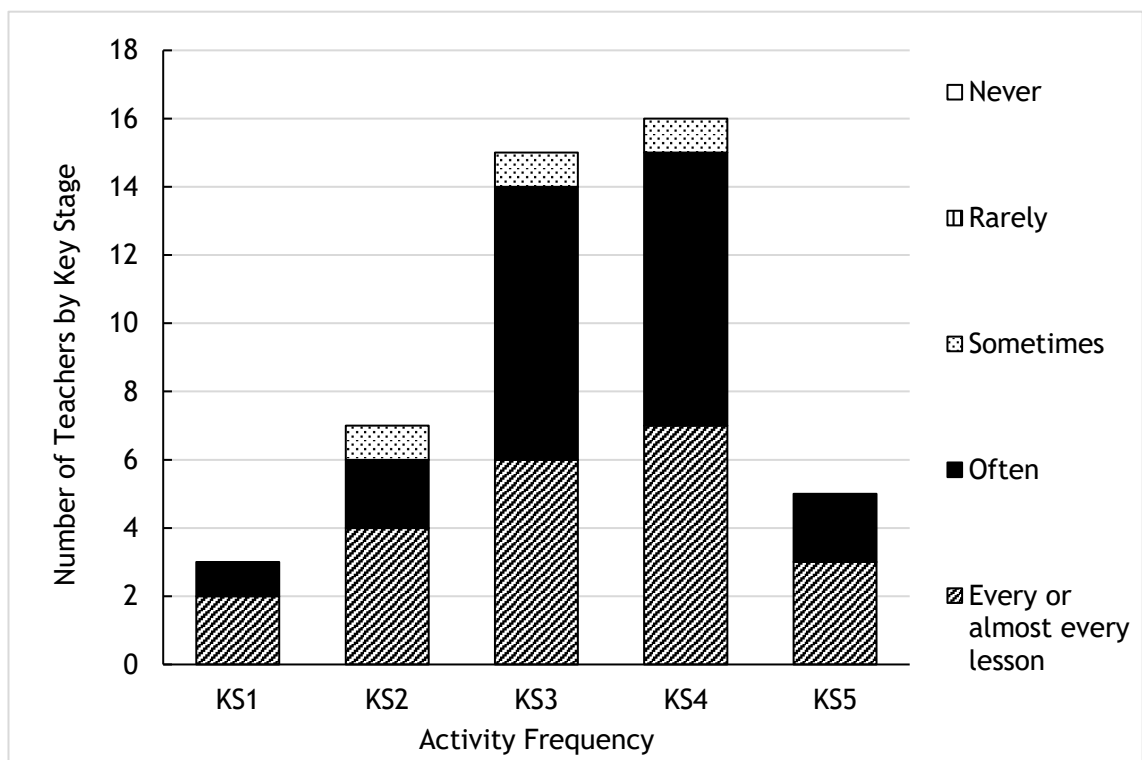
Figure 12: Frequency of teachers carrying out whole-class discussions at each key stage n=26

Practical activities were often considered to be the "fun" part of learning science and enjoyed by the majority of students (Toplis and Allen, 2012). This may explain why laboratory science activities were *often or nearly always* used by the participants in their lessons (see Figure 13). The primary teachers emphasised the importance for students to have direct hands-on experiences in practical investigations; adding that open-ended investigations tended to be spread across several lessons and the access to budgets and supplies for practical activities was always an issue. Teachers were concerned that budgetary constraints affected their creativity and ability to react to the unanticipated questions of the students by using practical activities. For instance, Olive said "…around science week, where you're asking people to come up with fun ideas but then you are saying that we've only got £20 or something….". Mustard remarked that equipment was often not available and that the school begged and borrowed what they needed.
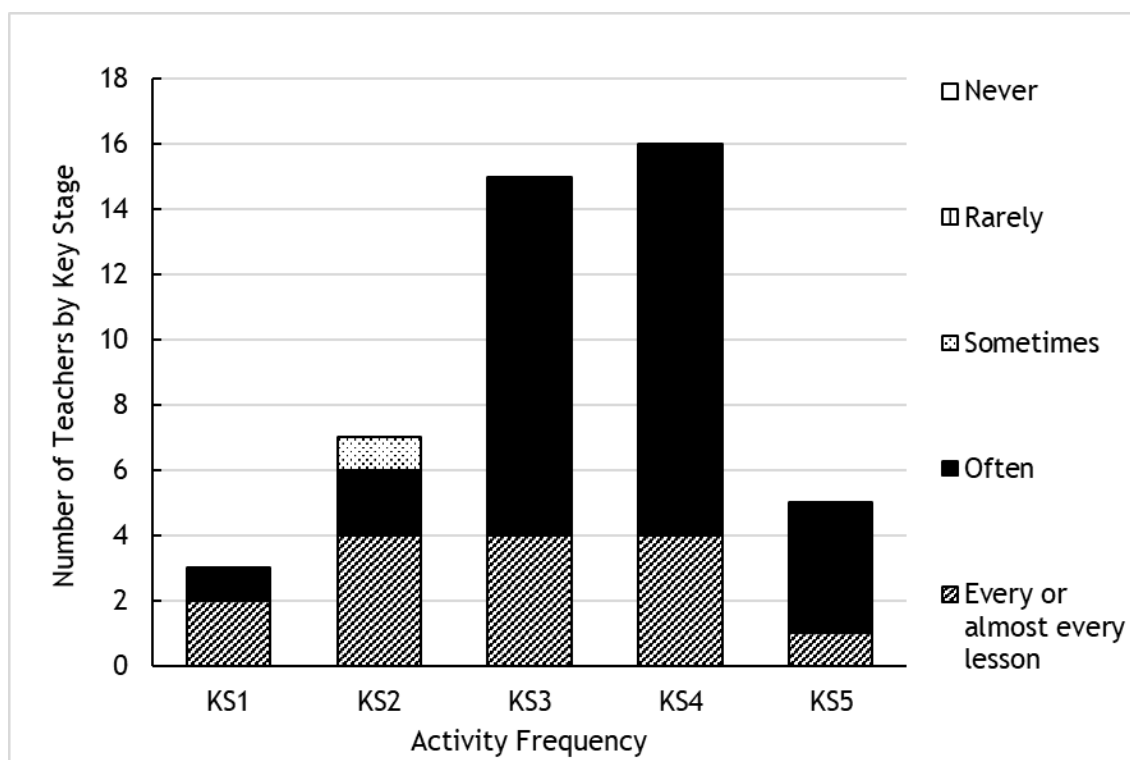
Figure 13: Frequency of teachers organising laboratory science activities at each key stage n=26

One secondary teacher, Blue, saw practical activities as "*a very happy excuse to get the kids to do science*", another teacher was adamant that he did no physics without a practical, but noticed that some colleagues were substituting whole class practicals for teacher demonstrations. Deemed as a time for KS3 students to gain experience of open-ended investigations, with practicals every lesson, laboratory science at KS4 was mainly concerned with ensuring that students completed the GCSE required practicals. However, provision for student's skills development in this type of assessed practical work was entering the KS3 curriculum and incorporated through active revisions to department schemes of learning.

8.1.3 Student activities by key stage

The extent to which student-led activities were used in the classroom was measured by questionnaire items which focussed on finding out whether students worked in small groups or on projects; were given the opportunity to hold whole class discussions or debate. This section reports on what the participants said about what the students did in the classroom and how frequently these activities occurred (see Table 27).

Students *evaluating their work* was reported as a common activity in classrooms with all the respondents enabling their students to do this at least *sometimes* during their lessons. Students were not always *grouped by ability*, with seven teachers saying that they *rarely* arranged their students in this way, whereas *working in small groups* was carried out *often or more,* possibly reflecting the organisation of science practicals. The types of teaching and learning activities students encounter varied across the key stages but it appeared that four student activities; debating, project work, use of textbooks and making a product were undertaken less regularly. In their primary classes, teachers spoke of how they encouraged their students to

routinely evaluate their experimental results and discuss how they could make

improvements to an investigation.

Table 27: Frequency of teachers indicating that students carry out different activities in the classroom (n=26)

Frequency	Every or in almost	Often	Sometimes	Rarely	Never
Students work in small groups to come up with a joint solution to a problem or task.	6	8	11	1	0
Students work on projects that require at least one week to complete.	0	3	8	11	4
Students evaluate and reflect upon their own work	2	12	12	0	0
Students work individually with the textbook or worksheets to practice newly taught subject matter	1	6	7	10	3
Students hold a debate and argue for a particular point of view which may not be their own.	0	0	10	15	1
Students work in groups based upon their abilities.	2	7	9	7	2
Students make a product that will be used by someone else.	0	0	7	9	7

Dedicated Improvement and Reflection Time (DIRT) (Lamb, 2016; Beere, 2012) is a

key element in the feedback cycle of assessment for learning. DIRT is used in six of

the secondary schools to describe student evaluation of their work, particularly

following a significant assessment task. For DIRT to be effective, teachers must

invest classroom time to develop the students' ability to self-assess additional class

time is also required to enable students to read and act on the feedback and to plan

how they will move their learning forward (Wiliam, 2011). Brown was concerned that DIRT time was being squeezed despite the school's policy on written feedback and the expectation that students respond to the feedback and correct their work. He was less likely to ask GCSE students to spend time trying to find their own solutions and is conscious of "geeing them along" because of having to move to the next topic. Cerulean also remarked that finding sufficient improvement time with larger science classes was challenging.

Holding debates in class did not occur as frequently at KS4 as in other key stages, with participants giving the increased curriculum content as a factor. Whilst, 6 KS2 and 8 KS3 classrooms *sometimes* hold debates, 14 teachers reported *rarely* using this pedagogical tool. Requiring a whole lesson or a series of lessons for student debate was seen as prohibitive by a number of the teachers. Primary teachers, such as Jade linked the debates to investigations and exploring open-ended questions within whole-class teaching setting. Although Green told me that debates and role-plays were written into the scheme of learning, these were not carried out often enough and felt that the students were missing out on opportunities for public speaking.

A similar picture emerged in the use of weeklong projects. Of the participants, 11 stated that they *rarely* gave students a week or more to complete a project, whilst 8 of them said that this was something that they did but only *sometimes*.   Carrying out investigations was considered by Blossom as project work and was done regularly with a year 7 group. However, it was acknowledged that the prescriptive nature of the GCSE course meant that project work was no longer carried out at KS4. Project work for primary teachers differed to that of the secondary participants as primary science, usually taught in the afternoon, was often subsumed as part of an overall theme or topic-based curriculum delivery.

Worksheets and textbooks were used to record the results of practical activities or as differentiation tools. From the table, it appears that textbooks and worksheets are *rarely or never* used at KS1 or EYFS, but they are *sometimes* used by the KS2 teachers interviewed. One primary science coordinator was working to convince her school to maintain the benefits of greater collaborative working seen among children at EYFS instead of "everyone filling out the same worksheet and it getting stuck in their books" as she saw in Year 1 lessons. In primary science, worksheets were more likely to be used if the lesson was being taught by a Higher-Level Teaching Assistant (HTLA). KS3 use of worksheets was seen as a short cut to completing assessed tasks which required scaffolding and at KS4 worksheets were used as a quick way to disseminate calculation questions to students. Physical textbooks were too expensive particularly with the introduction of interactive whiteboards, online resources and electronic textbooks giving students remote access via an appropriate internet connection.

The analysis above looked at the day-to-day practices of the participants in my study. Much of their pedagogical practice has been tried, tested and researched in the literature on what constitutes effective teaching to raise attainment. It was evident from the questionnaire and interview data that some aspects of teachers' practice had changed whilst others had remained the same, despite school policy, changes to curriculum or other innovations. Part 2 of this chapter continues with the description and exploration of teacher responses, looking closely at their confidence and feelings toward the new attainment, progress, and accountability measures and how these impact classroom practice.

8.2 Participants and External Reform

The previous section described the different teacher classroom practices of the participants in this study. In this section, the teachers' interaction with, and attitudes toward the recently introduced changes to the science curriculum, assessment and accountability measures are reported. Firstly, teacher's access to and engagement with professional development is described, followed by a discussion of what teachers said about their understanding and feelings toward the new measures. This part of the findings also draws upon the emergent, latent themes that arose from the interviews.

8.2.1 Professional Development and Working Collaboratively

Chapter 6 set out the scope of the recent changes to the science curriculum and its assessment, as expected, the participants had been involved in curriculum and professional development to accommodate these changes. In the last 12 months, 81% of them reported developing new schemes of learning in their department following the introduction of the new national curriculum, with 13 secondary respondents claiming to be at the forefront, or heavily involved in the curriculum development process. Only three primary science coordinators claimed to be involved in the writing or reviewing of their science scheme of learning.
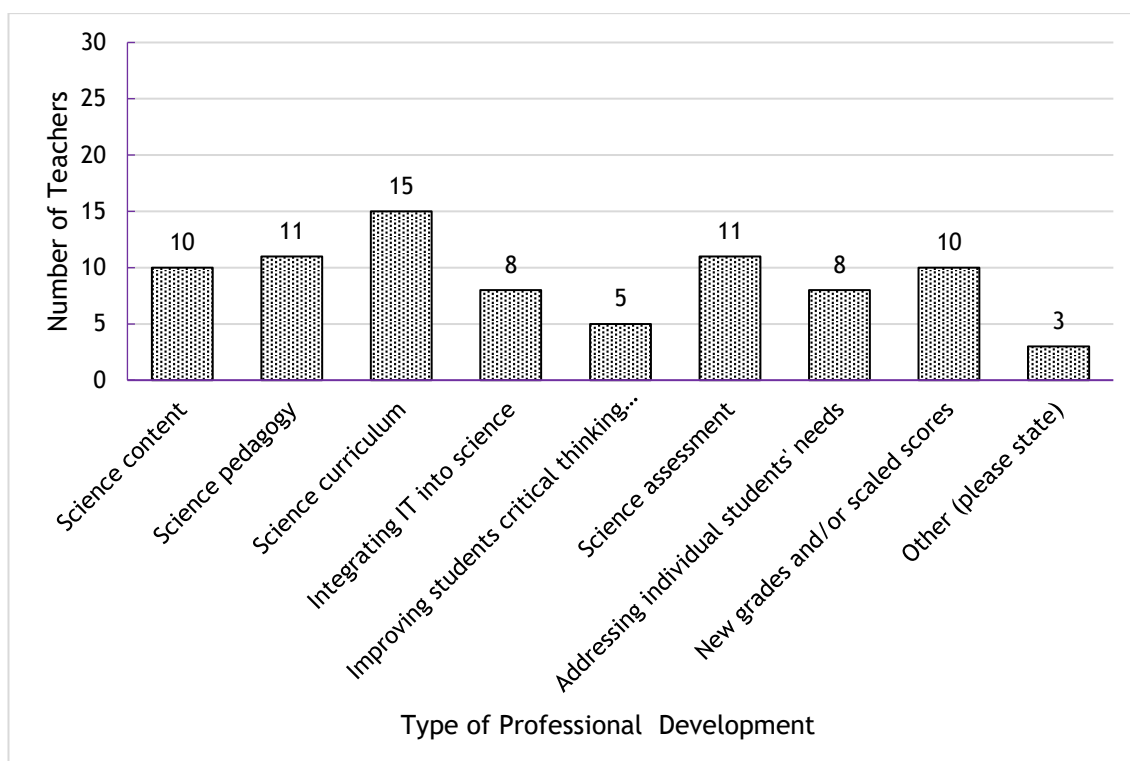
Figure 14: Frequency of the Number of teachers engaging in Professional Development by Type

A range of professional development activities had taken place over the past 12 months (see Figure 14) with teachers taking part in more than one session in that time. However, two teachers stated that they had received no CPD and a third teacher had had in-house CPD only in this time. Professional development for the new curriculum was the most commonly cited response (15 times), for the participants overall, followed by CPD focussed on science pedagogy and science assessment (11 times each).

Working as a group on curriculum development was a prominent activity and occurred at least termly for 85% of the participants. It also emerged that secondary science teachers were more involved with a diverse range of professional development than the primary science coordinators, who tended to work alone in their redrafting of schemes. The secondary teachers valued the opportunities to work together as a department or subject teams in rewriting the schemes of learning

for the new curriculum. It was clear that a number of my primary school participants often, but not exclusively relied on published schemes of learning to supplement their termly plans, others embraced the bought-in published scheme entirely.

Table 28: Frequency of teacher interaction by type

How often do you have the following types of interactions with other teachers?	Very Often (weekly)	Often (monthly)	Sometimes (termly)	Never or almost never
Work as a group on curriculum developments	11.5%	30.8%	42.3%	15.4%
Carry out joint lesson planning	8.0%	20.0%	28.0%	44.0%
Discuss how to teach a particular topic	23.1%	30.8%	34.6%	11.5%
Carry out joint marking and assessment	0.0%	16.0%	32.0%	52.0%
Visit another classroom to learn more about teaching	11.5%	11.5%	57.7%	19.2%
Share teaching experiences	20.0%	36.0%	36.0%	8.0%
Work with teachers from other phase to ensure continuity in learning	4.0%	0.0%	48.0%	48.0%
Work together to try out new ideas	3.8%	19.2%	50.0%	26.9%

Table 28 shows that taking the time to discuss how to teach a particular topic was a common activity, with 7 respondents claiming to do this *often or more*. Primarily seen in subject departments, this may be attributed to the regular curriculum team meetings held in secondary schools or access to informal spaces like a prep room. Joint marking and assessment, for 13 respondents was *never or almost never* undertaken despite being key to ensuring consistency. Moderation of the old-style GCSE science coursework in the past was a whole department exercise (QCA, 2005a) which, following the changes to the assessment of practical science, is no longer

required. Although, five primary teachers had taken part in moderation to agree on outcomes from examples of student's work both in-house, across a federation of schools and cross county, Lime admitted that there was very little time to moderate student's work in English and maths, thereby leaving science assessment to teacher's professional judgement. Keen to meet with colleagues and take part in the academy-wide review of students' work, Lime was frustrated at the cancellation of the planned moderation meeting, this may be indicative of the position of science in primary schools, in that it is no longer accorded the primacy of time and money symbolic of a core subject.

To explore collaborative working and networking outside of the participants own school context, the respondents were asked whether they had been involved or taken part in joint professional development with other schools. Almost 20% of respondents said that they had *never* been involved with joint professional development, whereas just over 42% of all teachers interviewed said that they did this once or twice a term (Figure 15).
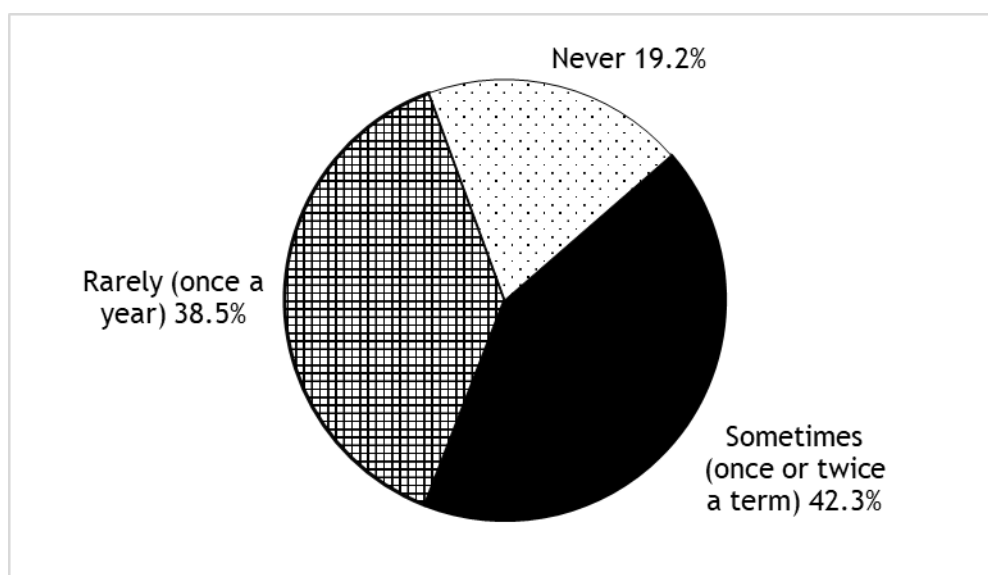
Figure 15: Percentage of teachers taking part in joint professional development in the past 12 months n=26

Joint professional development with other schools was associated with activities supported and funded in the past by the local authority. Whilst the local authority still carried out the annual health and safety checks, several secondary participants lamented the loss of interventions that brought teachers together for joint practice development days or heads of subject meetings. But there was evidence to suggest that sharing of resources, schemes of learning and assessments continued regardless. Those who were part of a larger academy chain or trust, like Blue, were still able to access joint professional development days working with schools from across the country. Subject leaders' network meetings appeared to be organised and run by clusters of schools, exam board hub meetings had also taken place at host schools around the county, but these had not always been successful. To quote Garnet … "the exam boards are as clueless as we are really…". Cerulean adds that he has delivered professional development internally, primarily to protect his colleagues from "some pretty poor stuff delivered from the exam boards". The timing for issuing specimen papers, the clarity in explaining the grade boundaries and accuracy of support material for tracking student's progress through practical activities were all cited as reasons why the relationship and trust between the exam boards and the secondary school curriculum leaders in this study appeared strained. There were more positive reports on working with subject associations and learned bodies. Three secondary heads of department praised their relationship with the Institute of Physics (IoP). The departments regularly received professional development on teaching particular topics to address the gaps in subject knowledge exposed by the introduction of the new science curriculum. This was seen as beneficial to improving the subject knowledge and confidence of non-specialists and enhanced the teaching of triple science.

In working with other organisations outside of the school, one primary science coordinator had attended training provided by the publishers of the teaching scheme the school used; subsequently, cascading the knowledge to train others. For an annual fee, which some found prohibitive, secondary schools had access to an organisation called PiXL- Partners in Excellence (PiXL.org.uk, 2017). As a replacement for support from the local authority, PiXL schedules events, provides mock exam papers and other resources, as well as introducing new ideas to schools on revision and assessment techniques such as Walking-Talking Mocks.

When asked about working with teachers from a different phase, all 26 participants wanted to do more of this. Collaborative working at the transition points was discussed as something that all teachers valued but that few had achieved systematically. The majority of secondary schools had good relationships with their feeder primary schools and visited regularly, taking boxes of equipment and working with the younger students and their teachers or hosting science week events. Getting release time for primary colleagues to work with secondary schools was raised as a barrier to greater liaison, with Coral expressing that they were hoping to arrange masterclasses for primary teachers in the future but was finding this difficult to organise in practice.

Common to both primary and secondary teachers in the study was that whilst they were positive about using non-contact time and department meetings to discuss teaching particular topics, trying out new ideas and sharing teaching experiences along with the time needed to explore the different aspects of teaching, learning and assessment was always constrained. Moreover, the teachers lamented the reduction of time and loss of opportunities to work with other colleagues from other schools, phases and organisations. The final section analyses the data around teachers'

understanding and use of the new attainment, progress and accountability measures to uncover similarities and differences between teachers from different phases.

8.2.2 Understanding the new reform measures

The final section of the questionnaire focused on teacher's understanding of the new education measures across all key stages, irrespective of the age groups they taught. Consisting of two items, teachers were asked to rate their confidence to illustrate their understanding and use of the recently introduced measures for reporting student attainment. The scale ranged from 0, *"little or no confidence"* to 100, *"highly confident";* with 50 being designated as "*confident.* Figure 16 shows the participants mean confidence rating for each of the attainment and progress measures introduced since 2014.
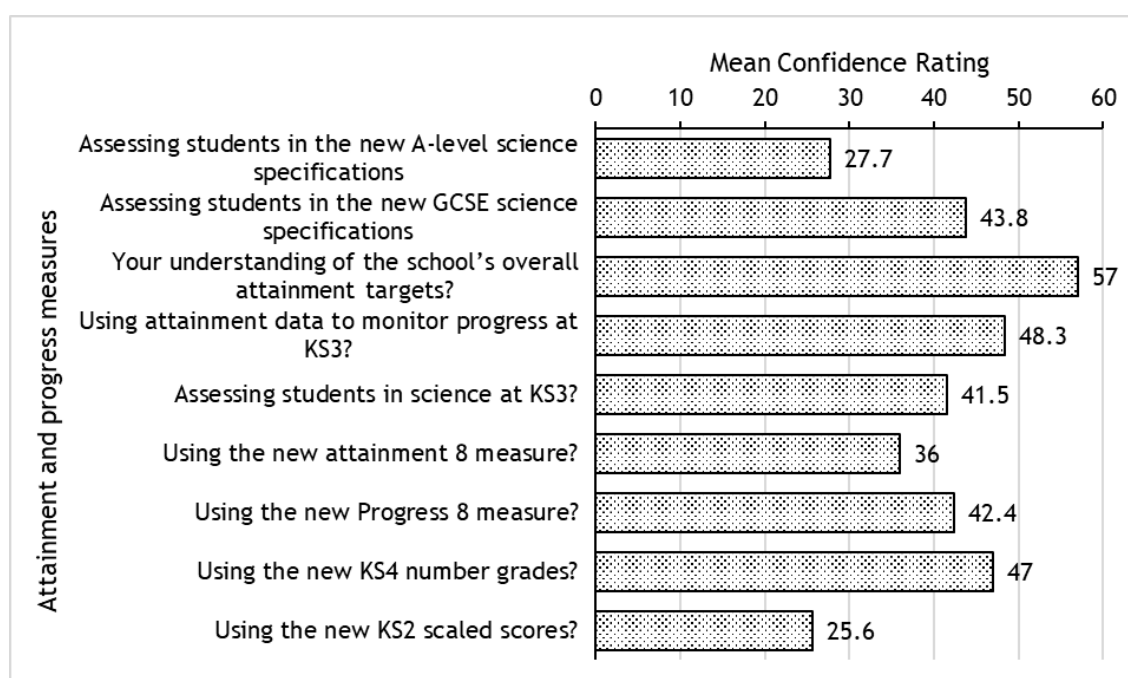
Figure 16: Teacher mean confidence rating in the different student attainment and progress measures

The data suggested little mutual understanding of the different reforms and measures across the phases. For instance, primary school colleagues had little or no understanding of the terms *Progress 8* and *Attainment 8*, and the highest rating

(max.) given by a primary respondent to any secondary school-based attainment issue was 2. Having said that, only 3 of the 10 primary respondents rated themselves with a score of 50 or above in their *understanding of the KS2 scaled scores*. Through further discussion during the interviews, it became apparent that higher levels of confidence were associated with the respondent being a parent of a child in a phase that did not teach in.

Separating the data into primary and secondary teacher's confidence levels suggested that the primary teachers have *"very low"* to *"no confidence"* (a rating of between 0 to 10) in many of the measures affecting secondary schools as can be seen in Table 29. The maximum rating in the secondary progress measures was 2, in comparison to the greater understanding of primary phase measures at a rating of 90.

Table 29: Primary teacher's maximum, minimum and mean confidence ratings in the new accountability measures (n=10)

Rate your confidence levels for the following	Min.	Max.	Mean
Using the new KS2 scaled scores?	0	90	30.1
Using the new KS4 number grades?	0	2	0.7
Using the new Progress 8 measure?	0	2	0.7
Using the new attainment 8 measure?	0	2	0.7
Assessing students in science at KS3?	0	1	0.6
Using attainment data to monitor progress at KS3?	0	2	0.7
Your understanding of the school's overall attainment targets?	1	90	42.3
Assessing students in the new GCSE science specifications	0	1	0.6
Assessing students in the new A-level science specifications	0	1	0.6

In contrast, only 6 of the 16 secondary teachers reported a zero confidence rating for the KS2 scaled scores with a maximum rating of 76 (see Table 30). Moreover, if the participant has a role, which involved supporting students through the transition from primary to secondary school, then a greater understanding of the final KS2

attainment data was demonstrated. This appeared to imply that secondary teachers were well-informed about the assessment and accountability measures that occur at the end of KS2 and possibly related to the use of KS2 assessment data in the generation of the GCSE targets on students' entry to secondary school.

Table 30: Secondary teacher's maximum, minimum and mean confidence ratings in the new accountability measures (n=16)

Rate your confidence levels for the following	Min.	Max.	Mean
Using the new KS2 scaled scores?	0	76	22.0
Using the new KS4 number grades?	30	100	67.3
Using the new Progress 8 measure?	5	100	58.0
Using the new attainment 8 measure?	5	100	49.1
Assessing students in science at KS3?	17	87	59.4
Using attainment data to monitor progress at KS3?	18	91	69.2
Your understanding of the school's overall attainment targets?	20	100	64.8
Assessing students in the new GCSE science specifications	33	95	62.8
Assessing students in the new A-level science specifications	0	95	32.6

The mean level of understanding of the KS2 scale scores for secondary colleagues was 22, which is comparable to that of the primary teachers (mean level of 30). The data in Table 30 suggested that confidence in other measures is mostly above average (50) except for the rating for *Understanding Attainment 8* and for *Assessing students at A-level*. As only five participants were reporting that they taught post 16 classes, it is understandable that confidence levels in this area were lower. However, the Attainment 8 measure was introduced in 2015, yet confidence and familiarity with this key benchmark seemed insecure.

In addition to asking teachers about their confidence in the new attainment, progress and accountability measures, it was also essential to ask about the possible impact of these new measures on their classroom practice. The frequencies of the participant

responses, Table 31 shows that the number of responses recorded as *N/A* is relatively high. For example, the new style A-levels did not apply to 18 of the respondents as they were either a primary school or an 11-16 school with no 6th form.

Table 31: Frequency of teachers views of the impact of the reforms on their teaching

	Extremely positive	Somewhat positive	Neither positive nor negative	Somewhat negative	Extremely negative	N/A
New style GCSEs	0	3	3	10	1	9
New national curriculum	0	7	6	9	0	4
New style A-levels	0	4	4	0	0	18
Focus on literacy and numeracy at KS2	2	12	3	4	1	4
Changes to the KS2 assessments	0	3	9	5	1	8
New floor targets	0	1	11	2	1	9
Progress 8	0	8	6	1	0	11
Attainment 8	0	6	9	0	0	11
Compulsory double science	1	6	4	8	0	7
EBacc	0	1	11	3	0	11
KS2 Scale scores	0	0	11	3	0	12

Analysing the data by secondary and primary phase gives a more detailed summary of what teachers felt would affect their classroom practice. Figure 17 shows the frequencies for secondary teachers and Figure 18 shows the data for primary teachers. Almost 90% of secondary teachers felt that the KS2 scale scores or the changes to the KS2 assessments were *not applicable* to them or considered their impact to be *neutral*. Significantly, 62% of secondary teachers feel *extremely or somewhat positive* about the focus on literacy and numeracy at KS2, 13% were *neither positive or negative* with only 6% considering this primary school initiative to have a *somewhat negative* effect in what they do in the classroom.

What was really interesting is that, when reported as a separate group, the secondary teachers were far more negative about the new style GCSEs compared to overall.   Eleven of the secondary teachers were *somewhat or extremely negative* about the new exams, only 3 of the 16 teachers were *positive*, of which none were exceptionally so. Furthermore, for 9 teachers (56%), their attitude toward the new national curriculum was *somewhat negative*; only 1 teacher thought that it would have a positive effect on what happened in the classroom. Contrastingly, the accountability measures of Attainment 8 and Progress 8 were seen as having a more *positive effect*, particularly the Progress 8 measure as it reflects the move away from secondary teachers concentrating on the students on the old C-D borderline to ensure that all children are given the opportunities to reach their potential.

Although, half of the secondary respondents had neutral feelings towards the Attainment 8 figure because it represented little change from what the science teams did overall in terms of contributing two GCSEs to the schools' KS4 attainment scores. The EBacc measure is viewed *negatively* by 5 of the secondary teachers, but due to the long-time compulsory inclusion of science in the curriculum, the EBacc measure was seen to have little impact on teacher activity as the vast majority of students had previously studied core and additional science GCSEs. Where the EBacc was deemed to be unfavourable was not connected to teaching *per se* but to do with a consideration of fairness and suitability for all students.
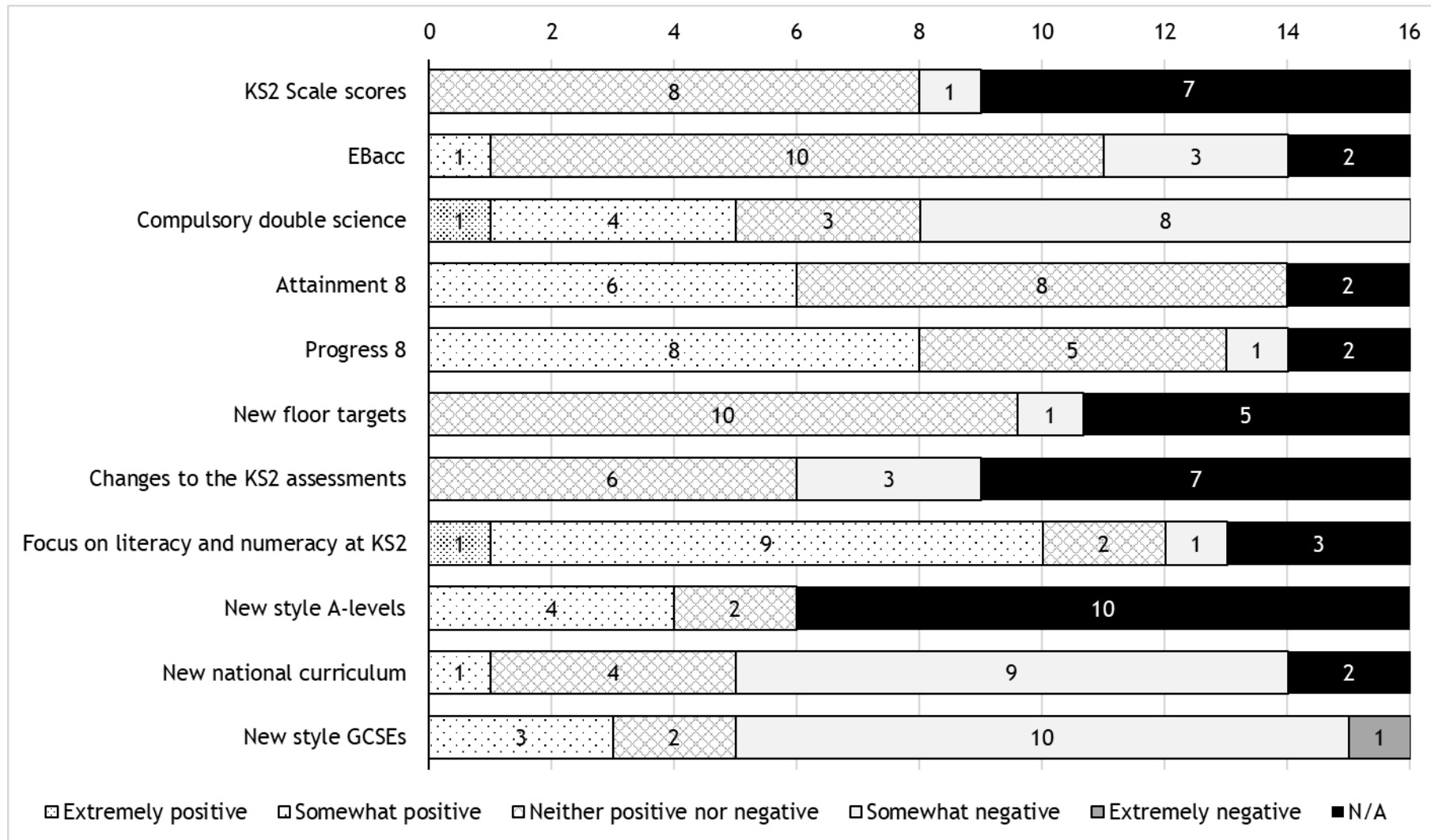
Figure 17: Frequency distribution of secondary teachers views of the impact of the reforms on their teaching (n=16)

With the primary participants, only five of the attainment, progress and accountability measures generated useful feedback (Figure 18). Reforms to the GCSEs and A-levels; Progress 8 and Attainment 8 and the EBacc were considered to be *not applicable* to the primary schooling and as such the teachers felt neutral about these measures as they would have no impact on their teaching. The new national curriculum was seen to have a *positive* impact on the classroom for the majority of the primary participants (6). But, responses to the item which asked about the impact of the focus on literacy and numeracy at KS2 was mixed, and varied depending on the perceived to impact on science due to the time spent on maths and English.
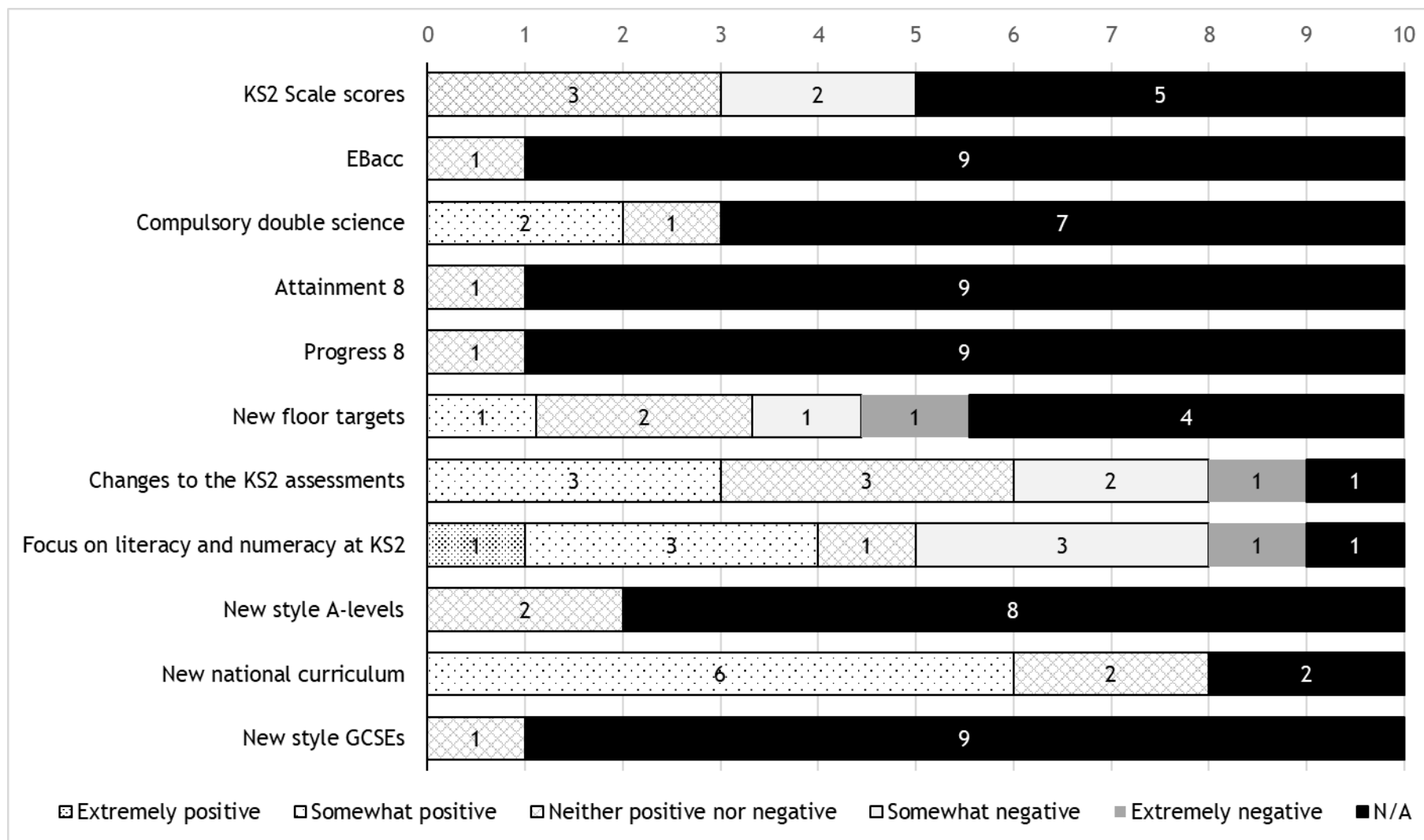
Figure 18: Frequency distribution of primary teachers views of the impact of the reforms on their teaching (n=10)

8.3 Summary

This chapter focussed on describing and quantifying teacher practice in line with the survey instrument to set the scene for a more in-depth and interpretative exploration of drivers behind teacher action which follows in chapter 9. Using the TALIS teacher survey enabled the discussion and comparison of teacher practice within an already validated framework. Strongly evident from the data was teacher's continued passion for practical, hands-on activities and their insistence on using every opportunity to enable their students to experience experimental science. Assessment for learning pedagogical practices also played a pivotal role in what teachers did in the classroom, where open-ended questions, stating explicit learning goals and giving students time to evaluate their own learning featured highly. Time constraints were a recurring theme of many teacher's narratives. Clearly, emerging was the impression that teachers never had enough time to do everything that they would like to do or everything that was expected of them. Collaborating with colleagues both in and outside of their own school context and phase was one such activity about which teachers felt constrained.

The introduction of the new curriculum, progress and accountability measures in recent years had certainly changed what some secondary teachers do with their classes in KS4. The use of examination questions for homework with the reliance on electronic software to reduce marking load and provide instant feedback was evident. The reframed focus on particular practical activities leading to fewer open-ended investigations in GCSE classrooms and the routinised memory techniques used to embed the compulsory physics equations were all geared towards ensuring that students achieve well in the new terminal exams. Although there were pockets of "resistance", this appeared within the realm of the highly confident and very experienced teacher.

In primary school settings, the time allocated to science remained a significant point of contention, as indicated in previous literature (Wellcome Trust, 2014). My research shows that little has changed. Science, despite being a core subject (DfE, 2014a), in line with English and maths, appeared to have a less critical role. In all but a couple of primary settings, science teaching received less curriculum time and was often led by staff who did not teach science at all, have the least experienced subject coordinators or staff who had been teaching for less than five years. The pace and intensity of the primary school classroom to an extent limited primary practitioner's engagement with educational developments outside of their immediate sphere of operation.

The next chapter analyses participants interview responses through the theoretical framework developed in chapter 5. This draws upon the concepts of agency and discretion to explore the impact of reform on teacher practice in greater depth.

Chapter 9 Analysis of Teacher Data – Interpretative

Introduction

This chapter focuses on interpretation of the participant's interview data through the exploration of the emergent and *a priori* themes. The chapter generates an understanding of the factors that impact teachers' classroom practice during times of policy change.

Using the methodological embodiment of Teachers as Street-Level Bureaucrats (Lipsky, 2010) and through the framework of agency, as defined by Emirbayer and Mische (1998), the chapter interprets the interview data and argues that teachers use their agency, prior experience, and future aspirations to make decisions and formulate actions, inside and outside the classroom. The research question asked *"What are teachers' perspectives on the 2014 science education reforms and how has this affected their practice?"* Drawing attention to the actions of my participants, following the changes to the curriculum and its assessment, uncovered new knowledge about how teachers mitigated change to their working landscape and, locates the study within the education literature.

The chapter is presented in five sections. The first section briefly explains the steps in the analysis and how the data was coded; detailing the semantic and latent themes, defining the key terms and laying out the assumptions and implications underpinning each theme. Next, each of the *a priori* themes, discretion, iterative, projective and practical-evaluative are applied to the data through the discussion of the semantic themes. This adds a deeper level of awareness to the specific teaching activities and attitudes to reform explored in part 1 of the Teacher Analysis Data discussed in Chapter 8. This chapter sets up the arguments for the discussion chapter (Chapter 10)

which follows and in which the analysis from the quantitative contextual chapter is integrated and this study's contribution to knowledge of the field is discussed.

9.1 Analysis Themes and Concepts

9.1.1 Coding

The teacher interviews were collected over a period of 7 months between October 2017 to April 2018. The interviews were recorded, transcribed and the resulting transcriptions uploaded to the qualitative data analysis software, NVivo. As detailed in the methodology chapter, from an initial reading of the early interview transcripts, extracts of text were highlighted which captured a cross-section of issues and experiences relevant to addressing the research question (Brooks *et al.,* 2015; Braun and Clarke, 2006). The key terms, representing common ideas or signifying emerging patterns, were identified and used to generate nodes. Through subsequent readings, the initial template of nodes was evaluated and revised to ensure that unexpected responses were included (Brooks *et al.,* 2015). Final revisions produced three semantic themes (Braun and Clarke, 2006) which were used to capture knowledge of the participants' interaction with the new reforms, these were:- *Assessment and Attainment; Teaching and Learning Activities; Curriculum.* The four conceptual *a priori* themes, devised through the theoretical framework, grouped the data as *Past, Present, Future and Discretion.* As before, through repeated reading of the text, the data selected illustrated teacher's perceptions of the challenges faced and situational contexts that guide their actions and decision-making along temporal lines.

9.1.2 The Semantic Themes

The first theme, "Assessment and Attainment", reflected teacher's responses to the areas of their practice and interaction with policy that involved the "measurement" of students. The theme brought together the different activities, expectations, and

accommodations that teachers employed in this aspect of their role. Concerned with the monitoring of student progress in science and the types of assessment used, relative to the changes to teacher practice, the theme grouped together the terms "Target Grades"; "Understanding the new number grades", "Examinations" and "Flightpath".

The second semantic theme, "Teaching & Learning Activities", drew out ideas related to teacher's beliefs, and use of constructivist and behaviourist teaching strategies. This theme coded for terms such as "Fun", "Practicals", "CPD" and "Classroom Activities" and aligned with the pre-interview questionnaire, adding greater insight to the questions on classroom activities. The theme reflected what was, is and might be carried out by teachers or students in the classroom; identifying how this might have changed and describing the challenges faced by teachers in their day-to-day teaching.

The final semantic theme concerned the "Curriculum" and was constituted by terms such as "rigour" "content", "time pressures" and "equations". This theme reflected aspects of science education which were more externally determined and in which teachers appeared to have less individual control. This theme captured responses about the appropriateness of the curriculum and the routes through GCSE science. This theme enabled the analysis to explore the extent to which teachers felt a loss of autonomy in their ability to develop an individualised curriculum for students.

9.1.3 The Conceptual Themes

Four themes, "Discretion", "Past", "Future", and "Present" were used to explore and interpret the data on a conceptual level. These captured knowledge of the participant's sense of agency in their capacities as individual classroom practitioners and as a member of a school, with a bureaucratic organisational structure.

The first section of the analysis used the theory of Street-Level Bureaucracy (Lipsky, 2010) to explore the data surrounding teacher's decision-making within rule guided contexts. As explained in chapter 5, teachers are defined as public sector workers who have regular interaction with citizens (children), where discretion in exercising their role determines the nature, amount and quality of the benefits and sanctions provided by their agencies (school) but who are also constrained by rules, regulations and directives from above. Analysing the data through this lens brought greater understanding of how the participant's actions and discretion were realised or constrained within the organisational structure.

The second section of the analysis was through the three temporal dimensions, *past, future and present* and, used the theorisation of agency as a Chordal Triad from the work of Emirbayer and Mische (1998) explained in chapter 5. Applying the iterative dimension explored how teachers drew upon established routines, selecting from practical repertoires of habitual activity and semiconscious schemas of action to get things done. This captured participant's past experience across a range of scenarios, for example, previous GCSE course work, KS2 science tests and their initial teacher training. The projective dimension explored teacher's plans for the future, their goals and aspirations, which determined how they behaved in the present. This captured knowledge about the participants' hopes for their students as well as their intentions for developing teaching and learning pedagogies to meet the increased cognitive demands of the curriculum. Analysing data from the practical-evaluative dimension, explored how teachers met the challenges of the present, as they moved through their day-to-day interactions. This theme acknowledged that teachers brought their experiences from the past and articulated their imaginings for the future into their current, concrete situation. This view of the data generated knowledge about how

teachers managed their time in class and how they talked about teaching, learning and their students.

Figure 19, is a reminder of the theoretical framework that illustrates the approach. The conceptual themes concerned with agency and discretion are used to explore and add a greater depth of understanding to the semantic themes.
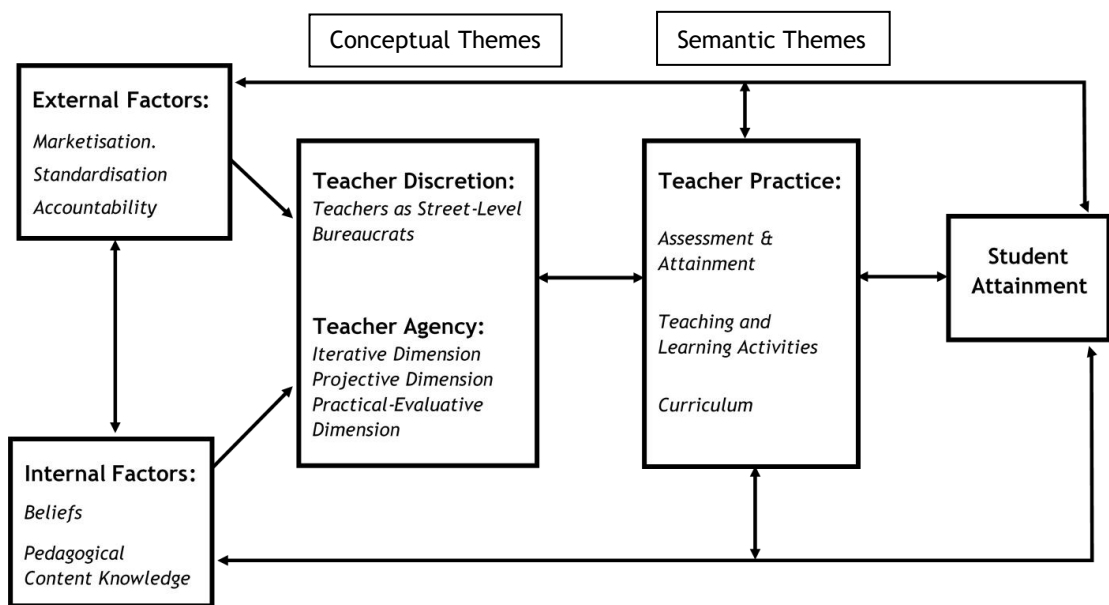
Figure 19: Theoretical Framework illustrating the association between teacher agency, discretion and practice and the structural factors in education

9.2 Classroom Practice and Teachers as Street-Level Bureaucrats

This section explores the three semantic themes of "Curriculum", "Assessment and Attainment", and "Teaching and Learning" through the lens of professional discretion in order to understand the interaction between policy and practice within the bureaucratic and structural contexts.

Irrespective of the national curriculum constraints, it can be argued that individual teachers have freedom to decide how they will teach a particular topic, what resources to use and tasks to set. However, I suggest that the accountability measures

have constrained teacher discretion for identifying suitable pathways for their students at GCSE. This was keenly expressed by one teacher, Ebony who described their concerns about "dragging everybody (students) through the really complicated stuff" and talked of their frustration following repeated attempts to introduce the entry-level certificate for particular groups of students:

> whereas if they did the certificate of science it would be more enjoyable, they get some science knowledge but we are not allowed to because it will affect the raw figure you know…. I've tried three times I've spoken to people higher up but they said that they won't do it they can't because the school's figure will go down… (Ebony)

Other discretionary decisions on who takes triple science (Archer *et al.*, 2017), enters foundation, higher tier or entry-level GCSE are now bound-up with the Ebacc performance measure, which stipulates that two GCSE science qualifications be counted in the student's Attainment 8 and Progress 8 scores. Every secondary science participant was concerned about the suitability of the curriculum for different groups of students in their school. With one biology teacher exclaiming:

> EBACC, why!... it's just why? I mean surely you want to go on to achieve stuff that you enjoy and become skilled in. And yet we're going to be having conversations about 13- and 14-year old students just to make school figures look better, and not depending on what the child actually enjoys (Green)

Referring back to the responses discussed in section 8.2.2 above, it was evident that 8 teachers were *somewhat negative* about compulsory double science. A remark by one teacher, Blue, illustrated this further by stating that there is little point delivering the more demanding curriculum material to students if they will only achieve a grade 1 at GCSE:

> I've got pupils getting grade 1s, grade 2s at best and they're expected to memorise mountains of information and it's pointless…. (Blue)

It appeared that teacher's ability to secure a science experience for the students that was appropriate to their needs had been limited by the high level of importance placed on examination outcomes. The majority of secondary teacher participants

remarked that the new curriculum was suitable for students progressing to study science to A-level but was not appropriate for the less academically able. Nearly all secondary science participants indicated that the physics equations and chemistry calculations were particular stumbling blocks. Arguably, this might be out of some need to protect the child's self-esteem, which teachers believed might be damaged by their failure to meet the more challenging and rigorous demands of the new curriculum. This extract demonstrates how one secondary biology teacher empathised with their students:

> …that's what's difficult, is trying to excite weak students who have got social problems coming out of their ears, problems at home they might be a LAC child they might be looked after they might have not slept last night cos' they you know, whatever it is and then you're …here you go bond energy! (Blue)

This extract, and others, relay the consensus among the secondary teachers interviewed, that the new science curriculum and its assessment, whilst aiming to raise academic rigour, was not always appropriate for every student. As such teachers often felt constrained by the lack of opportunity to provide alternative options to students or to present the material more engagingly. An alternative interpretation, however, might frame the difficulties in delivering the new curriculum across the ability range as a symptom of the teacher's pedagogical content knowledge (Park *et al.*, 2011) rather than the deficit lying with the student (Sharma, 2018).

However, teachers were seen to use their power to benefit classes, by changing the order of the topics taught, so that practical activities were introduced or enhanced. One teacher expressed how they introduced more light-hearted activities to an otherwise academically demanding Biology GCSE topic:

> the new B6 unit on … genetics, evolution and fossil records … yesterday we made some fossils… and I just thought you know sod it; I'm doing something fun… (Amber)

This extract suggested that a strategy for managing the increased curriculum demand was linked to teacher's beliefs and pedagogic skills, and displayed a high capacity for

discretionary decision-making, with an understanding of the possible positive and negative consequences.

Scrutiny by senior management impacted heavily on secondary participants because the outcomes for students in science are part of the overall school performance measures. Whereas, the scrutiny in primary science appeared to be driven by the leadership team and the need to meet the Ofsted criteria. To some extent, this confirmed Lipsky's suggestion that, for the most part, Street-Level Bureaucrats agree with the legitimacy of the formal structure of authority and worked towards shared goals (Lipsky, 2010), in the main, teacher's interests are the same as that of the school, both of which mirror government policy. An example from the data relates to the new accountability measure, Progress 8, a reform that has been positively received by secondary participants:

> … it is far from perfect …. But we don't have the boundaries that we had before where we were pressured to look at particular students on the C-D border. So, if a child at progress 8 gets a D and that's brilliant progress because actually they might have got an E. That's really important that is the main basis for me being positive about Progress 8 every child counts ...
> (Ecru)

This extract illustrates how the introduction of the accountability measure appeared to have shifted teachers' focus on to whole-class progress and away from students on the C-D grade boundary. However, evidence from the data also indicated that teachers had not accepted the Progress 8 model unquestioningly, in particular, where the student attainment targets were generated and circulated by senior leadership without negotiation between teacher or student.

Secondary participants displayed greater discretionary powers over how they planned the future learning of younger students. This was allied to what teachers needed to do to ensure that the students were prepared for GCSE studies, with Garnet explaining that the department planned to use "key stage 3 more strategically…" reflecting the

increased cognitive and content load. Several participants explained how, through revising their schemes of learning, more rigour was introduced into KS3 in recognition of the increased demands at KS4. A chemistry teacher reflected that, whilst the staff considered the work to be more challenging, the younger students did not feel the same way:

> So, I think for the younger students, it's kinda just how it is, you know, they have quite pacey, quite challenging lessons … but they kinda don't know any different… (Auburn).

The data suggested that teacher discretion varied according to the key stage taught but also in relation to the cycle of reform and policy implementation. Other participants explained that teaching and learning at KS3 would still retain the fun elements in science lessons that might have been removed from KS4 such as project work and open investigations. But this did not mean that there were no constraints or challenges in delivering the KS3 curriculum, time pressures and assessment issues impacted here too, as this teacher Green explained:

> We've kept key stage 3 to include as much fun as possible, with a little bit of juggling we have squeezed some bits out and narrowed it down in terms of time. (Green)

The primary teachers indicated that science was used to deliver different skills particularly through giving students hands-on science experiences. Achieving this was often at the discretion of the individual teacher as explained here:

> … what we've said to staff is, as long as you are teaching this knowledge, these skills and these concepts, then how you do it, if you have better ideas than I have then it's up to you how you teach it… (Lime)

The ability to make decisions about the timing and structure of lessons also reflected the high degree of autonomy in primary settings, as this coordinator told me:

> But teachers do not have to do science in any particular week, if they wanted to do a whole week of science in one go then they could … (Khaki)

However, another primary participant raised issues associated with organising the primary science curriculum in discrete weeks; in that there was little opportunity to revisit and build on new concepts and, that work and activities were often rushed or

bolted, on with little coherence, leaving no time to develop and practice skills or deepen understanding. This was a common theme in the data from the primary participants, from whom the impression arose that their teaching was less constrained by the curriculum but more so by time and resources.

Summary

Discretion varied, not simply between the secondary and primary sectors but within each key stage.   With the introduction of the new curriculum and its assessment, secondary teachers appeared to have higher levels of autonomy and discretion teaching KS3, than with KS4 and KS5. Primary teachers had more discretion teaching science than teaching English and maths for the same reasons, that of accountability and measurement by testing. Issues, such as the assessment of learning in KS3; procedures for recording attainment in the required practicals at GCSE or the source of primary science schemes of learning remained part of the day-to-day discretionary decision-making of my participants. However, I argue that it has become more difficult for teachers to circumvent the "rules" associated with the attainment measures. These manifest as control mechanisms, limiting discretion and autonomy, despite government policy suggesting that schools have more scope to determine what they do and how they are organised (Roberts, 2016a). My participants voiced concerns that science education had become more prescriptive and rule-governed with less scope for professional discretion, and appeared powerless to act in ways other than to manage in difficult circumstances. This was witnessed through the routinisation of activity and the adoption of client processing mentality (Lipsky, 2010) in the management of workload, for example through, rote learning of the physics equations, homework set electronically and paid for published schemes.

On one hand, the national curriculum, assessment and accountability regime eradicated some of the ambiguity in science education, on the other hand, there is

less flexibility for schools and departments to take different directions in curriculum management. Meaning that the embodiment of teachers as Street-Level Bureaucrats is not a perfect one, the boundaries around teacher's discretion and decision-making are therefore heavily context dependent. The next sections use the Chordal Triad of Agency to analyse and interpret the data. Drawing on teacher's experiences from the past to develop a future mediated through solving the present challenges, this links teacher's agentic capacity with their situational and relational dimensions.

9.3 Classroom Practice and Teacher Agency

This section explores the three semantic themes of "Curriculum", "Assessment and Attainment", and "Teaching and Learning" through the lens of agency. Providing an understanding of teacher practice and interaction with reform through a temporal perspective, I begin with the Iterative domain, followed by the Projective, with the main focus of the analysis centered on the Practical-Evaluative domain.

9.3.1 The Iterative Dimension: Past

Analysis through the iterative dimension of agency assumed that all participants had "past" experiences upon which to base their actions and decision making. More specifically, the teachers also possessed past knowledge related to science, drawn from a combination of their own experience as children, their undergraduate degree and their initial teacher training (Chen *et al.,* 2014). Four of the participants had less than 5 years' experience in the classroom, with two of the primary science coordinators new in post. One of these participants, Khaki, had not visited other science lessons and nor attended science coordinator meetings with other schools in the past. When asked how the curriculum changes had affected what they did in the classroom, the reply was:

> I haven't obviously seen anything specific because I'm not teaching science even though I'm the coordinator. The year 1 and 2 teachers haven't mentioned anything either so I think the only thing they say is obviously the issues with having the time to fit it all in. (Khaki)

In the context of curriculum change, new assessment and accountability measures, it is possible to argue that less experienced teachers do not have the resources to identify and match emerging experiences to previous types of change and possible action. Nor would it be easy for the less experienced teacher to locate changes to context, teaching practice and outcome in their limited matrix of pre-existing categories (Le Maistre and Paré, 2010), thus, making their reflection on the similarities and differences more problematic. Having said this, long years in teaching do not necessarily produce an experienced teacher with the wisdom and insight to interpret changing contexts and social relations (Brookfield, 2017).

From the data, it appeared that the more experienced teachers in my study, identified similarities between past and present, locating them within a matrix of people, context or events. In terms of action, experienced teachers compared the new with the old, looked at how circumstances had changed and how these changes were aligned with previous encounters. Take for example what Cerulean said to me about different government policies:

> I've seen all kinds of strategies come and go and group work be frowned upon and be happy and be frowned upon... So, I've pretty much stuck to what I found works for the majority of students and the happiness and the feeling of the classroom and the results as well… (Cerulean)

Locating and sorting each new experience into the already established lines to maintain familiar social structures required effort, however experienced teachers called upon existing working matrix for action, possibly rendering decision-making as taken for granted and unreflective (Emirbayer and Mische, 1998). This was exemplified by a chemistry teacher discussing the teaching of bonding:

> There's real dialogue in the learning… experience helps me with that because I almost know what's coming, what the children will get wrong… (Ecru)

Experienced or not, all of my participants were aware of the potential gaps in their pedagogic skills and the need to explore different teaching approaches for the new science curriculum. The use of purchased lesson plans and resources was one solution, as this primary teacher demonstrated:

> I came into this school four years ago and I just had a passion for science I have always like teaching science… I don't have a science degree either… we bought into a system and the head had a really big drive for science … the training is really, really thorough, the support is really good (Lavender)

My data pointed to teacher's use of selective attention in day-to-day teaching, to focus on a small area of reality, singling out the elements of response required to sustain a particular form of interaction. What this means is that teachers adhered to what they knew and repeated patterns, successfully used in the past, to generate an expected response. For example, one head of science described their teaching and interaction with students as being very traditional, remaining true to this style of teaching, routinely setting 20-question homework tasks. Although this created additional marking, the participant felt that it was an effective means to direct required interventions. This teacher, Brown, joked that "people have been trying to change me for years…".

The new curriculum, assessment and accountability measures raised the expectations for student attainment in science and outcomes for schools generally (Roberts, 2018; Long, 2017). All the secondary participants referred to the changes to the assessment of practical skills as a case in point. The imposition and subsequent accommodation of these changes created tensions and left previously held routines liable to break down, as this physics teacher exemplified:

> As an experienced teacher I know which practicals are important to keep for new teachers coming in that is a real difficulty… Gone are the days of doing a practical for investigative exploratory experience, for the fun of it, gone are those days with the time constraints we've got... (Coral)

The recent changes to the specifications for the new science GCSE examinations required teachers to challenge long-held understandings and find ways to accommodate new areas for action into pre-existing routines. One participant, Green, who had marked for an exam board for over 30 years, explained that they had benefited from the years of working in a particular way, locating their actions in a long-standing pattern. Although routines were established, actions need to be selected from practical repertoires of habitual activity and manoeuvre between possible responses is required (Emirbayer and Mische, 1998). This reflects the complex nature of teaching, in that despite established routines, unexpected interactions occur which require conscious in-situ action. The extract below refers to the participant's past levels of certainty, and how the introduction of the new curriculum had changed this:

> I've been doing the AQA course for so long now that I knew exactly what came and now you just have to keep checking back so all of us feel a little bit on slightly shakier ground than we have been… (Brown)

For this teacher and his department, what was previously known and understood was no longer fit for purpose but new interpretations had yet to be embedded.

The sometimes semiconscious incorporation of schemas of action into one's embodied practical activity, allowed the teacher to get things done quickly but contributed to the reproduction of structure in social relationships (Emirbayer and Mische, 1998). It was evident from the data that, the understanding of previous examination specifications and assessment criteria led teachers to categorise and label their students accordingly, as numbers or grades. This was seen in the data from several secondary teachers as these extracts illustrate:

Firstly, Auburn,
> …I think well I think they'll probably be about a grade B, so I think they're a 6…

And another, Coral,
> ….in the past we understood what a C grade student was we could feel it in their work, we could see it in the answers…

For my participants, the application of a generic identity to students, using knowledge built from experience with assessments and examinations, acted as a short cut to managing large amounts of information and dealing with complex interactions.

This was mirrored in the depiction of teachers as Street-Level Bureaucrats discussed above. Where teachers managed their increasing workload through routinisation of certain actions and adopted a client processing mentality. Past constructions of the "tools" required to teach can be difficult to surrender if what replaces them is in an unconstructed state of flux.

Summary

I contend that teachers drew upon their habitual activities, as the embodiment of their practice. Applying past schema gave individuals a matrix within which to locate and respond to events and social relationships, thus, allowing them to anticipate and predict what might happen in the future, bringing stability and continuity to their actions. Where this maintenance of expectation broke down was when disruptions, misunderstanding and changes in systems of relevance were introduced as seen with the revised curriculum and new accountability measures. The implications are that past experiences condition present actions, which it can be argued results in teachers relying more heavily on their experience as a coping mechanism in times of change and ambiguity.

9.3.2 The Projective Dimension: Future

Analysing participants' responses on their intentions and aspirations for the future gave an insight into how, over time, the new curriculum, assessment and accountability measures would become part of their schema of practice and routine. The internal structure of the projective dimension of agency supported the analysis of

the data by showing how teachers used their projections for the future to manage current actions and make decisions in the short and long term.

A range of future narratives stemmed from the participant's experience of previous educational changes and reflected their perceptions of current experiences. One teacher, Chestnut, predicted that poor attainment in the new science GCSE examinations, may drive the reintroduction of the coursework. This chemistry teacher equally anticipated that, due to the nature of the high-stakes terminal exams, females would no longer out-perform males in attainment or overall GCSE pass rates. Through the projective dimension, teacher's imaginations are unhindered, they have agentic power to suggest various ways to negotiate new terrain, and were expected to deliver what was asked of them, as Lemon, a primary teacher said "we asked for the change, so might as well get on with it".

The interview data also supported the idea that the projective dimension was both a construct of the individual and located in the social interactions of the department. Intersubjectively, joint projections of action gathered individuals together to formulate new strategies for the teachers in the study, this time and space was often located late in the summer term. Conflicts were re-examined, and new schemas introduced and merged with the old; planning for the next term, academic year or new year group addressed the "failures" of the past. Rewriting schemes of work was an activity born out of this reflective period, a time for evaluating teaching within a collegiate environment and planning as a team, as Green described:

> We have an inset day soon and my thinking around that is, we spend the day looking at the areas that we need to develop over the seven years of teaching that we deliver (Green)

Long term projections involved planning for student GCSE achievement from year 7. One head of department, Brown, discussed how this was carried out with their team:

> … it's going to take years for any of the changes that are taking place to see if we're teaching more in year 7 that those changes are going to come through cos we're making it harder in year 7, we're putting more formula in, we're putting more calculations in, we're toughening it up… (Brown)

This view of the future was shown in much of the data across both secondary and primary school participants. Throughout the data, participants acknowledged that the national curriculum for science was more demanding, and in terms of developing resources, reflecting on the new curriculum and new specification, planning time was eagerly awaited by all.

The data shows that hypothesising to reconfigure received schemas and generate alternative possibilities in response to problematic situations, was what teachers did to adapt their teaching on a day-to-day basis and in the longer term. This extract explains the overall picture imagined by Coral, a head of department:

> Then, if it's sticks after 2 years then you might start doing something, but most things move on. But this won't and in all my time of teaching this has been a moment of the greatest uncertainty … That's what really kills me about this whole thing because there's nothing that we as a teaching profession can do to stop it or to help it; we just need to let it run out for a couple of years and then we can see then we can respond. (Coral)

Actors insert themselves into a variety of possible trajectories and play out alternative means-ends sequences, thereby expanding their response to a given field (Emirbayer and Mische, 1998). In these hypothetical scenarios, narratives, schemas and codes are reconfigured, but possibly not in a way that reflects the teacher's beliefs about teaching (Toom *et al.*, 2015). Whilst most of the primary participants felt that changes to the science curriculum would lead to more constructivist teaching, the projective analysis of the data highlighted the uncertainty felt amongst all of the secondary participants. This was clearly stated by one teacher, Citrine, who remarked that they saw a time when he would be standing in front of the class, asking students to repeat the physics equations in a rote learning mantra akin to a Victorian classroom learning the times tables. However, many envisaged a time when their situation would alter

due to gaining an increased sense of agency within the delivery of the curriculum.

This, they believed, would bring a more creative element to their teaching in response

to the increased rigour of the curriculum.

The term *Flight Path* was used by many of the secondary participants to indicate the

future trajectories of student attainment. The Flight Path lays out an attainment

pathway against which student progress is tracked. Individual staff have little say over

the targets set for the students they teach. As Blossom explained,

> …. we use the same thing as flight paths, so they have their flight paths in year
> 7 and they stay on that same flight path right the way through to year 11
> (Blossom)

Mapping the progression of student's learning in a subject, against an increasing level

of conceptual demands is considered effective classroom practice (Krathwohl, 2002),

however, Flight Paths are constructed as aspirational, summative measures, framed in

the language of policy (Biesta *et al.,* 2015). Still, for individual teachers on a day-to-

day basis, it was evident that assessing students work and supporting their progress

was increasingly challenging, as a deeper interpretation of the grading system had not

yet been achieved. As this comment from Ecru, a member of SLT who teaches science

suggests:

> … we can then set a progress path and when you've got a child in Year 7 if their
> progress path is let's say a 5 or a 7 or a 9 we've got to look at Year 7 work and
> say is this the sort of work we would expect for that child to continue all the
> way through to year 11 and to get that final 5, 7 or 9 but we don't really know
> what the 5, 7, 9 look like yet so it's hard to tell. (Ecru)

Secondary teachers used modelling and mathematical projections to assess their

student's progress and monitor their attainment. In place of more concrete

estimations, short-term projections for attainment in the new science curriculum at

GCSE relied on the use of norm referencing and this was used to estimate the grades

that students might achieve in the examination. This head of department explained how this was done in his school:

> ... I haven't got a problem with the number grades because when they said that we're going to have the same number of... you know, similar numbers of A's and B's and C's as we always had then we just swap them over and we get 70-80% are getting an A-C so, we're expecting 70-80% to get a 9 to a 4. (Brown)

From the interviews, making projections and working towards students' future attainment played a key role in the actions of secondary science teachers. Preparing students for the next phase of school or work was also a narrative for future action projected by many of the participants. Ensuring that students were "secondary ready" as primary teacher Jade remarked and prepared to compete for jobs in an increasingly globalised work-place as Garnet mentioned:

> ... we need to be training them to come out in the end as competent and whole people and yes I want to have science of course I do, because the world is changing and I have no idea what jobs will be around in 20 years' time for science and STEM subjects... (Garnet)

All the same, participants showed awareness of the possible uncertainties surrounding the future and recognised the teacher's role in equipping students to meet these unknown demands.

Summary

Imaginative engagement with the future is considered to be crucial to human effort as actors (Bandura, 1977) therefore, as the data suggested, teachers continually project their designs for the future, reinventing and revising their practice with each year group. At times of major change, such as the introduction of the new national curriculum and the associated assessment regime, imaginings of the future appeared far more tentative. Lack of clarity and experience with the new curriculum and its assessment led teachers to utilise ideas from the past to predict the outcomes for the future. Secondary teachers recognised that this situation was unsustainable,

unpredictable and unreliable and would change once the first set of GCSE results were published.

9.3.3 The Practical-Evaluative Dimension: Present

The concrete activities of teachers have always been multiple and complex, situational and embedded in temporal locations (Wyatt-Smith and Looney, 2016); my participants were no different. On a lesson-by-lesson basis, teachers encountered situations and problems mediated through the engagement of existing schema, rules or routines (Emirbayer and Mische, 1998), for instance, behaviour management during practical activities. Exploring the participant's day-to-day actions in accommodating the new curriculum and accountability measures was particularly pertinent to this study.

Thirteen of my participants had been teaching for over 15 years and had experienced large-scale curriculum changes in the past. However, no new situation is ever the same as the old and so "work" has to be done and judgements made to render new situations less problematic. In this study, problematising the new GCSE grades was an example of the work done by my secondary participants in an attempt to integrate the new with the old. Many including the biology teacher Blue, from the extract below, explained how, when marking homework or giving assessments, the number grades were converted to the old letter grades and vice versa:

> I think until we've done one set of exams this summer... I don't think we're going to know what a 6, a 5, don't know what any of it means. And obviously, a 4 and 5 are confusing because we're trying to work out, is a 4 a C or is a 5 a C. (Blue)

Guidelines from the previous national curriculum exemplified the different levels of attainment across the key stages (QCA, 2010). Although considered to be of limited use, the descriptors for the new GCSE were untested and developed ad hoc before the terminal assessment took place (Cadwallader, 2014). It was evident from the interview data that secondary teachers required more explicit support to identify the

specific skills, knowledge and understanding students needed to demonstrate for a

given GCSE grade. This teacher, explained the complexities associated with day-to-day

assessments:

> … even a C grade student in old money, needs slightly different skills to be at that same level, so it isn't a C grade student anymore… a grade 4 to 5 has slightly different qualities to the former C grade student because they're assessed on slightly different things (Garnet)

The extract exemplified how habitual activities (Emirbayer and Mische, 1998) helped

teachers break down and characterise new situations and assign grades to student's

work.

Further evidence of the problematisation and execution of the recent changes was

found in the increased numeracy demands of the science curriculum. This extract

from another secondary teacher illustrates that whilst this was an on-going situation;

the teachers were taking ownership of the challenge and accepted the outcomes:

> Numeracy has gone through the roof and we're having to be maths teachers, we are having to teach about the tangent of a curve because of what the curriculum demands; the equations of motion are in further maths, we are having to go, right this is our little baby we are teaching this to everyone and we go yeah that's alright. (Coral)

In this practical-evaluative dimension, other themes connected to the delivery of the

new curriculum required teachers to find solutions to new problems faced in the

classroom. From the data, evidence of this was seen when talking to secondary

teachers about how they approached teaching students to memorise the required

science equations. Several teachers were experimenting and deliberating on the

various ways in which this could be done. With one young teacher remarking that

stepping into rote learning territory felt contrary to the way that he was taught to

teach, whilst a more longstanding teacher felt that a didactic style had always and

would continue to work successfully for him.

It was evident that the approach to managing the new required practicals varied by school, general guidance on exam specifications was available but was not prescriptive (Ofqual, 2015b). Some subject leaders shared good practice with other schools, but science departments decided their own path through this newly introduced element of the GCSE. One particular school explained that it was essential to revisit the practical activities before the GCSE examinations and, that this was achieved by their production of mini-videos. This new activity, having not taken previously, would feed into the future resources for the department to support the teaching and learning of the required practicals.

Few participants were able to use lesson time flexibly for activities such as pursuing student-led ideas for an open investigation. Others, as in the extracts below looked for ways to solve the problem of limited curriculum time by giving the students more homework:

> So, I'm setting a lot more homework than I ever did especially with the oldest ones because we've gone from 10 hours of contact time to 8 hours of contact time, I need them to be working. (Cerulean)

Although no primary teachers set homework in science, all secondary participants gave homework at least once a week. Indeed, the routine use of exam questions as homework was used to make up for the limited teaching time in class, as this teacher, explained:

> … so, we give a lot of exam question homework or a little pack of questions to consolidate what they've done in lessons because we, as I said, we wouldn't necessarily have time to do that in class, like we would have done in the past... (Auburn)

Time was a key factor in making finely tuned judgements on emerging situations. Teacher's deliberations looked at plausible choices, weighed in the light of practical perceptions and understanding against a broad field of possibilities and aspirations (Emirbayer and Mische, 1998). But expediency and pragmatism also determined teachers' actions in the "moment" (Heijden *et al.,* 2015; MacBeath *et al.,* 2004) such

that teachers acted in ways to cover the greatest content, for the greatest number of students, in the shortest time.

The perceived negative consequences of dealing with day-to-day time pressures were further highlighted by the secondary participants. Teachers told me that they moved on from teaching a concept before it was embedded or learned, or decided not to revisit a topic, despite knowing that this represented good practice and was beneficial to all learners. This biology teacher told me that:

> … and you can never fit it all in so, I feel like I'm continuously on a running track trying to teach them everything and so again, we've been observed and we've been told… oh you know when you do a lesson you should go and consolidate and repeat and check your learning, I said well no, every lesson we're teaching brand new pieces of information you have no time actually no time. (Blue)

This observation was confirmed by many of the secondary science teachers, who balanced their choice of teaching activities with the need to cover the curriculum subsequently, teachers were selective in what they delivered. From this evidence, it would appear that students were missing opportunities to experience a varied and enriching science education; as another teacher remarked:

> … so, I definitely say that whilst with my triple scientists I probably can take more risks, I know I can catch it up in 15 minutes probably with the higher attaining ones whereas with my… bottom set year 10 group, the thought of giving up a lesson for something …you know… something enriching, I probably would be less likely to do it. (Amber)

Time issues were common in primary settings too, making day-to-day management of learning challenging, as this teacher explained:

> The time issues are enormous in primary. We have so many things to teach and not enough time to do it in. We have about an hour a week for science but we really should have two hours… (Mocha)

I suggest that, despite the time constraints, predominantly working with the same class enabled primary teacher's agency, making them more able to adapt their teaching, react to changing situations and respond to student interactions. But I also

argue that science teaching at primary school is vulnerable; in that, it may not always be the class teacher who carries out the day-to-day teaching of science.

From the data, secondary teachers talked about losing the fun out of lessons; that students were tested more often, given less time for DIRT and extension tasks. Also, that the prescriptive required practicals replaced open-ended investigations and this had altered student's experiences of practical science and appreciation of how scientists work. One teacher commented that:

> … practical science is absolutely key and when that starts to go, which is what I see them doing in my triple science at the moment is that they don't have the time so there is more demonstration work… (Brown)

Whilst another, Chestnut, added:

> … we don't do investigations as much as we used to, I think they've kind of been squeezed out by the volume of content…

Nevertheless, there were differences of opinion on whether the challenges and uncertainty had impacted on the students, as this participant demonstrated:

> I don't think that my students realise there's any difference, they go through the system one year to the other to the other and they don't really actually see it obviously, it's an age thing, they're not particularly interested in what's happened before or after them really... (Coral)

A physics teacher, Citrine summed this up when they talked about the resources and experience called needed for teaching the new physics GCSE equations:

> … I think one of my biggest problems is that I fundamentally disagree with it and for that reason it makes it very difficult for me to teach…. If I don't believe in what I'm teaching the kids, it's very, very hard for me to do that so I think … that's causing me as much of a struggle as anything else… cos I could just stand at the front of the class every lesson for five minutes and just say we're going to go through every single formula …on and on… but it would kill me… I think it would absolutely drive me out of the teaching profession… if I felt that that was the only way to achieve it ... (Citrine)

This extract also demonstrated the on-going "struggle" faced by secondary teachers as they come to terms with developing their practice in response to the more prescriptive elements of the new curriculum.

Summary

In the current challenging situation where changes are whole-scale and wide-ranging, teachers would be forgiven for relying on their past experience to deal with their present circumstances. Applying an unconscious, unreflective reaction to situations that arise and by making decisions which have not been taken through a reflective/refractive process may, I suggest, work as a survival measure.

The practical-evaluative aspect of agency draws upon teachers' capacity to manage the every-day contingencies and uncertainties which are a feature of school life (Emirbayer and Mische, 1998). Teachers are engaged with new policy initiatives whilst managing varied classroom interactions in which unexpected issues can arise. This brings practical problem-solving skills into play which are framed within the individual teacher's personal experience of schooling, their initial teacher training and on-going CPD. Moreover, the enactment and application of school policies; social interactions and support within the department serve to generate causal mechanisms through which teaching practices are continually refined.

In the discussion chapter which follows, the analysis of the contextual quantitative data, from the national pupil database, and the qualitative data derived from the participant interviews, are brought together to address the issues raised in the research questions.

Chapter 10: Discussion

10.1.1 Situating the Discussion

This chapter draws together the analysis from the three findings chapters and in doing so, demonstrates how the new, and understood to be original, knowledge generated by this research contributes to the field. The introduction to this chapter briefly sets out the structure of this discussion and revisits the philosophical and conceptual underpinnings that were introduced earlier and used to frame the data. Taking a holistic look at the findings from the student attainment data (see chapter 7), teacher classroom practice questionnaire (see chapter 8), together with the teacher interviews (see chapter 9), I expand on how these different components of the inquiry are interrelated. The arguments arising from the contextual data and individual teacher perspectives work together to address the issues raised in the research questions. Through a critical engagement with the literature in chapter 4, this discussion chapter positions this research within the current educational debates outlined in chapter 2 and uses the theoretical framework developed in chapter 5 to bring understanding to this complex, multifaceted field. I propose that the broad topic of this research, science education and educational reform, lends itself to an exploration through the ontologically pluralist perspective offered by the critical realist lens as discussed previously in chapter 6. The diagram below Figure 20 was used to illustrate the interactive association between teacher agency, their discretion and the structural demands of science education and the impact of reform.
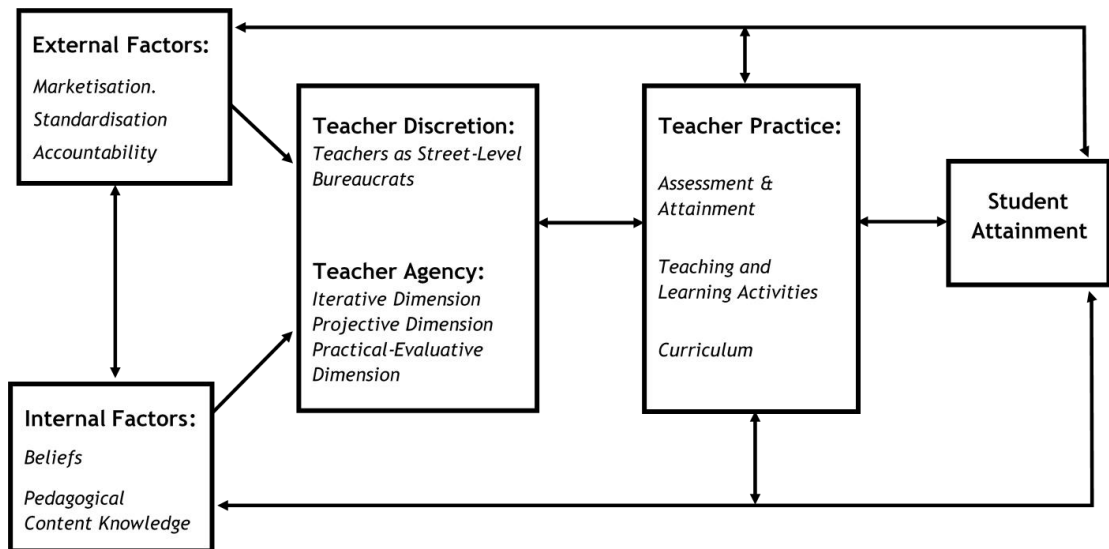
Figure 20: Theoretical Framework illustrating the association between teacher agency, discretion and practice and the structural factors in education

The interactions at the macro level of national government policy and indirectly with international policy, impact at the level of local authority and school which in turn can determine the interactions between teacher and student. The interactions at each level are imagined as a feedback loop, the illustrative framework uses double-headed arrows to symbolise where each level is connected and contributes to the outcomes of the next. The constant interaction and interrelation is complex, hosting many different actors and structures, each with their own causal powers. Consequently, to reflect this and to explore the impact of one level on another, I chose to situate the study within the conceptual framework of the structure/agency debate, embedded within the current standards-based reform agenda.

There are many sociological approaches for understanding how change is mediated, each bringing forward a means to investigate the reflexive actions of humans in particular contexts. I wanted to hone in on the temporal dimensions of teacher decision-making within the different strata of their everyday interactions to develop an understanding of the mechanism teachers employ to navigate changes in their

working environment over time. Therefore in the analysis of the teacher interview data in chapter 9, I wove together the theorisation of *agency* by Emirbayer and Mische (1998), representing factors which impact teacher classroom decision-making, with the exploration of *discretion* in the depiction of teachers as Lipsky's Street-Level Bureaucrats (2010) that reflected decision-making bounded by the structural rules of the education system. As far as I am aware, this is an analytical perspective that has not been undertaken in previous research and was used in order to characterise the different perspectives on professionalism (Ryder, 2015) which reflect teachers' contractual and responsive actions accountability (Halstead, 1994) in the process of change (Sahlberg, 2010).

With a concurrent nested (Biesta, 2010a) mixed methods study such as this, in which the theorisation is explicitly foregrounded, the quantitative data was secondary to the qualitative data and used to expand the range of inquiry. The NPD data used in the analysis, as a measure of student attainment in science education, provided secondary **evidence *for* reform**; this supported the **evidence *of* reform** gathered from the teacher perspectives. The illustrative theoretical framework provided one possible scaffold with which to generate meaningful understanding of the interactions between reform, student attainment, and teacher practice to answer the question *"How have recent reforms to the science curriculum and its assessment affected student attainment and science teacher classroom practice?"* There is an implicit understanding that reform will affect different groups of students and teachers in different contexts at different times.

Arriving at an understanding of the issues raised in the research question required mixing the quantitative and qualitative strands of the study. This was achieved by applying the two unifying themes which had emerged through the analysis and

interpretation of the findings. The first theme, *equity and fairness* juxtaposed the structural demands of equal opportunity inherent in the standardisation and accountability systems in education against the teacher's understanding of fairness in delivering a science education that responds appropriately to the differing needs of individual students. This runs parallel to the second theme, *time and resources,* which centred on the practicalities of teacher practice constrained by the structural demands of marketisation and the neoliberal agenda. This theme of *time and resources* reflected upon teacher discretionary decision-making powers in the quest to manage the learning of large groups of students with a range of diverse needs in an era where time and resources are stretched.

10.1.2 Outline of the Chapter

In this two-part discussion chapter, I explore the interaction and associations between student attainment data, teacher practice and reform as revealed in this research. The first section of the chapter reflects on the quantitative data analysis and positions student attainment data as evidence *for r*eform. I suggest that this evidence represents what a government chooses to measure; what it considers to be of value in education; what it deems as a valid proxy for learning and what are believed to be reliable indicators of educational success. Thus, in the analysis of historical secondary data, it was possible to explore the relative impact of policy on attainment by making year-on-year comparisons by student background characteristics. It should be noted that correlation of the attainment data with policy change does not imply causation.

The second part of the chapter reflects on the qualitative data analysis and positions my participants' perspectives on their classroom practice as evidence *of* reform. Teachers, among their multiple roles of reform implementor, curriculum interpreter, and professional pedagogue, are pivotal in supporting student attainment and ensuring the success of government policy. Through the qualitative data it was possible to

explore the relative impact of policy on teacher practice by the analysing my participant's interview responses along temporal lines. Although there are many interrelated factors which impact on student attainment and teacher practice, as discussed in the previous chapters, what unites both the student and the teacher are the structural rules in which they operate. As my data indicates, standardisation of the curriculum, expectations and accountability measures serve to move both teacher and student along a clearly defined path.

10.2 Evidence *for* Reform:

The quantitative research question asked, *"What do historical trends in student's end of key stage science attainment since 2008 show and how does this reflect policies to raise attainment?"* For me, this question incorporated and reflected the global and national context of the study by recognising the move toward gathering evidence of "what works" in education (Biesta, 2007, 2010b). My quantitative research question was rooted in my desire to find out whether government rhetoric and discourse (Gove, 2013; DfE and Gove, 2012) on the standards in science education had grounds for truth, and in doing so to explore how government policy was reflected in the attainment of students. This section briefly discusses the quantitative findings, in the context of historical attainment gaps for each student characteristic in turn, interlaced with the teachers' perspectives on the *fairness* of the current reforms.

Overall, the data in chapter 3 showed that attainment in end of key stage 2 teacher assessments increased year-on-year between 2008 and 2015. However, while there was an increase in the number of students awarded a level 4+, this masked the differences between the attainment of different groups of children. Subsequent data, from 2016 to 2018, although not strictly comparable, saw a decline in the percentage of students reaching the expected standards following the introduction of

the new national curriculum in 2014. Student attainment at KS4 shows more variation over the period 2011 to 2018 possibly reflecting the many changes to examinations and qualifications.   Like KS2, attainment gaps between groups of students persisted irrespective of the assessment regimes at KS4. Data drawn from the NPD for the period 2011 to 2016, indicated that in both of the GCSE measures analysed (percentage of students achieving 2 good science GCSEs and percentage students entered for triple science and achieving A*-B in Chem or Physics) the scale of these gaps varied for students with different characteristics.

Looking at attainment gaps between students with different background characteristics, my analysis supports previously reported research. This suggested that whilst the evidence for reform is indisputable, the direction of reform has yet to effectively address the differences. The gender gap is longstanding with girls outperforming boys (DfES, 2007, p. 2), the NPD data analysed in this study affirmed that girls outperformed boys at KS2 and KS4 in certain measures. One participant believed that the new Progress 8 measure might help to close the gender attainment gap as opposed to the previous contextual valued-added system which they thought had "allowed boys to do worse just because they were boys". However, my analysis also found that the attainment gap narrowed for higher achieving students. This was evident at KS2, with little difference between the percentage of male and females attaining a level 5+ and, at GCSE separate sciences, where equal numbers of male and females achieved an A or B in Chemistry or Physics.

Research suggested that boys performed better in high-stakes terminal examinations, whereas girls, possibly more anxious about linear exams, do better in longer-term, coursework-style assessments (Baird *et al.*, 2019; Elwood, 2005). However, the reported effect sizes from my analysis of the differences in attainment indicated that,

both male and female students improved their highest point score in science after reforms to GCSE examinations, although a gap remained and was statistically significant. My findings were consistent with recent research into the impact of GCSE reform (Baird *et al.,* 2019; Burgess and Thomson, 2019) but what my data cannot show and what remains to be seen is whether the new curriculum and the assessment of its content will in some way help close the gender attainment gap. Despite these attainment differences, there has been a poor translation of the academic success of girls in science to uptake in STEM careers (Bramley *et al.,* 2015; Ing, 2014), but one key factor to maintaining girls' interest and attainment in science is strong parental support (Ing, 2014; Perera, 2014). It is difficult to envisage that this source of positive encouragement will change or be hampered by government policy, leading to the tentative conclusion that the gender gap will continue to persist.

Established research shows that poverty and parental socio-economic status (SES) are predictors of student attainment (Jensen, 2013; Chowdry *et al.,* 2011; Gregg and Washbrook, 2011). In my study, the mean attainment gap in the highest point score between students receiving free school meals (FSM) and non-FSM students was the equivalent of approximately one and a half GCSE grades. At KS2, the data indicated that 14% fewer students in receipt of FSM achieved a Teacher Assessment level 4+ in science compared to non-FSM students. Recent research suggested that socio-economic differences produce significant gaps even for academically able girls; where able yet lower SES girls, lag behind equally academically able girls of higher SES by about 3 years (Jerrim, 2017, p. 4). My participants' response to this issue varied depending on the culture of the school. Some teachers made specific changes to the curriculum, homework tasks and lesson planning to address the socio-economic differences by, for instance, removing the model cell homework task from the scheme

of learning. While others used the accountability measures and annual review conversations to drive improvements for FSM or pupil premium students.

My analysis showed that the gap between FSM and non-FSM students achieving 2 good GCSEs had fallen over time, from 28.3% in 2011 to 23.2% in 2016, and that although the impact of reform at KS2 made little difference to improving attainment for FSM students, reform to GCSE examinations appeared to have brought about significant improvements for FSM students. The analysis of effect sizes in chapter 7 (Table 21 and Table 22), shows that mean highest point scores in science increased by more than one GCSE grade after the selected reforms were implemented, while this can be read as a positive step forward, the difference in attainment was still significant. Recent government data indicated that the attainment gaps between disadvantaged students and their peers had widened slightly in several key accountability measures (DfE, 2019b, p. 24) therefore, it may be problematic to suggest that the attainment of FSM students in the current, reformed science GCSEs will continue to improve. It has been suggested that a closer alignment with social reform rather than education reform would be a more effective way to tackle the differences in attainment associated with socio-economic background, poverty and deprivation (Parsons, 2016; Ball, 2010). From my analysis, it would appear that irrespective of the type of reform or policy, those students from economically disadvantaged backgrounds do not perform as well as their peers and this will continue to be the case under the new reforms despite the introduction of pupil premium (Andrews *et al.,* 2017).

Whilst ethnicity is considered to be a contributory factor in determining the outcomes of students, isolating its effect from other factors such as SES and gender is a complex task (Strand, 2015). From my analysis of NPD data from 2008 to 2016, the mean KS2 Teacher Assessment Level for white students was marginally higher than that for students from other ethnic backgrounds. This changed at GCSE, where overall students

from other ethnic groups or with missing ethnicity data achieved higher GCSE point scores in science than their white peers, whether the course was modular or linear, with or without the GCSE equivalents being allowed. However, aggregating the data in this way, as three groups (White, Other and Missing), masks the underlying differences between students from Asian, Black, Chinese and mixed backgrounds. It must be said that the analysis of attainment gaps between different ethnic groups gave mixed results. The effect sizes, which reflected the size of the difference between reform periods for each of the ethnic grouping, indicated that all groups increased their attainment post reform but not uniformly.

Only one of my participants discussed the impact of reform on changes to their practice relative to student ethnicity, when they spoke of students from the Traveller community attending their school. Reasons for this apparent lack of engagement with the issues around ethnically based attainment differences was unclear. I can only speculate that the tracking of student data by ethnicity, while important to government statistics and research into equity in education, plays little or no role in the conscious day-to-day lesson planning of the majority of the teachers in this study. Although, the poor attainment of white working class boys has been well documented (Stahl, 2016; Strand, 2014a) and the low expectations of teachers for their Black-Caribbean boys has been raised (Strand, 2012, 2014a) it is possible, that this lack of overt scrutiny of the outcomes for different ethnic groups at classroom level allows gaps to remain unseen and therefore not addressed. The gaps are persistent and difficult to eradicate through reform to the curriculum, assessment or teaching (Kirby and Cullinane, 2016; Parsons, 2016) and as suggested previously, ethnicity and poverty often intersect (Parsons, 2016; Strand, 2014a) and whilst this is an important factor in determining attainment it is beyond the scope of this study to explore the implications further.

Historically, students with special educational needs have performed less well in terms of their science attainment than their peers (Villanueva *et al.,* 2012) and my various analysis across KS2 and KS4 measures bears this out. Just over 50% of SEN students achieved a teacher assessment level 4+ in science at KS2 as opposed to over 90% of their non-SEN peers. At KS4, the percentage of non-SEN students achieving 2 good GCSEs was almost 5 times that of SEN students; with 8 times more non-SEN students entering triple science and achieved an A*-B grade in chemistry or physics than those identified as SEN. However, SEN students showed the greatest increases in GCSE highest point score in science, improving by more than one grade in both of the reforms explored in this study. But the attainment gap persists (DfE, 2019h) despite changes to the SEN code of practice and the drive toward greater inclusion and improved access to high-quality education (DfE and Department of Health, 2015). The primary school participants in this study, although faced with the increased subject demands, told me that they found ways to extend and assess their less able and SEN students through the use of verbal assessment, recording student responses and noting their engagement in science lessons. The formal assessment used to report progress in science at primary school is based on whether a student has or has not *met expected progress* (DfE, 2018d, p. 9). I suggest, that this relies upon teachers using their pedagogic knowledge to ensure that SEN students are given opportunities to demonstrate their learning. It should be noted, however, that as science is no longer reported as part of the KS2 accountability measures, the drive to show progress in primary science may not be as urgently felt when compared to the secondary phase. As with all students, those with SEN are not a homogeneous group and often present with a range of varied and complex needs. However, Black-Hawkins et al. (2017) suggested that the neo-liberal reforms to education which focus on performance in standardised tests, may be at odds with the aims of inclusive education such that the

lines that determine who receives additional support and who does not remain arbitrary.

The arguments that used student attainment data to justify the review and revisions to the national curriculum and, its assessment and accountability measures were many. These arguments included the so called "dumbing down" of GCSEs (DfE, 2010b); grade inflation due to the opportunities to re-sit modular examinations (Baird et al., 2019); the need to improve literacy at key stage 2 (DfE, 2015c); teaching to the test (Bew, 2011); coasting schools (Roberts and Bolton, 2016); schools "gaming" the system (DfE, 2010b, p. 8) and too few students going on to study science at A-level (Gill and Bell, 2013). Epitomising the negative discourse surrounding education at that time, these arguments, it was suggested, could be associated with the misuse and misinterpretation of attainment data (Mansell, 2013). I argue that the cumulative evidence for reform which preceded the 2014 changes to the national curriculum, its assessment and accountability measures, along with England's poor PISA tests ranking (DfE and Gove, 2012), have generated policy which now shapes student's experience of science in intended and unintended ways.

Recent data points to the success of the reforms aimed at increasing the number of students entered for double science from around 697,000 to almost 741,000 (see Table 4, chapter 3) and the number of students following separate science GCSEs which has increased by around 30% (DfE, 2019b). The introduction of the Progress 8 accountability measure appears to have achieved its intended outcomes, where student progress, from their starting points, is recognised regardless of their final grade. Several secondary teachers remarked that the Progress 8 measure relieved the some of the pressure that existed at GCSE, shifting the focus away from students on the C-D borderline, and helped them focus on students more widely. The unintended consequences resulting from the introduction of the new accountability measures, like

Attainment 8 and Progress 8, and the use of standardised targets, progress matrices (DfE, 2018h) and flight paths in particular, placed teachers under increasing pressures in the classroom. However, as Leckie and Goldstein (2017) pointed out, teachers have operated in the context of punitive accountability frameworks since the introduction of league tables. Therefore, for my participants, the added pressure may not represent a significant change from past experiences. Wilkins (2011) also suggested that younger teachers who had passed through the era where accountability measures, standardisation and targets were the norm, develop their teaching to meet these targets within the rules and citing little justification for doing otherwise.

This section discussed the *evidence for reform* and in choosing to analyse the attainment of students over time, it is presupposed that this knowledge about student's learning constitutes a valid outcome measure of the implementation of education policy. However, this omits the contribution made by other valid measures of the impact of education, such as developing citizenship values, health benefits, social and communication skills (Wolfe and Haveman, 2002; Vila, 2000). Implicit within the question, I argue, is the notion that reform in education is based on tackling an existing, perceived deficit; one which is thought to be a barrier to the achievement of a critical objective or goal. The premise of any major government reform in education is to raise standards and overall attainment for all students (DfE, 2010b) monitoring student attainment data plays a major part in this. Through the mechanisms of standardisation and accountability England legislates the education system to deliver equity and equal opportunity; to close the gap between opportunities available to the rich and poor, increase social mobility and overcome the fatalism brought about by deprivation (DfE, 2010b, p. 7). Consequently, student attainment data is used by the government to shine a light on the outcomes for particular groups of students (Macleod *et al.*, 2015; Sharp *et al.*, 2015), to monitor the

progress towards these goals, to direct additional funding or to design and implement relevant intervention programmes to tackle underachievement (Jarrett *et al.*, 2016). The unintended consequences are that attainment gaps remain even when different assessment models associated with different reforms are taken into account. This, as Parsons (2016) also suggested highlights the need to tackle the underlying structural factors which enable attainment gaps to persist. Reform to education does not take place in a vacuum, it involves a large number of socially connected stakeholders, with teachers being prominent within this mix. Teachers become the object of policies which one could argue standardises and measures the impact of their practice, and which aims to shape them into an image of the "effective" teacher. This next section discusses the responses from my participant interviews and draws upon the overarching theme of *time and resources* to illustrate the impact of reform on teacher practice.

10.3 Evidence *of* Reform

In this part of the chapter, as *evidence of reform,* and linked to the semantic themes driven by the pre-interview questionnaire, I discuss how constraints in time and resources have impacted participant's teaching practice. The focus is on the teaching, learning and assessment activities used, the curriculum content and time available to deliver it and the ways in which teachers have been supported in their engagement with reform.

10.3.1 Time and Resources in the Context of Teaching and Learning Activities

Practical and experimental work, emerged as one of the most important and enjoyable activities in the science classroom making it distinct from other subjects (Toplis and Allen, 2012; Woodley, 2009). Although the effectiveness of practical science to bring

about student conceptual change has been debated (Ofqual, 2015a; Abrahams and Reiss, 2012; Toplis and Allen, 2012), of the 26 teachers that I interviewed, organising practical activities was frequently cited and remains a key feature of their classroom practice. At GCSE, the required practicals dictated the range and scope of the knowledge, skills and understanding of the working scientifically strand of the KS4 curriculum. Yet, secondary teachers explained how this high level of prescription had bled into the activities undertaken by students in KS3, effectively squeezing out opportunities for younger students to take part in more open-ended investigative tasks. On the one hand, the removal of the school-based assessments may help to eradicate the differential interpretations found in coursework and controlled assessment across different schools (Opposs, 2016). On the other hand, participants indicated that, at GCSE, the uncertainty surrounding successful delivery of practical work had introduced a variety of different coping mechanisms and strategies. It also appeared that schools look to each other for support where there was little guidance from the examination boards or central government. Evidently, the new assessment strategy for practical work at GCSE has implications for teacher workload, school budgets and student's experience of experimental science but it was uncertain how impactful these would be (Ofqual, 2015a). I suggest that this area of the science curriculum presents fertile ground for future research.

Resourcing practical activities, experiments and equipment requires considerable investment in time, funding and technician support (Ofqual, 2015a; SCORE, 2013a). These factors arose in discussion with my participants at both primary and secondary level and had ramifications for the different types of teaching and learning activities undertaken. Primary school teachers spoke of the need to spend their own money on purchasing resources for practical activities, perhaps reflecting the suggestion that only 41% of schools have the appropriate budget for science resources (Wellcome

Trust, 2017, p. 6). As with the findings from Blackmore et al. (2018), my teachers had to be creative with the resources available, with one participant explaining how they were able to develop a science lesson as a result of a recent snowfall. However, at times the lack of equipment also meant that teacher demonstrations were carried out more often in place of student hands-on practical work.   In line with the recent Wellcome Trust report (2017), my findings indicated that organising "science week" was a major part of a primary science coordinator's role and an opportunity to liaise with their secondary school colleagues. This block of time immersed students in the subject, filled gaps in their science experience and accommodated longer investigative activities. The availability of sufficient curriculum time for science, especially investigative practical work, was a consistent issue for my participants, particularly in response to the more demanding curriculum content.

The data collected through the pre-interview questionnaire confirmed that much of what the literature considers to be good classroom practice was replicated in my participant's classrooms (Muijs and Reynolds, 2017; Rosenshine, 2012; Vieluf *et al.,* 2012; Hattie, 2009). The participants in primary and secondary school settings used many of the common classroom practices such as using open-ended questions to explore thinking and framed questions to check understanding; whole-class discussions and sharing learning goals. Whilst still regarding activities such as open-ended questions (Chin, 2007), DIRT time (Beere, 2012) and investigative skills development as important features in their lessons, secondary teachers at key stage 4 overwhelmingly responded that the reform to the curriculum and its assessment meant that the time invested in these pedagogical classroom practices was constrained. In their place was a concerted drive towards delivering more teacher-led lessons which enabled the speedy transmission of knowledge to meet the increased cognitive demand of the new specifications. Furthermore, the preparation for the new linear examinations involved

repeated drilling through exam question practice and limiting the focus of practical science to that concerned primarily with the statutory required practicals. Consequently, for many secondary teachers in my study, their interpretation and perception were that the "fun has gone out of key stage 4 lessons" with the new curriculum. This gave the sense that over time, the more engaging activities have been sacrificed in exchange for time to backfill the gaps in curriculum knowledge and exam practice.

The terms "fun" and "rigour" were cited frequently by my respondents in association with changes to their teaching and the students' learning experience. Fun was often allied to activities which had no or low-stakes assessments e.g. Kahoot quizzes, with open-ended practicals or modelling with various materials. The fun activities were educational but also the delivered lesson content in an interesting and engaging way which, it is argued, produced a positive learning experience for students (Shirazi, 2017). Whereas, rigour was linked to the increased level of factual recall of scientific knowledge and cognitive understanding expected of the students. For instance, in learning the physics equations or calculating bond energies in chemistry. For my participants, the combination of increased rigour and increased subject content appeared to have shifted KS4 teaching towards a more transmissive teacher-led style in order to meet the timelines for course completion.

The interviews highlighted a divergence within the approach to the KS3 curriculum and the extent to which the content of the GCSE curriculum was introduced into this earlier key stage. Many of the secondary teachers indicated that they felt it was important to retain the fun elements in the KS3 schemes of learning. These schemes contained time-intensive lesson activities for developing students' thinking and investigation skills and encouraged engagement with research projects and

presentations. Other secondary teachers explained that their KS3 schemes focussed on preparing students for the new GCSE curriculum through the introduction of more rigorous practical activities, mathematical skills and higher order content, e.g. introduction of more formula into year 7. This preparation was often associated with the pursuit of the student's predicted grades, trajectories and "Flight Paths". It has been suggested that introducing GCSE science too early may leave students feeling overburdened trying to cope with new, more complex topics (Shirazi, 2017). Having said this, all of the secondary participants acknowledged that formal study for GCSE began in year 9, thereby shortening the time spent on studying the KS3 curriculum.

10.3.2 Time and Resources in the Context of Curriculum Content and Curriculum Time

From the interviews, it appeared that the differences in the teaching of science between primary schools was reflected in the relative levels of autonomy and professional discretion available with respect to curriculum planning. Several respondents had embraced the new curriculum and the inclusion of more demanding topics such as genetics at key stage 2. Chief among these were coordinators with a science qualification, who had undertaken research into primary science teaching, or whose schools were completing the Primary Science Quality Mark (PSQM) (University of Hertfordshire, 2018). A school leader's confidence in science has a major impact on the level of support received by individual classroom teachers (Wellcome Trust, 2017; White *et al.*, 2015) and where involvement with the PSQM was directed by the vision of the school leadership for the particular schools in my study, it appeared to have enhanced the profile of science across the entire school. Both the Wellcome Trust (2017) and White *et al.,* (2015) report that attaining PSQM gave wide support to primary teacher's ability to teach science and assess the students appropriately, whilst also improving understanding of how to harness appropriate resources for teaching and learning. However, it was also evident from the interviews that respondents relied

heavily upon purchasing resources, such as published teaching schemes, to support lesson planning, practical activities and assessment. Companies like Empiribox, offer a ready-made attractive solution to schools in science which includes teacher training and support three times a year with all the equipment for the experiments every week (Empiribox, 2018). But this comes at a cost of £40 per pupil per year and at a time when school budgets are under considerable strain (Whittaker, 2019), it remains to be seen how long schools can maintain this level of commitment. Although, the presence in the classroom of a TA was deemed a welcome resource by many of my participants their effectiveness can vary (EEF, 2018c).

In line with previous research (Gess-Newsome *et al.,* 2017; Berry and Loughran, 2010), faced with the increased academic demand of the new curriculum, my findings indicated that secondary teachers recognised the need to develop their collective and individual pedagogical content knowledge to support the creation of engaging learning opportunities and classroom activities. In this respect, the secondary teachers planning for KS3, saw themselves moving towards a time when their knowledge of the new curriculum and its assessment was secure and was reflected in the variety of activities used with younger students. Earlier research which looked at science teacher response to reform (Ryder and Banner, 2013); large-scale changes to the national curriculum (Jenkins, 2000); the introduction Earth Science (King, 2001) or the removal of the end of KS2 national testing (Collins *et al.,* 2010) indicated that it takes time for teachers to become familiar with any new curriculum content and how it is best taught across the ability range. Teachers might be teaching outside of their specialism or not fully understand particular concepts themselves, and this adds to the cognitive load in developing practice around the new content (Kind, 2009b, 2009a; Childs and McNicholl, 2007). Confidence building through shared practice requires time and additional professional development resources (Weißenrieder *et al.,* 2015;

Park *et al.,* 2011; Berry and Loughran, 2010) which in the context of these reforms to the curriculum, my participants said was in short supply. Thinking ahead, previous research suggested that it takes about 4 or 5 years for new reforms to become embedded and routinised (Ryder and Banner, 2013), and that progressive cognitive commitments and changes to teacher's epistemologies, as well as language and beliefs, are also required in order to change practice (Wallace and Priestley, 2017). I believe that as the new curriculum becomes embedded, and schools and science departments accommodate and become more at ease with the changes, secondary teachers will establish new ways of working and create wider fields of action.

The number of teaching hours devoted to science in the school week has long and repeatedly been highlighted as an area of concern, with research (Boyle and Bragg, 2005) showing that science was given less time and resources than the other core subjects. In primary schools, time devoted to preparing for the high-stakes national assessments associated with literacy and numeracy ultimately impact on the time that can be spent on science learning (Wellcome Trust, 2014). Other recent reports, Wellcome Trust (2017) and Ofsted (2013), continued to find that primary school science lessons were not daily occurrences and opportunities for inquiry-led teaching were being missed; with 54% of classes not receiving the equivalent of two hours of science per week and only 15% of respondents were happy about the time spent teaching science (Wellcome Trust, 2017, p. 7). The primary schools in my study delivered an average of 1 hour of science a week and what emerged from the participants was that this was not always delivered by the classroom teacher themselves. From my data, it appeared that in some primary schools a qualified teacher was not always deployed to teach science; a Teaching Assistant (TA) takes on the subject in the afternoon during the teacher preparation (PPA) time. The presence of an effective teacher in the classroom has been found to have a very big impact on

student attainment especially for pupils from disadvantaged backgrounds (Sutton Trust, 2011). Whilst this is true for all curriculum subjects, access to the *resource* of a science specialist in primary school classrooms was rare (Cutler, 2015). Without additional support primary teachers and TAs may lack the confidence to deliver the newly introduced more demanding subject knowledge and may be unable to address misconceptions or use inquiry-based methods (Wellcome Trust, 2017; Aalderen-Smeets and Molen, 2015). The Association for Science Education warns schools that "a scheme of work is only as good as the teachers using it" (Hiscock, 2019). I suggest that despite the use of published scheme or science education service provider like Empiribox, the shortage of science expertise may have implications for future CPD, teaching, learning and assessment.

Despite there being no minimum hours required for any national curriculum subject (Roberts, 2018) it appeared that curriculum time for science in secondary schools was as constrained as it was in primary school settings. Although double award science GCSE often accounts for approximately 20% of the curriculum time (Tomei *et al.*, 2015, p. 171), single science GCSE subjects – Biology, Chemistry and Physics are rarely timetabled in a way that gives each subject the 10% of the curriculum time as other single GCSE subjects. Many of the secondary science teachers in the study felt constrained by the lack of curriculum teaching time, particularly with their separate science students. Secondary science teachers remarked that they were faced with making difficult decisions about their teaching; either to devote more time for deep learning to improve student's understanding of difficult concepts or to move onward to ensure that the entire specification is covered in time for the summer examinations. This relates to the previous discussion which explored the equity issues arising from the introduction of the reforms and also echoes the research by Moore and Clarke (2016). Their research found that many teachers believed official policy

was often at odds with their personal beliefs about education. Although not every teacher in their study faced the same difficulties, many felt that the system of public education was "faulty or unfair or as simply serving to perpetuate socio-economic inequalities in the wider world" (2016, p. 667). The teachers in my study maintained that a lack of support in their classes impacted the ways in which they delivered the curriculum to the students in need.

My findings show that the relatively short time-scale for the implementation of the reforms to GCSEs and A-levels placed science staff under additional pressures and generated a high degree of uncertainty. As Clement (2014) suggested, this left teachers feeling rushed, confused and at times cynical as they tried to cope with the consequences of top-down mandated change. My participants discussed how policies were rapidly incorporated across the science curriculum whilst student attainment was being closely monitored by senior management in order to maintain the school position in the education marketplace. The key measures that were once well understood by my secondary teachers now required greater levels of support to untangle. The volume and content of the science knowledge to be covered has increased, along with the cognitive demands bringing added uncertainty to the day-to-day lives of my participants. For example, establishing the difference between a GCSE grade 4 *standard pass*, compared to grade 5 *good pass* against the previous C grade. Similarly, drawing out the nuanced responses which differentiated between grade 8 and the highest grade 9 against the old A* for instance, were highlighted as major areas of concern to the secondary teachers. For secondary teachers in my study, the interactions between multiple yet distinct policies required significant adjustments to their practice and involved multiple expressions of agency. This conclusion reflects that reported by Ryder *et al.,* (2018) in their study with Swedish teachers.

It is well documented that England's science teachers operate within a context that has been subject to many policy changes directly affecting their work (Ryder and Banner, 2013; Millar, 2011; Braun *et al.,* 2010). Although educational change is not uncommon, I suggest that the newest changes represent a greater challenge due to their scope and reach, particularly in secondary school contexts. Analysis from the interview and teacher questionnaire data led me to conclude that accommodating the scale of these reforms introduced new constraints on time and resources available to teachers. This then limited teacher agency through the restriction of pedagogic choices and shut down areas of discretional decision-making in the management of large groups of students. Inevitably, the yearly publication of performance tables means that teachers and schools have nowhere to hide. All this is taking place during a period where resources are stretched across rising pupil numbers (Taylor, 2018) and a continued teacher recruitment and retention crisis (Foster, 2018; Vignoles *et al.,* 2018).

I have used the term *resource* to mean all of the material requirements and embodied representations of teaching. The commodification of science resources includes teaching resources such as textbooks or science equipment as well as human resources, like teaching assistants or technician support. I also conceptualised the notion of *time*, as a particular type of resource. *Time* manifests in this study in many different ways, for example, time spent on a particular topic or different activities; curriculum time for science; time spent working with individual students and time for CPD or collegiate working and as the means through which the temporal cycle of reform is enacted. The experience of the past and aspirations for the future meet to deliver the actions of the here and now so that the material and the temporal resources enable the teacher in the development and enactment of their practice but exist within contextual limitations of the school bureaucracy.

This section discussed the *evidence of reform* and notwithstanding the research that seeks to identify how teachers can and do make a difference to student outcomes (Miller *et al.,* 2017; Kane *et al.,* 2011; Kennedy, 2010), it can be argued that education reform has, to a greater or lesser extent, consequences for the teaching profession in intended and unintended ways. I suggest however, that an understanding of what constitutes "effective" classroom practice in England should also be read in the light of the embeddedness of National Strategy teaching and learning pedagogies aligned with Assessment for Learning models; and with the Teacher Standards, which set out the guidelines on teacher classroom behaviour (DfE, 2013a). This may indicate that over time particular practices appear and disappear from a teacher's repertoire and practices become nested as part of an individual's pedagogical content knowledge (Shulman, 1986). In certain circumstances, these new ways of teaching take on a structural significance as they take their place in the teaching and learning policies of a school (Braun *et al.,* 2010). I suggest that in response to the changes in *time and resources* available to them and, as a direct consequence of recent reforms to the science curriculum and its assessment, the teaching practices of my participants have changed. The final chapter concludes the thesis and draws on teacher agency and discretion and their interaction with the three external, contextual mechanisms of marketisation, standardisation and accountability.

Chapter 11: Conclusion, Limitations and Implications

11.1 Conclusion: Exploring the Dynamic Nature of Educational Reform

This study researched the experiences of 26 teachers in the South East of England and explored their interactions with the reforms to science education introduced in 2014. In doing so, the study contributed to knowledge on how top-down policy reform is understood and implemented by different policy actors, in different school settings. My findings revealed that questioning – open-ended and to check learning; stating learning goals and, carrying out practical activities remain central to science teacher classroom practice. However, the majority of secondary science teachers in the study expressed negative attitudes towards the new style GCSE and the new science curriculum as the increased content and rigour was not deemed appropriate for all students and put the teaching under increased time pressures. Furthermore, the analysis of student attainment data showed that despite the overall increase in attainment for students in the various GCSE science qualifications over time, there remained significant gaps between students from SEN and socio-economically disadvantaged backgrounds and their peers.

The illustrative theoretical framework used double-headed arrows to signify the interactive nature of the relationship between teacher practice, student attainment, and the external mechanisms of marketisation, standardisation and accountability that drive action and outcomes. Teacher's self-efficacy, their beliefs and their experiences are interconnected within this framework linking historic and current student attainment with accountability measures and the bureaucratic organisation of the school context. From these findings I concluded that maintaining a belief that they can influence the future, to yield improvements for students and themselves, is what appeared to sustain participant's present-day efforts; they consistently allied their

ability to adapt to and navigate the recently introduced reforms with student's future attainment.

The premise was that changes in teacher practice in response to policy initiatives provided evidence of the impact *of* reform, resultant student attainment data fed back as evidence *for* reform, particularly if government objectives were not met. Therefore, it appears that the mechanisms of standardisation and datafication used to drive the relationship shifts accountability for student attainment from government to teachers (Winter, 2017), which I suggest commits teachers to regulatory frameworks, limiting their interactions with students to delivering a narrowly defined core curriculum. This does not rule out pockets of resistance (Robertson, 2015) or prevent teachers from attempting to use their discretion to deliver what they believe is the best for their students (Moore and Clarke, 2016; Taylor, 2007). These beliefs stem from a teacher's past experience or through a collaborative, supportive working environment (Bandura, 1994, 2000). However, my findings suggested that enactment of the discretionary decision-making power, once held by secondary participants, was constrained as teachers felt less able to offer students alternative science provision. An overwhelming sense of unfairness was conveyed by the secondary teachers as they judged the impact of these constraints on their less academically able students.

The theorisation applied to the data, positioned discretion and agency as distinct from each other, and contributed to the possible explanations of how change takes place in educational settings over time. My secondary and primary participants also talked of the uncertainty and ambiguity surrounding the assessment focussed elements of the current reforms. The impact of this uncertainty in their day-to-day decision-making within the confines of the standardised curriculum, led teachers to create their "own" rules for managing cohorts of students; something Lipsky (2010) claimed was

synonymous in the theorisation of teachers as Street-Level Bureaucrats. In the classroom, during this period of transition from the familiar to the new, I conclude that teachers used their agency, beliefs and pedagogical content knowledge to conduct their day-to-day practice. They have less "voice" in negotiating the top-down mandated curriculum and accountability regime and relied, instead, on their ability to get through the period of instability and ambiguity. They faced real challenges in the present and projected forward to a time when, collectively, their efforts were rewarded through improved outcomes for them and their students.

Consequently, the more experienced teachers who have "lived" through many previous government reforms in education and would "wait out" the changes, as they relied on their successful track record, felt no need to justify their pedagogic practice. What has been under-theorised in my work is how the younger, less experienced teachers dealt with reform. Maguire *et al.*, (2015) positioned NQTs as being policy dependent actors, using compliance to policy as a coping strategy while Wilkins (2011, 2015) highlighted the inculcation of early careers teachers into the current performative environment of education, to which they have no other point of reference. For all social actors, it is argued that particular behaviours remain in place until a better understanding of the new situation had been constructed by the individual or socially co-constructed within a group (Bandura, 2000). In this dynamic process, new models of working are tried, tested and evaluated, incorporating new ideas by building on what has gone before. This seemingly simple outline understates the reality of the complex process in which teachers engagement in educational change involves professional efficacy beliefs, a willingness to adopt and adapt, and is rooted in layers of contextual factors (Ryder *et al.*, 2018; Pyhältö *et al.*, 2014).

I positioned the quantitative data as providing the evidence for the reforms which guided the Conservative government's 2010 education agenda (DfE, 2010b). Although the research did not set out to look in detail, interacting with the NPD data has allowed me to take a glimpse at the impact of reform on students with different background characteristics. It is therefore, rather disappointing to see that the conditions discussed by Ball (2010) have not changed in that the unrelenting focus on student attainment in a narrowly defined curriculum has failed to bring about the improvements so keenly sought through reform. The research study was conducted during a period of transition with the numerous reforms to the curriculum and examinations and accountability measures yet to be ingrained and incorporated into the daily lexicon of teacher's lives. A whole cohort of children are still to pass from key stage 2 to key stage 4, therefore, the full impact of the reform is unknown.

Chapter 2 of this thesis argued that the mechanisms used to create change include the standardisation of the curriculum, standardisation of teaching and assessment; the selection of particular attainment measures as key accountability targets; and the introduction of competition to create a market-driven ideology framed in the guise of parent choice (Rudd and Goodson, 2017; Apple, 2005; Hargreaves *et al.*, 2002). These factors have been at the forefront of England's education landscape since the introduction of the Education Reform Act in 1988 (Acquah, 2013). My research has found that these mechanisms manifest differently for science in primary settings. Although primary teachers are required to teach science and report on student attainment, the lack of punitive accountability measures associated with science in this phase means that the lines between personal agency and professional discretion in decision-making are more blurred. Therefore, I put forward that the balance for primary teachers lies between finding the time, resources and expertise to teach science against devoting more time to literacy and numeracy.

From the perspective of a former science teacher, I may be somewhat biased in the conclusion that there are positive implications for the increased focus on science. In that it cements the central role of science as a core subject in the curriculum; it allows a great number of students the opportunity to engage with the cognitively demanding areas of knowledge, skills and understanding within the science curriculum, thereby enabling more students to consider studying science in their post-compulsory education. Looking more globally, the increased demands in the science curriculum introduced through the new programmes of study may also be seen as a push towards improving the place that England holds in the PISA international rankings. It could be argued that if students attain highly in the examinations in the more rigorous English science curriculum, then they are more likely to perform well in the international tests. Thus, sending an outward message that England's education system is comparable to any high achieving nation state and is capable of producing a skilled labour force to meet the needs of the competitive and increasingly technological world.

Next, I present a summary of the main limitations which have been previously discussed in the methodology and throughout the findings chapters. I then discuss the implications of this research for teachers, policymakers and other stakeholders before proposing ideas for further research and indicating what new knowledge remains to be explored. The thesis then ends with a reflection of my final thoughts on the outcomes and the research process.

11.2 Limitations

11.2.1 Limitations to the Qualitative Study

The outcomes from this qualitative study like any other, are subject to the decisions taken by the researcher throughout the process (Creswell, 2018; Cohen *et al.*, 2017) and the validity of the qualitative study relied upon confirming a valid description of what I set out to do. The qualitative study aimed to uncover knowledge and meaning about how government policy in science education impacted on teacher classroom practice. The key to the new knowledge is an understanding that the impact on teacher practice is *relative*, and through a temporal lens, related to teachers' beliefs in their ability to adapt to changes. This perspective acknowledges that participants will interpret the changes in different ways (Creswell, 2018). Opting to take a philosophical stance which reflects my beliefs as a researcher and choosing to use particular theoretical and conceptual lenses to analyse the data undoubtedly limited other avenues of inquiry and certainly, the arguments developed in this study may be open to alternative interpretation should different theories and concepts be applied. In discussing the specific limitations, this section draws together the areas in which the methodological and data-driven decisions may have impacted the knowledge generation.

In 2017, there were approximately 498,220 contracted teachers in state-funded schools in England (DfE, 2019i), therefore this research cannot presume to represent the varied perspectives and lived experiences of each practitioner. And whilst all teachers in state schools are subject to the control mechanism of the Teacher Standards document (DfE, 2013a), as individuals it is supposed that they enact these guidelines in ways which fit their context. Although it was not the original intention, due to the high degree of educational standardisation and the depth and breadth of the accountability frameworks, there is a case for suggesting that the findings from

this study reveal features which show a degree of generalisability (De Vaus, 2014) and, on reflection, the modification of several specific features of the study may have yielded additional insights. For example, increasing the sample size may have ensured that data saturation was fully achieved (O'Reilly and Parker, 2013), whilst widening the range of teachers interviewed to include a greater number of year 6 teachers could support the exploration of classroom practice during the final year of KS2; a time at which the pressure to prepare students for the end of key stage examinations is at its greatest.

In an attempt to reduce any potential bias, the content of the pre-interview questionnaire originated from the validated TALIS Teacher Survey (OECD, 2013), consequently, no distinctions were made between the secondary and primary school context. From the responses to the electronic questionnaire and interview, it was evident that specific questions were more relevant to the secondary teachers than the primary teachers, meaning that the interpretations may have been biased towards the changes occurring in secondary education. However, the follow-up interviews provided the means to explore the differences and issues in a more in-depth and nuanced way. The methodology of the study did not include data collection via lesson observations with my participants. Observing teachers in action, although still open to researcher interpretation (Wragg, 2011) may have introduced a degree of confirmatory evidence to the qualitative findings, adding reliability to the data collected from the pre-interview questionnaire (Greene *et al.,* 1989). However, I suspect that incorporation of lesson observations could also have reduced the number of teachers who were willing to take part in the study due to the increased demand on time and a possible increase in stress and anxiety on the participants (Cohen *et al.,* 2017), thus undermining the ethical principles (BERA, 2018) I had agreed to abide by.

11.2.2 Limitations to the Quantitative Study

Using the NPD provided the opportunity to carry out statistical analysis on more than 10 years of student attainment data, it also presented some limitations and challenges. As with all secondary quantitative data, although it was derived from a trusted source it was not free of error (Smith, 2008; Gorard, 2002). The complexity of the application process meant that the data from two cohorts of KS2 students was not obtained (years 2012/2013 and 2013/2014). This meant that all years for the period under investigation were not covered in detail. Furthermore, the reform to data protection laws (Defend Digital Me, 2018; DfE, 2018e) meant that applying for additional data for the years 2017 to 2018 to supplement the inquiry proved problematic, and required a more involved application process than undertaken previously (DfE, 2018f). This reflected the evolving nature and complexities in using big datasets as they are often far removed from the original people from whom the data was collected, therefore their rights can easily be overlooked in the secondary data research process. I decided not to go ahead with an application and relied upon the aggregate data published by the DfE. Although this did not hamper the analysis, it meant that any unusual changes to the data could have been missed. Over the years, the definitions and codes used to describe key student variables have changed (see Appendix Table H, p359). This was particularly relevant to the introduction of the new SEN Code of Practice where the designation relating to School Action plus was removed (DfE, 2015; DfE and Department of Health, 2015) meaning that data for year 2014/2015 on the number of students with SEN statements action plus was limited.

Ensuring that the data was as accurate as possible was one of the challenges faced during the research process. The change from whole cohort reporting to a 5% sample, of students examined in science at the end of KS2, introduced different methodologies and quantification of the outcomes for students, making year on year comparisons

between different groups of students problematic. Similarly, within the NPD, there is a limitation in the way that the different GCSE examination specifications are reflected (Gorard, 2012). In that, due to the many changes to GCSE qualifications and their equivalents, it was not possible to make a full comparison based on the types of assessment and examinations in force over the period studied. It was difficult to isolate the effects of the changes to the practical assessments, namely: coursework, Investigative Skills Assessments (ISA), controlled assessments and the recently introduced Required Practicals from the other assessment changes, namely: modular vs linear, GCSE equivalents, new curriculum and new grading 9 to 1 (Baird *et al.*, 2013). Furthermore, the precise examination format followed by each student could not be determined, and while the option to follow a modular route through science GCSE was available to all schools prior to 2012, science departments may not have followed this route and were free to select the most appropriate assessment strategy for their school context.

Statistical modelling can help to predict attainment and account for student characteristics such as gender, ethnicity and socio-economic status (Hamlyn *et al.*, 2017) but it is more challenging for a study such as this to account for the impact of other external factors or unplanned interruptions such as a teacher's strike (Jaume and Willén, 2017; Wills, 2014) or examination boycott (Busby, 2018; Turner, 2017); the poor quality of marking for the tests themselves (Ofqual, 2015c), and the room temperature when sitting an exam (Park, 2017). Data which examined the progress of students beyond compulsory education and into the A-level sciences was requested but not used in this study and therefore the impact of government reform on the STEM pipeline did not go beyond exploring the numbers of students who follow and succeed in the separate science route. Despite this, using the NPD allowed the analysis of

student attainment for whole cohorts, meaning that the results of my study can be justifiably generalised more widely.

11.3 Implications of this Research

The findings from this research may have implications for teachers and wider education stakeholders. Accordingly, I maintain, that this research has highlighted a possible need for further investment in time and resources to support teacher pedagogic subject and content knowledge. Unlike the era of the National Strategies which ushered in a system-wide change to teaching, learning and assessment - that was funded by central government and supported by a networked structure of consultants with training for schools and teachers (DfE, 2011)- the reforms in 2014, although widespread, appeared not to have attracted an equivalent investment level from central government to those of the Strategies (National College for School Leadership, 2013). Whilst it is unlikely that the support for policy implementation on the scale of the national strategies will occur again it was evident from the research that teachers required more time to develop their understanding of the new curriculum and to address the issues relating to the teaching of the more rigorous content to a wider ability range of students. The results from this research implied that there is a continuing, possibly increasing role for subject associations such as the Royal Society of Biology, Royal Society of Chemistry and the Institute of Physics to provide additional CPD for teachers. In addition to forging connections with the learned bodies, the possible involvement of STEM learning centres (STEM Learning, 2018) could give schools the support required to ensure that science teaching not only remains relevant but is also at the cutting edge of new scientific knowledge and ideas. This may involve the loan of specialised resources beyond the normal school budget or include the opportunity for teachers to work with scientific researchers or with a range of different employers from the science and technology field (STEM

Learning, 2018). The cost of having teachers out of the classroom may appear prohibitive, but, this additional time would represent an investment in staff development and a welcome re-professionalisation of teachers. Acknowledging that the success of any reform is a two-way process of interaction (Braun *et al.,* 2010, 2011), I suggest that the money invested in staff training could be drawn from that spent on purchasing electronic or published schemes of work. Furthermore, giving teachers greater involvement with the decision-making on the direction of their professional learning with the opportunity to reflect on and adapt their practice may have long-term gains for student progress, teacher pedagogical content knowledge and teacher self-efficacy (Murphy *et al.,* 2015; Pedder and Opfer, 2011; Guskey, 2002).

My research also suggested that a more formal structure of support, with allocated time for collaborative working between and within primary and secondary schools over and above the occasional transition days or moderation meetings, may be necessary. This could take the form of curriculum planning for student progress involving teachers of science from different phases working together over the coming 4 to 5 years to better integrate the teaching and learning in science across the age and ability range, thereby creating communities of practice (Chandler-Grevatt, 2012). Similar conclusions had already been drawn (Royal Society, 2010) but I believe that the argument for creating better links between schools remains strong, only now with a greater imperative.

As previously mentioned, the new science curriculum significantly increased the level of detail and scientific knowledge required at GCSE with greater inclusion of mathematical skills, ultimately to stretch the most academically able students and encourage further study (DfE and Gove, 2014). My research suggested that there are implications for cross-curricular working between maths and science departments in

secondary schools. With an aim to develop better awareness of the mathematical and cognitive skills needed to address the requirements of the new science curriculum and to reflect this in the sequencing of lessons in schemes of work and long-term curriculum planning. Moreover, there are also implications for initial teacher education in ensuring that University ITE and school-centred initial teacher training (SCITT) courses take account of the increased demand of the science curriculum especially at primary level. Currently, primary school teachers are only required to have a science qualification at GCSE (National Careers Service, 2019), and although there are subject knowledge enhancement (SKE) courses for primary maths there no pre-ITE training courses for primary science, unless offered by an individual ITE programme (DfE, 2019j). Few primary teachers have an A-level in a science subject (Wellcome Trust, 2017) and as non-subject specialists teaching all curriculum subjects (Royal Society, 2010) their ITE training in science may be restricted by the time available on the course. A case can be made for the restructure of primary education ITE to include more focus on developing primary practitioners' pedagogy and subject knowledge in science. Additionally, a review of the curriculum content of existing SKE and science teacher education courses, school- or university-based, could ensure that the more demanding topics are covered in sufficient depth to build trainee confidence. A focus on differentiation and adapting the teaching for the needs of all pupils, regardless of ability, may also ensure that the more academically able students benefit from engagement with the more cognitively demanding work, and the less academically able students remain sufficiently engaged with the material to prevent disruptive behaviour.

One further implication from this research relates to the impact of the new accountability measures on the number of students who pursue science post-16. The EBacc accountability measure ensures that a greater number of students follow the

academic GCSE route towards science qualifications (DfE, 2019a, Parameshwaran and Thomson, 2015) and its introduction has increased the number of students entered for GCSE Combined Science (DfE, 2019a). However, research also suggested that while students enjoy science, they do not often aspire to be scientists or continue the study of sciences post-16 (IoP, 2018; DeWitt and Archer, 2015). To translate the benefits of reform into visible increases in young people's engagement with STEM careers, government policy should aim to develop improved access to careers advice and guidance. This is particularly important for students from disadvantaged or ethnic minority backgrounds, who are underrepresented in the field (Mcmaster, 2017; Archer *et al.,* 2015). The choice to continue the study of science post-16 is a complex one and there are many factors that influence a student's decision, one of which is how enjoyable and fun the subject is at KS3 (Bennett *et al.,* 2013; Hampden-Thompson *et al.,* 2011). The increased rigour of the new curriculum appears to have removed the "fun" aspects from science, breaking its associations with enjoyable activities (Shirazi, 2017). Therefore, the policy which makes an already difficult subject even more challenging and by closing off alternative routes to vocational scientific careers may embed negative perceptions of science as being a really difficult subject and therefore put more students off studying it than previously (Ofqual, 2015c, 2015b).

The reality is that a substantial minority of less academically able students take no science qualifications at all (DfE, 2019a) so effective policy work needs to be undertaken to address this. The implication is that rather than the expected increase in the numbers of students going forward to study science post-16; the numbers may instead decline. A policy review which looks at a possible return of vocational qualifications pertinent to careers in science may be a way forward to widen the involvement of all students.

11.4 Further Research

An underlying theme of this study was time, and the temporal nature of agency in the face of change. The study was situated at the early stages of the changes to the science curriculum and its assessment, as of February 2020, only two cohorts of students have taken the new, more rigorous GCSE examinations. Teaching and assessment of the new curriculum, introduced in 2014 (Roberts, 2018), has yet to feed through from KS1. Consequently, I suggest that there is a real need to conduct further research with a larger sample of science teaching professionals within the next 5 years to revisit the associated themes, time and resources, and equity and fairness. This would offer a further opportunity to explore any long-term embedded changes to teaching practice. I also contend that a deeper theoretical understanding of how teachers balance their decision-making in the face of curriculum constraint, against their beliefs related to science and science pedagogical practice, is needed to illustrate why certain teachers respond to policy change in particular ways and other do not. This may provide a continued reflection on the impact of standardisation of practice and the de/professionalisation of teachers.

The routine collection, tracking and publication of student attainment data in science will inevitably continue whether by government, learned societies or international surveys. I suggest that this would benefit from the inclusions of a widespread cohort survey which details the post-16 destinations and study options taken up by the students from the 2018 KS4 cohort over the next 5 years. This may provide some evidence of the impact of reform on the STEM pipeline. Additionally, further research to examine the extent to which the attainment gaps have narrowed, could be used to build on the research by Allen and Thomson (2016) to look at the extent to which the EBacc accountability measure has supported the progress of disadvantaged or less able students.

11.5 Final Thoughts

This research study aimed to shed light on teacher practice during a period of transition, a point at which reform to science education was being enacted. To do so, the evidence *for* reform was sought through the analysis of student attainment data and evidence *of* reform was gained through talking to teachers about their classroom practice. The complexity of the field lies in its embeddedness in a global context, which pits country against country in a bid to be the "best" in the international rankings; and which relies on policy borrowing from a variety of different jurisdictions to replicate what is perceived as good practice, often failing to acknowledge the importance of the cultural and structural differences that have a major influence in the way in which a policy is experienced. As a contribution to the field, the theoretical framing included insights into teachers' responses to education reform from a structural, rules-based context in addition to an individual agency perspective. The study uncovered teacher responses to the 2014 science curriculum reforms in terms of the impact on their classroom practice and the time and resources available; the "translation" work required to bring better understanding of new policies processes and, the constraints imposed on their ability to deliver an alternative, enriching science experience to their students.

Ultimately, what emerged was the depth to which England's education system relied on the standardisation of the work of its teachers, the curriculum and assessment. Individual teacher actors navigated this world on stratified levels encompassing the personal, the school collective, the national and international; a more in-depth and wide-reaching study could begin to do justice to the intricacies of this multifaceted, interconnected world. Drawing together more nuanced understandings of teachers' approach to reform over time and revisiting the participants to gain a picture of what

has been incorporated as part of their beliefs or pedagogical content knowledge; what they feel they can change through their agency and influence and how they feel they have been constrained to work within the rules of the imposed accountability frameworks.

In drawing together this work, it is inevitable that particular aspects have been explored in more detail than others and that some questions remain unanswered, whilst others are yet to be uncovered. However, the responses of the participants in this study have provided a unique glimpse into how reform to science education, for secondary and primary schools in England, has impacted on teachers' classroom practice.

Bibliography

Aalderen-Smeets, S. I. van and Molen, J. H. W. van der (2015) 'Improving Primary Teachers' Attitudes Toward Science by Attitude-Focused Professional Development', *Journal of Research in Science Teaching*, vol. 52, no. 5, pp. 710–734.

Abell, S. K. (2008) 'Twenty Years Later: Does Pedagogical Content Knowledge Remain a Useful Idea?', *International Journal of Science Education*, vol. 30, no. 10, pp. 1405–1416.

Abrahams, I. and Millar, R. (2008) 'Does Practical Work Really Work? A Study of the Effectiveness of Practical Work as a Teaching and Learning Method in School Science', *International Journal of Science Education*, vol. 30, no. 14, pp. 1945–1969.

Abrahams, I. and Reiss, M. J. (2012) 'Practical Work: Its Effectiveness in Primary and Secondary Schools in England', *Journal of Research in Science Teaching*, vol. 49, no. 8, pp. 1035–1055.

Abrahams, I., Reiss, M. J. and Sharpe, R. (2014) 'The Impact of the "getting Practical: Improving Practical Work in Science" Continuing Professional Development Programme on Teachers' Ideas and Practice in Science Practical Work', *Research in Science & Technological Education*, vol. 32, no. 3, pp. 263–280.

Abrahams, I., Reiss, M. J. and Sharpe, R. M. (2013) 'The Assessment of Practical Work in School Science', *Studies in Science Education*, vol. 49, no. 2, pp. 209–251.

Acquah, D. (2013) 'School Accountability in England: Past, Present and Future', Centre for Education Research and Policy. AQA, Manchester, UK.

Adami, S. (2014) 'Street-Level Bureaucrats: Role and Dimension in the Policy Arena', [Online]. Available at https://steveadami.myefolio.com/Uploads/Street-Level%20Bureaucracy.pdf.

Akiba, M. and Liang, G. (2016) 'Effects of Teacher Professional Learning Activities on Student Achievement Growth.', *The Journal of Educational Research*, vol. 109, no. 1, pp. 99–110.

Albergaria-Almeida, P. (2010a) 'Questioning Patterns and Teaching Strategies in Secondary Education', *Procedia-Social and Behavioral Sciences*, vol. 2, no. 2, pp. 751–756.

Albergaria-Almeida, P. (2010b) 'Classroom Questioning: Teachers' Perceptions and Practices', *Procedia - Social and Behavioral Sciences*, Innovation and Creativity in Education, vol. 2, no. 2, pp. 305–309.

Alexander, R. (2012a) 'Moral Panic, Miracle Cures and Educational Policy: What Can We Really Learn from International Comparison?', *Scottish Educational Review*, pp. 4–21.

Alexander, R. (2012b) 'Pedagogy, Culture and the Power of Comparison', in Daniels, H., Lauder, H., and Porter, J. (eds), *Educational Theories, Cultures and Learning: A Critical Perspective*, Routledge, pp. 10–26.

Alexander, R. (2014) 'Evidence, Policy and the Reform of Primary Education: A Cautionary Tale', *Forum*, vol. 3, no. 56, pp. 349–375.

Alexander, R. and Armstrong, M. (2010) *Children, Their World, Their Education: Final Report and Recommendations of the Cambridge Primary Review*, Taylor & Francis US.

Allen, M. (2017) 'Interviews, Recording and Transcribing', in *The SAGE Encyclopedia of Communication Research Methods*, California, United States, SAGE Publications, Inc, pp. 800–803.

Allen, R., Burgess, S. and McKenna, L. (2014) *School Performance and Parental Choice of School: Secondary Data Analysis*, DfE.

Allen, R. and Thompson, D. (2016) *Changing the Subject: How Are the Ebacc and Attainment 8 Reforms Changing Results?*, Changing the Subject, Research Brief, The Sutton Trust.

Almeida, P. and Neri de Souza, F. (2010) 'Questioning Profiles in Secondary Science Classrooms', *International Journal of Learning and Change*, vol. 4, no. 3, pp. 237–251.

Altrichter, H. and Kemethofer, D. (2015) 'Does Accountability Pressure Through School Inspections Promote School Improvement?', *School Effectiveness and School Improvement*, vol. 26, no. 1, pp. 32–56.

Anderson, J. A. (2005) *Accountability in Education*, Paris, International Institute for Educational Planning; International Academy of Education.

Anderson, K. J. (2012) 'Science Education and Test-Based Accountability: Reviewing Their Relationship and Exploring Implications for Future Policy', *Science Education*, vol. 96, no. 1, pp. 104–129.

Anderson, M. (2017) 'Transformational Leadership in Education: A Review of Existing Literature', *International Social Science Review*, vol. 93, no. 1, p. 15.

Andrews, J., Robinson, D. and Hutchinson, J. (2017) *Closing the Gap? Trends in Educational Attainment and Disadvantage*, London, Education Policy Institute.

Apostolou, A. and Koulaidis, V. (2010) 'Epistemology and Science Education: A Study of Epistemological Views of Teachers', *Research in Science & Technological Education*, vol. 28, no. 2, pp. 149–166.

Apple, M. W. (2005) 'Education, Markets, and an Audit Culture', *Critical Quarterly*, vol. 47, no. 1–2, pp. 11–29.

Apple, M. W. (2012) *Education and Power*, New York and London, Routledge.

AQA (2019) *GCSE Combined Science: Trilogy Specification* [Online]. Available at https://filestore.aqa.org.uk/resources/science/specifications/AQA-8464-SP-2016.PDF (Accessed 9 June 2020).

Archer, L. (2015) *Increasing the Uptake of Science in Schools*, Evidence Briefing, Swindon, ESRC.

Archer, L., Dewitt, J. and Osborne, J. (2015) 'Is Science for Us? Black Students' and Parents' Views of Science and Science Careers', *Science Education*, vol. 99, no. 2, pp. 199–237.

Archer, L., Moote, J., Francis, B., DeWitt, J. and Yeomans, L. (2017) 'Stratifying Science: A Bourdieusian Analysis of Student Views and Experiences of School Selective Practices in Relation to "Triple Science" at KS4 in England', *Research Papers in Education*, vol. 32, no. 3, pp. 296–315.

Archer, M. S. (1995) *Realist Social Theory: The Morphogenetic Approach*, Cambridge, Cambridge University Press.

Archer, M. S. (2003) *Structure, Agency and the Internal Conversation*, Cambridge, Cambridge University Press.

Archer, M. S. (2010) 'Morphogenesis Versus Structuration: On Combining Structure and Action', *The British Journal of Sociology*, vol. 61, pp. 225–252.

Ashton, P. (2015) 'Historical Overview and Theoretical Perspectives of Research on Teachers' Beliefs', in Fives, H. and Gill, M. G. (eds), *International Handbook of Research on Teachers' Beliefs*, London, Routledge, pp. 31–47.

Avalos, B. (2011) 'Teacher Professional Development in Teaching and Teacher Education Over Ten Years.', *Teaching and Teacher Education*, vol. 27, no. 1, pp. 10–20.

Baird, J.-A., Ahmed, A., Hopfenbeck, T., Brown, C. and Elliott, V. (2013) *Research Evidence Relating to Proposals for Reform of the GCSE*, Centre for Educational Assessment Report, Oxford, Oxford University.

Baird, J.-A., Caro, D., Elliott, V., El Masri, Y., Ingram, J., Isaacs, T., de Moira, A., Randhawa, A., Stobart, G. and Meadows, M. (2019) *Examination Reform: Impact of Linear and Modular Examinations at GCSE*, London, Ofqual.

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J. and Shepard, L. A. (2010) *Problems with the Use of Student Test Scores to Evaluate Teachers*, London, Economic Policy Institute.

Ball, S. (2004) 'Education for Sale! The Commodification of Everything?' London, University of London.

Ball, S. J. (2003) 'The Teacher's Soul and the Terrors of Performativity', *Journal of Education Policy*, vol. 18, no. 2, pp. 215–228.

Ball, S. J. (2007) *Education Plc: Understanding Private Sector Participation in Public Sector Education*, London; New York, Routledge.

Ball, S. J. (2009) 'Privatising education, privatising education policy, privatising educational research: network governance and the "competition state"', *Journal of Education Policy*, vol. 24, no. 1, pp. 83–99.

Ball, S. J. (2010) 'New Class Inequalities in Education: Why Education Policy May Be Looking in the Wrong Place! Education Policy, Civil Society and Social Class', *International Journal of Sociology and Social Policy*, vol. 30, no. 3/4, pp. 155–166.

Ball, S. J. (2015) 'What is policy? 21 years later: reflections on the possibilities of policy research', *Discourse: Studies in the Cultural Politics of Education*, vol. 36, no. 3, pp. 306–313.

Ball, S. J. (2016) 'Neoliberal Education? Confronting the Slouching Beast', *Policy Futures in Education*, vol. 14, no. 8, pp. 1046–1059.

Ball, S. J., Maguire, M., Braun, A. and Hoskins, K. (2011a) 'Policy Actors: Doing Policy Work in Schools', *Discourse: Studies in the Cultural Politics of Education*, vol. 32, no. 4, pp. 625–639.

Ball, S. J., Maguire, M., Braun, A. and Hoskins, K. (2011b) 'Policy Subjects and Policy Actors in Schools: Some Necessary but Insufficient Analyses', *Discourse: Studies in the Cultural Politics of Education*, vol. 32, no. 4, pp. 611–624.

Bandura, A. (1977) 'Self-Efficacy: Toward a Unifying Theory of Behavioral Change.', *Psychological review*, vol. 84, no. 2, p. 191.

Bandura, A. (1986) *Social Foundation of Thought and Action: A Social-Cognitive Theory*, Englewood Cliffs, NJ., Prentice Hall.

Bandura, A. (1994) 'Self-efficacy', in Ramachaudran, V. S. (ed), *Encyclopedia of Human Behavior*, New York, Academic Press, vol. 4, pp. 71–81.

Bandura, A. (2000) 'Exercise of Human Agency Through Collective Efficacy', *Current Directions in Psychological Science*, vol. 9, no. 3, pp. 75–78.

Barker, B. (2008) 'School Reform Policy in England Since 1988: Relentless Pursuit of the Unattainable', *Journal of Education Policy*, vol. 23, no. 6, pp. 669–683.

Barkley, E. F., Cross, K. P. and Major, C. H. (2014) *Collaborative Learning Techniques: A Handbook for College Faculty*, San Francisco, John Wiley & Sons.

Bartlett, W. (1993) 'Quasi-Markets and Educational Reforms', in Grand, J. L. and Bartlett, W. (eds), *Quasi-Markets and Social Policy*, London, Palgrave Macmillan UK, pp. 125–153.

Bartos, S. A. and Lederman, N. G. (2014) 'Teachers' Knowledge Structures for Nature of Science and Scientific Inquiry: Conceptions and Classroom Practice', *Journal of Research in Science Teaching*, vol. 51, no. 9, pp. 1150–1184.

Bathmaker, A. M. (2010) 'Dealing with Data Analysis', in *The Routledge Doctoral Student's Companion: Getting to Grips with Research in Education and the Social Sciences*, pp. 200-212.

Baum, S., Cunningham, A. and Tanenbaum, C. (2015) 'Educational Attainment: Understanding the Data', *Change: The Magazine of Higher Learning*, vol. 47, no. 6, pp. 18-27.

Beere, J. (2012) *The Perfect Ofsted Lesson-Revised and Updated*, Carmarthen, Wales, Crown House Publishing.

Bell, D. and Skiebe-Corrette, P. (2016) 'Science Education and Economic Growth: Some Implications for Curriculum, Pedagogy and Assessment', in Wyse, D., Hayward, L., and Pandya, J. (eds), *The SAGE Handbook of Curriculum, Pedagogy and Assessment: Two Volume Set*, London, SAGE Publications Ltd, pp. 470–486.

Bennett, J., Holman, J., Lubben, F., Nicolson, P. and Otter, C. (2005) 'Science in context: The Salters approach', in Nentwig, P. and Waddington, D. (eds), *Making It Relevant: Context Based Learning of Science*, Münster, Germany, Waxmann Verlag, pp. 121–154.

Bennett, J., Lubben, F. and Hampden-Thompson, G. (2013) 'Schools That Make a Difference to Post-Compulsory Uptake of Physical Science Subjects: Some comparative case studies in England', *International Journal of Science Education*, vol. 35, no. 4, pp. 663–689.

BERA (2018) 'BERA Ethical Guidelines for Educational Research 4th Edition 2018', [Online]. Available at https://www.bera.ac.uk/researchers-

resources/publications/ethical-guidelines-for-educational-research-2018 (Accessed 26 June 2018).

Bernstein, B. (1975) 'Class and Pedagogies: Visible and Invisible', *Educational Studies*, vol. 1, no. 1, pp. 23–41.

Bernstein, B. (2000) *Pedagogy, Symbolic Control, and Identity: Theory, Research, Critique*, Lanham; Oxford, Rowman & Littlefield.

Bernstein, B. (2004) 'The Structuring of Pedagogic Discourse, Volume IV: Class, Codes and Control', in London, Routledge.

Berry, A. and Loughran, J. (2010) 'What Do We Know About Effective CPD for Developing Science Teachers' Pedagogical Content Knowledge?', National Science Learning Centre, York.

Bevan, Y., Brighouse T, Mills G, Rose J and Smith M (2009) *Report of the Expert Group on Assessment*, London, Department for Children, Schools and Families.

Bew, P. (2011) *Independent Review of Key Stage 2 Testing, Assessment and Accountability*, London, DfE.

Bhaskar, R. (2008) *A Realist Theory of Science*, Classical Texts in Critical Realism, London, Routledge.

Biesta, G. (2007) 'Why "What Works" Won't Work: Evidence-Based Practice and the Democratic Deficit in Educational Research', *Educational Theory*, vol. 57, no. 1, pp. 1–22.

Biesta, G. (2010a) 'Pragmatism and the Philosophical Foundations of Mixed Methods Research', in *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, California, SAGE Publications, pp. 95–118.

Biesta, G. (2010b) 'Why "What Works" Still Won't Work: From Evidence-Based Education to Value-Based Education', *Studies in Philosophy and Education*, vol. 29, no. 5, pp. 491–503.

Biesta, G., Priestley, M. and Robinson, S. (2015) 'The Role of Beliefs in Teacher Agency', *Teachers and Teaching*, vol. 21, no. 6, pp. 624–640.

Biesta, G. and Tedder, M. (2006) *How Is Agency Possible? Towards an Ecological Understanding of Agency-as-Achievement*, Learning Lives: Learning, Identity and Agency in the Life-Course, Working Paper Series, Exeter, ESRC.

Billimoria, J. (2017) *Is STEM Education All It's Cracked up to Be?* [Online]. Available at https://www.weforum.org/agenda/2017/05/is-stem-education-all-it-s-cracked-up-to-be/ (Accessed 8 March 2019).

Black, P. (1995) '1987 to 1995 — The Struggle to Formulate a National Curriculum for Science in England and Wales', *Studies in Science Education*, Routledge, vol. 26, no. 1, pp. 159–188.

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2004) 'Working Inside the Black Box: Assessment for Learning in the Classroom', *Phi delta kappan*, vol. 86, no. 1, pp. 8–21.

Black, P. and Wiliam, D. (1998) 'Assessment and Classroom Learning', *Assessment in Education: Principles, Policy & Practice*, vol. 5, no. 1, pp. 7–74.

Black, P. and Wiliam, D. (2010) 'Inside the Black Box: Raising Standards through Classroom Assessment', *Phi Delta Kappan*, vol. 92, no. 1, pp. 81–90.

Black-Hawkins, K., Florian, L. and Rouse, M. (2017) *Achievement and Inclusion in Schools*, 2nd edn, Abingdon, England, Routledge.

Blackmore, K., Howard, C. and Kington, A. (2018) 'Trainee Teachers' Experience of Primary Science Teaching, and the Perceived Impact on Their Developing Professional Identity', *European Journal of Teacher Education*, vol. 41, no. 4, pp. 529-548.

Blatchford, P. and Russell, A. (2018) 'Class Size, Grouping Practices and Classroom Management', *International Journal of Educational Research*, vol. 96, pp. 154–163.

Boaler, J. (2013) 'Ability and Mathematics: The Mindset Revolution That Is Reshaping Education', *Forum*, vol. 55, pp. 143–152.

Boberg, J. E. and Bourgeois, S. J. (2016) 'The Effects of Integrated Transformational Leadership', *Journal of Educational Administration*, vol. 54, no. 3, pp. 357–374.

Bol, T. and Van de Werfhorst, H. G. (2011) *Measuring Educational Institutional Diversity: External Differentiation, Vocational Orientation and Standardization*, Amsterdam Centre for Inequality Studies Working Paper Series, 1., Amsterdam Centre for Inequality Studies.

Borer, M. I. and Fontana, A. (2012) 'Postmodern Trends: Expanding the Horizons of Interviewing Practices and Epistemologies', in *The SAGE Handbook of Interview Research: The Complexity of the Craft*, California, SAGE Publications, Inc., pp. 45–60.

Borko, H. (2004) 'Professional Development and Teacher Learning: Mapping the Terrain', *Educational Researcher*, vol. 33, no. 8, pp. 3–15.

Borko, H., Jacobs, J. and Koellner, K. (2010) 'Contemporary Approaches to Teacher Professional Development.', *International Encyclopedia of education*, vol. 72, no. 2, pp. 548-556.

Boslaugh, S. (2010) 'Secondary Data Source', in *Encyclopedia of Research Design*, California, United States, SAGE Publications, Inc., pp. 1330–1331.

Boud, D. (2013) *Enhancing Learning Through Self-Assessment.*, London; New York, Routledge Falmer.

Bowe, R., Ball, S. J. and Gold, A. (2017) *Reforming Education and Changing Schools: Case Studies in Policy Sociology*, London; New York, Routledge.

Boxer, A. and Bennett, T. (eds.) (2019) *The Research(ED) Guide to Explicit & Direct Instruction: An Evidence-Informed Guide for Teachers*, The Research(ED) series, Woodbridge, John Catt.

Boyle, B. and Bragg, J. (2005) 'No Science Today—the Demise of Primary Science', *The Curriculum Journal*, vol. 16, no. 4, pp. 423-437.

Bradbury, J. (2019) 'Off-Rolling: An Update on Recent Analysis - Ofsted Blog: Schools, Early Years, Further Education and Skills', [Online]. Available at https://educationinspection.blog.gov.uk/2019/09/06/off-rolling-an-update-on-recent-analysis/ (Accessed 13 January 2020).

Brady, K. and Woolfson, L. (2008) 'What Teacher Factors Influence Their Attributions for Children's Difficulties in Learning?', *British Journal of Educational Psychology*, vol. 78, no. 4, pp. 527–544.

Bramley, T., Rodeiro, C. and Vitello, S. (2015) *Gender Differences in GCSE*, Cambridge Assessment Research Report, Cambridge UK.

Braun, A., Ball, S. J., Maguire, M. and Hoskins, K. (2011) 'Taking Context Seriously: Towards Explaining Policy Enactments in the Secondary School', *Discourse: Studies in the Cultural Politics of Education*, vol. 32, no. 4, pp. 585–596.

Braun, A., Maguire, M. and Ball, S. J. (2010) 'Policy Enactments in the UK Secondary School: Examining Policy, Practice and School Positioning', *Journal of Education Policy*, vol. 25, no. 4, pp. 547–560.

Braun, V. and Clarke, V. (2006) 'Using Thematic Analysis in Psychology', *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101.

Brookfield, S. D. (2017) *Becoming a Critically Reflective Teacher*, San Francisco, John Wiley & Sons.

Brookhart, S. M. (2010) *How to Assess Higher-Order Thinking Skills in Your Classroom*, Virginia, ASCD.

Brooks, C. (2010) 'Embodied Transcription: A Creative Method for Using Voice-Recognition Software', *The Qualitative Report*, vol. 15, no. 5, pp. 1227–1242.

Brooks, J., McCluskey, S., Turley, E. and King, N. (2015) 'The Utility of Template Analysis in Qualitative Psychology Research', *Qualitative Research in Psychology*, vol. 12, no. 2, pp. 202–222.

Brown, C. (2017) '"Favourite Places in School" for Lower-Set "ability" Pupils: School Groupings Practices and Children's Spatial Orientations', *Children's Geographies*, vol. 15, no. 4, pp. 399–412.

Brown, P., Friedrichsen, P. and Abell, S. (2013) 'The Development of Prospective Secondary Biology Teachers PCK', *Journal of Science Teacher Education*, vol. 24, no. 1, pp. 133–155.

Bryman, A. (2006) 'Integrating Quantitative and Qualitative Research: How Is It Done?', *Qualitative research*, vol. 6, no. 1, pp. 97–113.

Bryman, A. (2016) *Social Research Methods*, Oxford, Oxford University Press.

Buchanan, R. (2015) 'Teacher Identity and Agency in an Era of Accountability', *Teachers and Teaching*, vol. 21, no. 6, pp. 700–719.

Buchs, C., Filippou, D., Pulfrey, C. and Volpé, Y. (2017) 'Challenges for Cooperative Learning Implementation: Reports from Elementary School Teachers', *Journal of Education for Teaching*, vol. 43, no. 3, pp. 296–306.

Buehl, M. M. and Fives, H. (2011) 'Spring Cleaning for the "messy" Construct of Teachers' Beliefs: What Are They? Which Have Been Examined? What Can They Tell Us?', in Harris, K., Graham, Steve., and Urdan, T. (eds), *APA Educational Psychology Handbook: Vol. 2. Individual Differences and Cultural and Contextual Factors*, Washington, American Psychological Association, vol. 2, pp. 471–499.

Burgess, S., Greaves, E., Vignoles, A. and Wilson, D. (2015) 'What Parents Want: School Preferences and School Choice', *The Economic Journal*, vol. 125, no. 587, pp. 1262–1289.

Burgess, S. and Thomson, D. (2019) 'The Impact of the Wolf Reforms on Education Outcomes for Lower-Attaining Pupils', *British Educational Research Journal*, vol. 45, no. 3, pp. 592–621.

Burton, N. (2010) 'Where Did the Science Go? Disappeared or Dissolved?', *Education 3-13*, vol. 38, no. 4, pp. 369–379.

Busby, E. (2018) *Parents told that sick children must sit all SATS exams as calls for boycott grow* [Online]. Available at https://www.independent.co.uk/news/education/education-news/sats-primary-school-exams-parents-ill-boycott-children-mental-health-a8333296.html (Accessed 20 June 2018).

Bush, T. and Glover, D. (2014) 'School Leadership Models: What Do We Know?', *School Leadership & Management*, vol. 34, no. 5, pp. 553–571.

Cadwallader, S. (2014) 'Developing Grade Descriptions for the New GCSEs: Considerations and Challenges', Centre for Education Research and Practice, Manchester.

Calhoun, C. (2002) *Dictionary of the Social Sciences*, New York, Oxford University Press.

Carter, L. (2016) 'Neoliberalism and STEM Education', *Journal for Activist Science and Technology Education*, vol. 7, no. 1.

Chaiklin, S. (2003) 'The Zone of Proximal Development in Vygotsky's Analysis of Learning and Instruction', in Kozulin, A., Gindis, B., Ageyev, V., and Miller, S. (eds), *Vygotsky's Educational Theory in Cultural Context*, Cambridge University Press, pp. 39-62.

Chan, K. K. H. and Yung, B. H. W. (2018) 'Developing Pedagogical Content Knowledge for Teaching a New Topic: More Than Teaching Experience and Subject Matter Knowledge', *Research in Science Education*, vol. 48, no. 2, pp. 233-265.

Chandler-Grevatt, A. (2012) *Transition in Science Education using Connecting Technologies*, Research Report, Falmer, University of Sussex.

Chen, J. A., Morris, D. B. and Mansour, N. (2014) 'Science Teachers' Beliefs: Perceptions of Efficacy and the Nature of Scientific Knowledge and Knowing', in *International Handbook of Research on Teachers' Beliefs*, New York, Routledge, pp. 370–386.

Cheng, E. C. K. and Ling, L. M. (2013) *The Approach of Learning Study: Its Origin and Implications (Vol. 94)*, Education Working Papers, OECD.

Childs, A. and McNicholl, J. (2007) 'Science Teachers Teaching Outside of Subject Specialism: Challenges, Strategies Adopted and Implications for Initial Teacher Education', *Teacher Development*, vol. 11, no. 1, pp. 1–20.

Chin, C. (2006) 'Classroom Interaction in Science: Teacher Questioning and Feedback to Students' Responses', *International Journal of Science Education*, vol. 28, no. 11, pp. 1315-1346.

Chin, C. (2007) 'Teacher Questioning in Science Classrooms: Approaches That Stimulate Productive Thinking', *Journal of research in Science Teaching*, vol. 44, no. 6, pp. 815–843.

Chizhik, E. and Chizhik, A. (2016) '(Re)Conceptualizing the Purpose of the Lesson Plan', *The Journal of Educational Thought (JET) / Revue de la Pensée Éducative*, vol. 49, no. 2, pp. 210–225.

Chowdry, H., Crawford, C. and Goodman, A. (2011) 'The Role of Attitudes and Behaviours in Explaining Socio-Economic Differences in Attainment at Age 16', *Longitudinal and Life Course Studies*, vol. 2, no. 1, pp. 59–76.

Clarke, D. and Hollingsworth, H. (2002) 'Elaborating a Model of Teacher Professional Growth', *Teaching and Teacher Education*, vol. 18, no. 8, pp. 947–967.

Clement, J. (2014) 'Managing Mandated Educational Change', *School Leadership & Management*, vol. 34, no. 1, pp. 39–51.

Coates, S. (2011) *First Report of the Review of Teachers Standards*, London, DfE.

COBIS (2018) *Teacher Supply in British International Schools Final Report*, London, COBIS.

Coe, R. (2002) 'It's the Effect Size, Stupid: What Effect Size Is and Why It Is Important.', University of Exeter, England.

Coe, R., Aloisi, C., Higgins, S. and Major, L. E. (2014) *What Makes Great Teaching? Review of the Underpinning Research*, London, Sutton Trust.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Hillsdale, N.J, L. Erlbaum Associates.

Cohen, L., Manion, L. and Morrison, K. (2017) *Research Methods in Education*, 7th edn, London ; New York, Routledge.

Collins, S., Reiss, M. and Stobart, G. (2010) 'What Happens When High-Stakes Testing Stops? Teachers' Perceptions of the Impact of Compulsory National Testing in Science of 11-Year-Olds in England and Its Abolition in Wales', *Assessment in Education: Principles, Policy & Practice*, vol. 17, no. 3, pp. 273–286.

Connell, R. (2013) 'The Neoliberal Cascade and Education: An Essay on the Market Agenda and Its Consequences', *Critical Studies in Education*, vol. 54, no. 2, pp. 99–112.

de Corte, E. (2010) 'Historical Developments in the Understanding of Learning', in Hanna, D., David, I., and Francisco, B. (eds), *Educational Research and Innovation The Nature of Learning Using Research to Inspire Practice: Using Research to Inspire Practice*, Paris, OECD Publishing, pp. 35–68.

Cotton, K. (2001) 'Classroom Questioning', *School Improvement Research Series*, vol. 3.

CPD for Teachers (2018) *Courses* [Online]. Available at https://www.cpdforteachers.com/courses (Accessed 21 September 2018).

Creative Education (2018) *Mock Ofsted & Department Reviews | Creative Education* [Online]. Available at https://www.creativeeducation.co.uk/consultancy/ofsted (Accessed 21 September 2018).

Creswell, J. W. (2010) 'Mapping the Developing Landscape of Mixed Methods Research', in Tashakkori, A. and Teddlie, C. (eds), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, California, SAGE Publications, pp. 45–68.

Creswell, J. W. (2018) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 5th edn, California, US., Sage.

Crouch, C., Fagen, A. P., Callan, J. P. and Mazur, E. (2004) 'Classroom Demonstrations: Learning Tools or Entertainment?', *American Journal of Physics*, vol. 72, no. 6, pp. 835–838.

Cuban, L. (2013) *Inside the Black Box of Classroom Practice: Change without Reform in American Education.*, Massachusetts, Harvard Education Press.

Cullingford, C. (2017) 'Secondary School Pupils' Experience of the National Curriculum', in Cullingford, C. and Oliver, P. (eds), *The National Curriculum and its Effects*, Routledge, pp. 45–62.

Cullingford, C. and Daniels, S. (2013) 'Effects of Ofsted Inspections on School Performance', in Cullingford, C. (ed), *An Inspector Calls: Ofsted and Its Effect on School Standards*, London ; New York, Routledge, pp. 59–69.

Cullingford, C. and Oliver, P. (2017) *The National Curriculum and its Effects*, Abingdon; New York, Routledge.

Cumming, C. (2011) 'CPD: Support Strategies for Professional Learning, National Initiatives and Major Curriculum Reform', *Improving Schools*, vol. 14, no. 2, pp. 145–155.

Cutler, M. (2015) *Primary science: the poor relation?* [Online]. Available at http://cprtrust.org.uk/cprt-blog/primary-science-the-poor-relation/ (Accessed 20 March 2017).

Daniel, J. (2011) *Sampling Essentials: Practical Guidelines for Making Sampling Choices*, California, Sage.

Davis, K. S. (2003) '"Change Is Hard": What Science Teachers Are Telling Us About Reform and Teacher Learning of Innovative Practices', *Science Education*, vol. 87, no. 1, pp. 3–30.

Day, C., Elliot, B. and Kington, A. (2005) 'Reform, Standards and Teacher Identity: Challenges of Sustaining Commitment', *Teaching and Teacher Education*, vol. 21, no. 5, pp. 563–577.

DCSF (2008) *Pupil Characteristics and Class Sizes in Maintained Schools in England: January 2008 (Provisional)* [Online]. Available at https://webarchive.nationalarchives.gov.uk/20080610163225/http://www.dfes.gov.uk/rsgateway/DB/SFR/s000786/index.shtml.

DCSF (2009) *GCSE and Equivalent Results in England* [Online]. Available at https://webarchive.nationalarchives.gov.uk/20120505061657oe_/http://www.education.gov.uk/a00195931/gcse-and-equivalent-results-in-england.

DCSF (2010) *KS4 Results in England, 2008/09 (Revised)* [Online]. Available at https://webarchive.nationalarchives.gov.uk/20120505061506oe_/http://www.education.gov.uk/a00196306/ks4-results-in-england-2008-09-(revised).

De Houwer, J., Barnes-Holmes, D. and Moors, A. (2013) 'What Is Learning? On the Nature and Merits of a Functional Definition of Learning', *Psychonomic Bulletin & Review*, vol. 20, no. 4, pp. 631–642.

De Vaus, D. (2014) *Surveys in Social Research*, 6th edn, London, Routledge.

Dearing, R. (1993) *The National Curriculum and Its Assessment*, National Curriculum Council York, England.

Defend Digital Me (2018) 'The State of Data: Lessons for Policymakers Report', [Online]. Available at http://defenddigitalme.com/wp-content/uploads/2018/05/StateOfDataReport_policymakers_ddm.pdf (Accessed 3 July 2018).

Desimone, L. (2009) 'Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures', *Educational Researcher*, vol. 38, no. 3, pp. 181–199.

DeWitt, J. and Archer, L. (2015) 'Who Aspires to a Science Career? A Comparison of Survey Responses from Primary and Secondary School Students', *International Journal of Science Education*, vol. 37, no. 13, pp. 2170–2192.

DfE (2010a) *The Importance of Teaching: The Case for Change*, London, DfE.

DfE (2010b) *The Importance of Teaching: The Schools White Paper*, London, DfE.

DfE (2011) *The National Strategies 1997-2011: A Brief Summary of the Impact and Effectiveness of the National Strategies*, London, DfE.

DfE (2013a) *Teachers' Standards - Guidance for School Leaders, School Staff and Governing Bodies*, London, DfE.

DfE (2013b) *National Curriculum in England Consultation Report - FINAL*, London, DfE.

DfE (2013c) *Assessment without Levels* [Online]. Available at http://webarchive.nationalarchives.gov.uk/20130802141012/https://www.education.gov.uk/schools/teachingandlearning/curriculum/nationalcurriculum2014/a00225864/assessing-without-levels (Accessed 9 August 2017).

DfE (2013d) *Primary Assessment and Accountability Under the New Curriculum Consultation Document*, London, DfE.

DfE (2014a) 'The National Curriculum in England: Framework Document', DfE.

DfE (2014b) 'National Curriculum and Assessment from September 2014: Information for Schools', DfE.

DfE (2014c) 'Science KS4 Programme of Study', [Online]. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/381380/Science_KS4_PoS_7_November_2014.pdf (Accessed 18 May 2020).

DfE (2014d) *GCSE and Equivalent Results: 2012 to 2013 (revised)* [Online]. Available at https://www.gov.uk/government/statistics/gcse-and-equivalent-results-in-england-2012-to-2013-revised.

DfE (2015a) *Reading: The Next Steps: Supporting Higher Standards in Schools*, London, DfE.

DfE (2015b) *National Curriculum Assessments at Key Stage 2 in England, 2015 (revised)*, London, DfE [Online]. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/483897/SFR47_2015_text.pdf (Accessed 2 February 2017).

DfE (2015c) 'National Pupil Database: User Guide and Supporting Information', DfE.

DfE (2016a) *Revised GCSE and Equivalent Results in England: 2014 to 2015* [Online]. Available at https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2014-to-2015.

DfE (2016b) *Section 8: Science Sampling - 2016 Key Stage 2: Assessment and Reporting Arrangements* [Online]. Available at https://www.gov.uk/guidance/2016-key-stage-2-assessment-and-reporting-arrangements-ara/section-8-science-sampling (Accessed 28 January 2018).

DfE (2017) *GCSE and Equivalent Results: 2016 to 2017 (provisional)* [Online]. Available at https://www.gov.uk/government/statistics/gcse-and-equivalent-results-2016-to-2017-provisional.

DfE (2018a) *Bursaries and Scholarships for Teacher Training - Get into Teaching* [Online]. Available at https://getintoteaching.education.gov.uk/funding-my-teacher-training/bursaries-and-scholarships-for-teacher-training (Accessed 2 May 2019).

DfE (2018b) 'School Teachers' Pay and Conditions: 2018', DfE.

DfE (2018c) *Literacy and Numeracy Catch-up Strategies*, London, DfE.

DfE (2018d) *Attainment in Primary Schools in England: Quality and Methodology Information*, London.

DfE (2018e) *National Curriculum Assessments at Key Stage 2 in England, 2018 (revised)*, London.

DfE (2018f) *National Pupil Database: User Guide and Supporting Information* [Online]. Available at https://www.gov.uk/government/publications/national-pupil-database-user-guide-and-supporting-information (Accessed 18 June 2018).

DfE (2018g) *Statistics: School and Pupil Numbers* [Online]. Available at https://www.gov.uk/government/collections/statistics-school-and-pupil-numbers (Accessed 13 February 2019).

DfE (2018h) *Ready Reckoners and Transition Matrices for Key Stage 4: 2017* [Online]. Available at https://www.gov.uk/government/publications/ready-reckoners-for-key-stage-4-2017 (Accessed 21 November 2018).

DfE (2019a) *Secondary Accountability Measures: Guide for Maintained Secondary Schools, Academies and Free Schools*, London, DfE.

DfE (2019b) *Key Stage 4 Including Multi-Academy Trust Performance, 2018 (revised)*, London, DfE.

DfE (2019c) *Key Stage 4 and Multi-Academy Trust Performance 2018 (revised)* [Online]. Available at https://www.gov.uk/government/statistics/key-stage-4-and-multi-academy-trust-performance-2018-revised.

DfE (2019d) *Schools, Pupils and Their Characteristics: January 2019*, Sheffield, DfE [Online]. Available at

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812539/Schools_Pupils_and_their_Characteristics_2019_Main_Text.pdf (Accessed 18 December 2019).

DfE (2019e) *Statistics: Special Educational Needs (SEN)* [Online]. Available at https://www.gov.uk/government/collections/statistics-special-educational-needs-sen.

DfE (2019f) *Statistics: School and Pupil Numbers* [Online]. Available at https://www.gov.uk/government/collections/statistics-school-and-pupil-numbers.

DfE (2019g) *Special Educational Needs in England: January 2019*, London, DfE.

DfE (2019h) *Special Educational Needs: An Analysis and Summary of Data Sources*, London, DfE.

DfE (2019i) *School Workforce in England: November 2018*, London, DfE.

DfE (2019j) *Subject Knowledge Enhancement (SKE) Courses | Get into Teaching* [Online]. Available at https://getintoteaching.education.gov.uk/explore-my-options/teacher-training-routes/subject-knowledge-enhancement-ske-courses (Accessed 3 July 2019).

DfE (n.d.) *Types of school - GOV.UK* [Online]. Available at /types-of-school/academies (Accessed 2 January 2017).

DfE and Department of Health (2015) 'Special Educational Needs and Disability Code of Practice: 0 to 25 Years', DfE.

DfE and Gove, M. (2011) *National Curriculum Review Launched - GOV.UK* [Online]. Available at https://www.gov.uk/government/news/national-curriculum-review-launched (Accessed 4 August 2017).

DfE and Gove, M. (2012) *2012 OECD PISA results - Oral statements to Parliament* [Online]. Available at https://www.gov.uk/government/speeches/2012-oecd-pisa-results (Accessed 21 February 2017).

DfE and Gove, M. (2013) *Curriculum, Exam and Accountability Reform* [Online]. Available at https://www.gov.uk/government/speeches/curriculum-exam-and-accountability-reform (Accessed 14 August 2017).

DfE and National Statistics (2010) *National Curriculum Assessments at Key Stage 2 in England, 2010 (revised)*, Statistical First Release, London, DfE.

DfEE (1998) 'Teaching: High Status, High Standards Circular 4/98', HMSO.

DfES (1991) *Science for Ages 5 to 16 (1991). Proposals of the Secretary of State for Education and Science*, London, Her Majesty's Stationery Office.

DfES (2003) 'GCSEs in Vocational Subjects: A General Guide and Overview of the Qualifications', [Online]. (Accessed 18 May 2020).

DfES (2004) *Pedagogy and Practice: Teaching and Learning in Secondary Schools Unit 12: Assessment for learning* [Online]. Available at http://webarchive.nationalarchives.gov.uk/20110809174429/http://wsassets.s3.amazonaws.com/ws/nso/pdf/2deff878cffd2cdcd59a61df29e73105.pdf (Accessed 3 May 2018).

DfES (2006a) 'Primary National Strategy: Primary Framework for Literacy and Mathematics', DfES Publications.

DfES (2006b) *A Condensed Key Stage 3: Improving Key Stage 2 to Key Stage 3 Transfer.*, Curriculum and Standards, Secondary National Strategy for School Improvement: Guidance, London, DfES.

DfES (2007) *Gender and Education: The Evidence on Pupils in England*, London, DfES Publications [Online]. Available at http://webarchive.nationalarchives.gov.uk/20090108131527/http://www.dcsf.gov.uk/research/data/uploadfiles/RTP01-07.pdf (Accessed 26 February 2018).

DfES and QCA (2004) 'Science in the National Curriculum', DfES Publications [Online]. Available at https://www.stem.org.uk/system/files/elibrary-resources/legacy_files_migrated/8969-Science%20in%20the%20National%20Curriculum%202004%20A.pdf (Accessed 12 May 2020).

Diamond, J. B. and Spillane, J. P. (2016) 'School Leadership and Management from a Distributed Perspective: A 2016 Retrospective and Prospective.', *Management in Education*, vol. 30, no. 4, pp. 147–154.

van Dijk, E. M. and Kattmann, U. (2007) 'A Research Model for the Study of Science Teachers' PCK and Improving Teacher Education', *Teaching and Teacher Education*, vol. 23, no. 6, pp. 885–897.

Dillon, J. and Manning, A. (2010) 'Science Teachers, Science Teaching: Issues and Challenges', in Osborne, J. and Dillon, J. (eds), *Good Practice In Science Teaching: What Research Has To Say: What research has to say*, McGraw-Hill Education (UK), pp. 6–19.

Dixson, D. D. and Worrell, F. C. (2016) 'Formative and Summative Assessment in the Classroom', *Theory into practice*, vol. 55, no. 2, pp. 153–159.

Dong, L., Clarke, D., Cao, Y., Wang, L. and Seah, W. T. (2019) 'Teacher Questioning Practices over a Sequence of Consecutive Lessons: A Case Study of Two Mathematics Teachers', *Sustainability*, vol. 11, no. 1, p. 139.

Donnelly, J. and Ryder, J. (2011) 'The Pursuit of Humanity: Curriculum Change in English School Science', *History of Education*, vol. 40, no. 3, pp. 291–313.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J. and Willingham, D. T. (2013) 'Improving Students' Learning with Effective Learning Techniques: Promising Directions from Cognitive and Educational Psychology.', *Psychological Science in the Public Interest*, vol. 14, no. 1, pp. 4–58.

Dunne, M., Pryor, J. and Yates, P. (2005) *Becoming a Researcher: A Research Companion for the Social Sciences: A Companion to the Research Process*, McGraw-Hill Education (UK).

Dweck, C. S. (2007) 'Is Math a Gift? Beliefs That Put Females at Risk', in *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, Washington, DC, US, American Psychological Association, pp. 47–55.

Dweck, C. S. (2008) 'Can Personality Be Changed? The Role of Beliefs in Personality and Change', *Current Directions in Psychological Science*, SAGE Publications Inc, vol. 17, no. 6, pp. 391–394.

Dweck, C. S. (2014) *Mindset: The New Psychology of Success*, New York, Random House.

Earl, L., Watson, N. and Katz, S. (2003) *Large-Scale Education Reform: Life Cycles and Implications for Sustainability*, Reading, CfBT.

Earl, L., Watson, N., Levin, B., Leithwood, K., Fullan, M. and Torrance, N. (2003) *Final Report of the External Evaluation of England's National Literacy and Numeracy Strategies. Final Report. Watching & Learning 3.*, ERIC.

EdComms (2016) 'Why Your Business Should Go Back to School: How You Can Inspire the Next Stem Generation', [Online]. Available at https://www.stem.org.uk/system/files/elibrary-resources/2016/08/ProjectEnthuse_FINAL_0.pdf (Accessed 29 June 2020).

EEF (2018a) *Teaching and Learning Toolkit* [Online]. Available at https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit (Accessed 5 September 2018).

EEF (2018b) 'The Attainment Gap 2017', Education Endowment Foundation [Online]. Available at https://educationendowmentfoundation.org.uk/public/files/Annual_Reports/EEF_Attainment_Gap_Report_2018_-_print.pdf (Accessed 26 February 2018).

EEF (2018c) *Teaching Assistants - Teaching and Learning Toolkit*, EEF [Online]. Available at https://educationendowmentfoundation.org.uk/pdf/generate/?u=https://educationendowmentfoundation.org.uk/pdf/toolkit/?id=149&t=Teaching%20and%20Learning%20Toolkit&e=149&s= (Accessed 28 March 2019).

EL-Deghaidy, H., Mansour, N., Aldahmash, A. and Alshamrani, S. (2015) 'A Framework for Designing Effective Professional Development: Science Teachers' Perspectives in a Context of Reform.', *Eurasia Journal of Mathematics, Science & Technology Education*, vol. 11, no. 6.

Elliott, A. (2012) 'RISE Review: Twenty Years Inspecting English Schools - Ofsted 1992-2012', Research and Information on State Education Trust.

Elwood, J. (2005) 'Gender and Achievement: What Have Exams Got to Do with It?', *Oxford Review of Education*, vol. 31, no. 3, pp. 373–393.

Emirbayer, M. and Mische, A. (1998) 'What Is Agency?', *American Journal of Sociology*, vol. 103, no. 4, pp. 962–1023.

Empiribox (2018) *Discover Empiribox* [Online]. Available at https://www.empiribox.org/discover-empiribox/ (Accessed 12 September 2018).

Enderle, P., Dentzau, M., Roseler, K., Southerland, S., Granger, E., Hughes, R., Golden, B. and Saka, Y. (2014) 'Examining the Influence of Rets on Science Teacher Beliefs and Practice: Science Teacher Beliefs and Practice', *Science Education*, vol. 98, no. 6, pp. 1077–1108.

Engelmann, S. (1980) *Direct Instruction*, The Instructional Design Library, Englewood Cliffs, NJ., Educational Technology, vol. 22.

English, E., Hargreaves, L. and Hislam, J. (2002) 'Pedagogical Dilemmas in the National Literacy Strategy: Primary Teachers' Perceptions, Reflections and Classroom Behaviour', *Cambridge Journal of Education*, vol. 32, no. 1, pp. 9–26.

Erdogan, I. and Campbell, T. (2008) 'Teacher Questioning and Interaction Patterns in Classrooms Facilitated with Differing Levels of Constructivist Teaching Practices', *International Journal of Science Education*, vol. 30, no. 14, pp. 1891–1914.

EU STEM Coalition Team (2016) 'STEM Skills for a Futureproof Europe', EU STEM Coalition, Netherlands.

European Commission (2018) *2018 Reform of EU Data Protection Rules* [Online]. Available at https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en (Accessed 3 July 2018).

Evans, L. (2011) 'The "shape" of Teacher Professionalism in England: Professional Standards, Performance Management, Professional Development and the Changes Proposed in the 2010 White Paper', *British Educational Research Journal*, vol. 37, no. 5, pp. 851–870.

Evans, T. and Harris, J. (2006) 'A Case of Mistaken Identity? Debating the Dilemmas of Street-Level Bureaucracy with Musil et al.', *European Journal of Social Work*, vol. 9, no. 4, pp. 445–459.

Farrimond, H. (2017) 'The Ethics of Research', in Wyse, D., Selwyn, N., Smith, E., and Suter, L. E. (eds), *The BERA/SAGE Handbook of Educational Research*, London, SAGE Publications, pp. 72–89.

Field, A. (2000) *Discovering Statistics Using SPSS for Windows*, London, SAGE Publications Ltd.

Fisher, D. and Frey, N. (2014) *Checking for Understanding: Formative Assessment Techniques for Your Classroom*, Virginia, US., ASCD.

Fives, H. and Buehl, M. M. (2016) 'Teachers' Beliefs, in the Context of Policy Reform', *Policy Insights from the Behavioral and Brain Sciences*, vol. 3, no. 1, pp. 114–121.

Florian, L. and Black-Hawkins, K. (2011) 'Exploring Inclusive Pedagogy', *British Educational Research Journal*, vol. 37, no. 5, pp. 813–828.

Forde, C., McMahon, M. A., Hamilton, G. and Murray, R. (2016) 'Rethinking Professional Standards to Promote Professional Learning', *Professional Development in Education*, vol. 42, no. 1, pp. 19–35.

Foster, D. (2018) *Teacher Recruitment and Retention in England*, House of Commons Briefing Paper, London, House of Commons.

Foucault, M. (1972) *Archaeology of Knowledge*, London, Tavistock.

Furtak, E. M., Seidel, T., Iverson, H. and Briggs, D. C. (2012) 'Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis', *Review of educational research*, vol. 82, no. 3, pp. 300–329.

Galton, M., Hargreaves, L. and Pell, T. (2009) 'Group Work and Whole-Class Teaching with 11-to 14-Year-Olds Compared', *Cambridge Journal of Education*, vol. 39, no. 1, pp. 119–140.

Gamlem, S. M. and Smith, K. (2013) 'Student Perceptions of Classroom Feedback', *Assessment in Education: Principles, Policy & Practice*, vol. 20, no. 2, pp. 150–169.

Gayle, V., Murray, S. and Connelly, R. (2016) 'Young People and School General Certificate of Secondary Education Attainment: Looking for the "Missing Middle"', *British Journal of Sociology of Education*, vol. 37, no. 3, pp. 350–370.

Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D. and Stuhlsatz, M. A. M. (2017) 'Teacher Pedagogical Content Knowledge, Practice, and Student Achievement', *International Journal of Science Education*, vol. 0, no. 0, pp. 1–20.

Giddens, A. (1984) *The Constitution of Society: Outline of the Theory of Structuration*, Los Angeles, University of California Press.

Gilbert, C. (2012) 'Towards a Self-Improving System: The Role of School Accountability', NCSL.

Gill, K. (2017) 'Two Decades of Reading: An Analysis of English Policy Affecting Literacy 1997-2016', Literacy Trust.

Gill, M. G., Ashton, P. T. and Algina, J. (2004) 'Changing Preservice Teachers' Epistemological Beliefs About Teaching and Learning in Mathematics: An Intervention Study', *Contemporary Educational Psychology*, Epistemological Development and Its Impact on Cognition in Academic Domains, vol. 29, no. 2, pp. 164–185.

Gill, T. (2012) *Provision of GCSE subjects 2010*, Statistics Report Series.

Gill, T. and Bell, J. F. (2013) 'What Factors Determine the Uptake of A-level Physics?', *International Journal of Science Education*, vol. 35, no. 5, pp. 753–772.

Gillies, R. M. (2013) 'Productive Academic Talk During Inquiry-Based Science', *Pedagogies: An International Journal*, vol. 8, no. 2, pp. 126–142.

Gillies, R. M. and Boyle, M. (2010) 'Teachers' Reflections on Cooperative Learning: Issues of Implementation', *Teaching and Teacher Education*, vol. 26, no. 4, pp. 933–940.

Gilson, L. (2015) 'Michael Lipsky - Street-Level Bureaucracy: Dilemmas of the Individual in Public Service', in Lodge, M. and Balla, S. (eds), *The Oxford Handbook of Classics in Public Policy and Administration*, Oxford, Oxford University Press.

Glackin, M. (2016) '"Risky Fun" or "Authentic Science"? How Teachers' Beliefs Influence Their Practice During a Professional Development Programme on Outdoor Learning', *International Journal of Science Education*, Routledge, vol. 38, no. 3, pp. 409–433.

Goddard, R. D., Hoy, W. K. and Hoy, A. W. (2004) 'Collective Efficacy Beliefs: Theoretical Developments, Empirical Evidence, and Future Directions', *Educational Researcher*, vol. 33, no. 3, pp. 3–13.

Goertz, G. and Mahoney, J. (2012) 'Concepts and Measurement: Ontology and Epistemology', *Social Science Information*, vol. 51, no. 2, pp. 205–216.

Goldstein, H. (2001) 'Using Pupil Performance Data for Judging Schools and Teachers: Scope and limitations', *British Educational Research Journal*, vol. 27, no. 4, pp. 433–442.

Goodman, R. and Burton, D. (2012) 'What Is the Nature of the Achievement Gap, Why Does It Persist and Are Government Goals Sufficient to Create Social Justice in the Education System?', *Education 3-13*, vol. 40, no. 5, pp. 500–514.

Gorard, S. (2002) 'The Role of Secondary Data in Combining Methodological Approaches', *Educational Review*, vol. 54, no. 3, pp. 231–237.

Gorard, S. (2010) 'Research Design, as Independent of Methods.', in Tashakkori, A. and Teddlie, C. (eds), *SAGE Handbook of Mixed Methods in Social & Behavioral Research.*, 2nd edn, London, SAGE Publications, pp. 237–252.

Gorard, S. (2016) 'A Cautionary Note on Measuring the Pupil Premium Attainment Gap in England', *British Journal of Education, Society & Behavioural Science*, vol. 14, no. 2, pp. 1–8.

Gorard, S. and Fitz, J. (1998) 'The More Things Change ... The Missing Impact of Marketisation?', *British Journal of Sociology of Education*, vol. 19, no. 3, pp. 365–376.

Gove, M. (2009) 'What is Education For?', Speech, Royal Society of Arts [Online]. Available at https://www.thersa.org/globalassets/pdfs/blogs/gove-speech-to-rsa.pdf (Accessed 8 November 2019).

Gove, M. (2011) *Speech: Improving the Standard of Education in England* [Online]. Available at https://www.gov.uk/government/speeches/michael-gove-to-ofqual-standards-summit (Accessed 13 March 2018).

Gove, M. (2013) *Oral statement by Michael Gove on education reform* [Online]. Available at https://www.gov.uk/government/speeches/oral-statement-on-education-reform (Accessed 4 September 2017).

GOV.UK (2019) *Pupil progress between 11 and 16 years old ('Progress 8')* [Online]. Available at https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/11-to-16-years-old/pupil-progress-progress-8-between-ages-11-and-16-key-stage-2-to-key-stage-4/latest (Accessed 25 November 2019).

GOV.uk (n.d.) *Get Information About Schools* [Online]. Available at https://get-information-schools.service.gov.uk/Establishments/Search?tok=zcWvdb1d (Accessed 17 October 2018).

Graham, J. and Haidt, J. (2010) 'Beyond Beliefs: Religions Bind Individuals Into Moral Communities', *Personality and Social Psychology Review*, SAGE Publications Inc, vol. 14, no. 1, pp. 140–150.

Gray, C. and MacBlain, S. (2015) *Learning Theories in Childhood*, 2nd edn, London, Sage Publications.

Gredler, M. E. (2009) *Learning and Instruction: Theory into Practice*, 6th ed., New Jersey, Merrill Pearson.

Greene, J. C., Caracelli, V. J. and Graham, W. F. (1989) 'Toward a Conceptual Framework for Mixed-Method Evaluation Designs', *Educational Evaluation and Policy Analysis*, vol. 11, no. 3, pp. 255–274.

Gregg, P. and Washbrook, E. (2011) 'The Role of Attitudes and Behaviours in Explaining Socio-Economic Differences in Attainment at Age 11', *Longitudinal and Life Course Studies*, vol. 2, no. 1, pp. 41–58.

Gronn, P. (2003) *The New Work of Educational Leaders: Changing Leadership Practice in an Era of School Reform*, London, SAGE.

Guest, G., Namey, E. E. and Mitchell, M. L. (2012) *Collecting qualitative data: A field manual for applied research*, Sage.

Gunter, H. (2017) 'Corporate Consultancy Practices in Education Services in England', in Gunter, H. and David, H. (eds), *Corporate elites and the reform of public education*, Policy Press, pp. 147–160.

Guskey, T. R. (2002) 'Professional Development and Teacher Change', *Teachers and Teaching*, vol. 8, no. 3, pp. 381–391.

Hallinger, P. and Heck, R. H. (2010) 'Leadership for Learning: Does Collaborative Leadership Make a Difference in School Improvement?', *Educational Management Administration & Leadership*, vol. 38, no. 6, pp. 654–678.

Halstead, M. (1994) 'Accountability and Values', in Scott, D. (ed), *Accountability and Control in Educational Settings*, London, Cassell, pp. 146–165.

Hamlyn, R., Matthews, P., Shanahan, M. and Kantar Public (2017) *Young People's Views on Science Education*, Science Education Tracker Research Report, London, Wellcome Trust.

Hammersley, M. (2011) 'Reflexivity', in Lewis-Beck, M., Bryman, A., and Futing Liao, T. (eds), *The SAGE Encyclopedia of Social Science Research Methods*, California, Sage Publications, pp. 934–935.

Hammersley, M. (2014) *The Limits of Social Science: Causal Explanation and Value Relevance*, London, SAGE Publications Ltd [Online]. Available at http://methods.sagepub.com/book/the-limits-of-social-science (Accessed 1 June 2018).

Hampden-Thompson, G. and Bennett, J. (2013) 'Science Teaching and Learning Activities and Students' Engagement in Science', *International Journal of Science Education*, vol. 35, no. 8, pp. 1325–1343.

Hampden-Thompson, G., Lubben, F. and Bennett, J. (2011) 'Post-16 Physics and Chemistry Uptake: Combining Large-Scale Secondary Analysis with in-Depth Qualitative Methods', *International Journal of Research & Method in Education*, vol. 34, no. 3, pp. 289–307.

Haney, J. J., Lumpe, A. T., Czerniak, C. M. and Egan, V. (2002) 'From Beliefs to Actions: The Beliefs and Actions of Teachers Implementing Change', *Journal of Science Teacher Education*, vol. 13, no. 3, pp. 171–187.

Hanna, D., David, I. and Francisco, B. (eds.) (2010) *Educational Research and Innovation: The Nature of Learning Using Research to Inspire Practice: Using Research to Inspire Practice*, OECD Publishing.

Hanushek, E. A., Jamison, D. T. and Wößmann, L. (2008) *Education and Economic Growth: It's Not Just Going to School, but Learning Something While There That Matters,* IDEAS Working Paper Series from RePEc, St. Louis., Federal Reserve Bank of St Louis.

Hargreaves, A., Earl, L., Moore, S. and Manning, S. (2002) *Learning to Change: Teaching Beyond Subjects and Standards*, John Wiley & Sons.

Hargreaves, D. (2011) *The Challenge for the Comprehensive School: Culture, Curriculum and Community*, Routledge.

Harlen, W. (2012) *Primary Science Assessment*, London, Nuffield Foundation.

Harris, L. R. and Brown, G. T. L. (2013) 'Opportunities and Obstacles to Consider When Using Peer- and Self-Assessment to Improve Student Learning: Case Studies into Teachers' Implementation', *Teaching and Teacher Education*, vol. 36, pp. 101–111.

Hartley, D. (2010) 'Paradigms: How Far Does Research in Distributed Leadership "Stretch"?', *Educational Management Administration & Leadership*, vol. 38, no. 3, pp. 271–285.

Harvey, W. (2011) 'Strategies for Conducting Elite Interviews.', *Qualitative Research*, vol. 11, no. 4, pp. 431–441.

Hatcher, R. (2006) 'Privatization and Sponsorship: The Re-Agenting of the School System in England', *Journal of Education Policy*, vol. 21, no. 5, pp. 599–619.

Hattie, J. (2009) *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*, Abingdon, England, Routledge.

Hattie, J. and Timperley, H. (2007) 'The Power of Feedback', *Review of educational research*, vol. 77, no. 1, pp. 81–112.

Havnes, A. and Prøitz, T. S. (2016) 'Why Use Learning Outcomes in Higher Education? Exploring the Grounds for Academic Resistance and Reclaiming the Value of Unexpected Learning', *Educational Assessment, Evaluation and Accountability*, vol. 28, no. 3, pp. 205–223.

Hay, C. (2002) *Political Analysis*, Basingstoke, Palgrave.

Heijden, H. R. M. A. van der, Geldens, J. J. M., Beijaard, D. and Popeijus, H. L. (2015) 'Characteristics of Teachers as Change Agents', *Teachers and Teaching*, vol. 21, no. 6, pp. 681-699.

Heineke, A. J., Ryan, A. M. and Tocci, C. (2015) 'Teaching, Learning, and Leading: Preparing Teachers as Educational Policy Actors', *Journal of Teacher Education*, vol. 66, no. 4, pp. 382–394.

Heinrich, M. and Stringer, N. (2012) *The Effects on Schools and Pupils of Modularising GCSEs*, Manchester, Centre for Education Research and Policy.

Hennah, N. (2018) 'A Novel Practical Pedagogy for Terminal Assessment', *Chemistry Education Research and Practice* [Online]. Available at https://pubs.rsc.org/en/content/articlelanding/2019/rp/c8rp00186c (Accessed 17 December 2018).

Heritage, M. and Heritage, J. (2013) 'Teacher Questioning: The Epicenter of Instruction and Assessment', *Applied Measurement in Education*, vol. 26, no. 3, pp. 176–190.

Herr, K. and Anderson, G. (2005) 'The Continuum of Positionality in Action Research', in *The Action Research Dissertation: A Guide for Students and Faculty*, California, SAGE Publications, pp. 29-48.

Herrington, D. G., Bancroft, S. F., Edwards, M. M. and Schairer, C. J. (2016) 'I Want to Be the Inquiry Guy! How Research Experiences for Teachers Change Beliefs, Attitudes, and Values About Teaching Science as Inquiry.', *Journal of Science Teacher Education*, vol. 27, no. 2, pp. 183-204.

Hicks, T. (2015) 'Inequality, Marketisation and the Left: Schools Policy in England and Sweden', *European Journal of Political Research*, vol. 54, no. 2, pp. 326–342.

Hill, C. (2014) 'Using DIRT as a Learning Journey', *UKEdChat* [Online]. Available at https://ukedchat.com/2014/01/07/dirt/ (Accessed 16 December 2018).

Hillcrest (2017) *Revision Courses for GCSE, IGCSE, AS, A2, A-Level & IB* [Online]. Available at https://www.hillcrestrevision.com/ (Accessed 21 September 2018).

Hirsch, D. (1994) 'Schools: A Matter of Choice', *OECD Observer*, vol. 187, no. 2, pp. 12–15.

Hiscock, N. (2019) *ASE: How to Select a Scheme of Work?* [Online]. Available at https://www.ase.org.uk/resources/how-select-scheme-of-work (Accessed 23 April 2019).

HM Treasury (2004) *Science & Innovation Investment Framework: 2004-2014*, London, Stationery Office.

HM Treasury (2014) *Our Plan for Growth: Science and Innovation*, London, Department for Business Innovation & Skills.

HMSO (1988) *Education Reform Act 1988* [Online]. Available at http://www.legislation.gov.uk/ukpga/1988/40/pdfs/ukpga_19880040_en.pdf (Accessed 14 November 2016).

Hoadley, U. (2006) 'Analysing Pedagogy: The Problem of Framing', *Journal of Education*, vol. 40, no. 1, pp. 15–34.

Hoch, J. E., Bommer, W. H., Dulebohn, J. H. and Wu, D. (2018) 'Do Ethical, Authentic, and Servant Leadership Explain Variance Above and Beyond Transformational Leadership? A Meta-Analysis', *Journal of Management*, vol. 44, no. 2, pp. 501–529.

Holloway, J. and Brass, J. (2018) 'Making Accountable Teachers: The Terrors and Pleasures of Performativity', *Journal of Education Policy*, vol. 33, no. 3, pp. 361–382.

Holman, J. (2018) 'Response to the Professional Bodies' Articles on Developing Uk Science Curriculum Frameworks', *School Science Review*, vol. 100, no. 370, pp. 44–46.

Holmes, A. (2014) 'Researcher Positionality – a Consideration of Its Influence and Place in Research', University of Hull.

House of Commons (2009) *National Curriculum Fourth Report of Session 2008-09 Volume I Report, together with formal minutes*, London, House of Commons Children, Schools and Families Committee.

House of Commons Education Committee (2015) *Academies and Free Schools: Fourth Report of Session 2014-15*, London, House of Commons.

House of Commons Select Committee (2008) *Testing and Assessment: Third Report of Session 2007-2008. Volume 1 Report, together with formal minutes*, London, House of Commons Children, Schools and Families Committee.

Hoy, W. K. (2002) 'An Analysis of Enabling and Mindful School Structures: Some Theoretical, Research and Practical Considerations', *Journal of Educational Administration*, vol. 41, no. 1, pp. 87–109.

Hupe, P. and Hill, M. (2007) 'Street-Level Bureaucracy and Public Accountability', *Public administration*, vol. 85, no. 2, pp. 279–299.

Husbands, C. and Pearce, J. (2012) 'What Makes Great Pedagogy? Nine Claims from Research', National College for Teaching and Leadership.

Hutner, T. L. and Markman, A. B. (2016) 'Proposing an Operational Definition of Science Teacher Beliefs', *Journal of Science Teacher Education*, vol. 27, no. 6, pp. 675–691.

Hyam, H. (2006) *Increasing Uptake of Science Post-16: Report of a Royal Society Conference*, London, Royal Society.

IEA (2007) *TIMSS 2007 Teacher Questionnaire Science Grade 8* [Online]. Available at https://timssandpirls.bc.edu/TIMSS2007/PDF/T07_TeacherQ_Science_G8.pdf (Accessed 29 May 2018).

IEA (2011) *Creating and Interpreting the TIMSS and PIRLS 2011 Context Questionnaire Scales* [Online]. Available at https://timssandpirls.bc.edu/methods/t-context-q-scales.html (Accessed 29 May 2018).

IEA (2014) *TIMSS 2015 Teacher Questionnaire Science Grade 8* [Online]. Available at https://timssandpirls.bc.edu/timss2015/questionnaires/downloads/T15_TQS_8.pdf (Accessed 29 May 2018).

Iksan, Z. H. and Daniel, E. (2016) 'Types of Wait Time During Verbal Questioning in the Science Classroom', *International Research in Higher Education*, vol. 1, no. 1, pp. 72–80.

Ing, M. (2014) 'Gender Differences in the Influence of Early Perceived Parental Support on Student Mathematics and Science Achievement and STEM Career Attainment', *International Journal of Science and Mathematics Education*, vol. 12, no. 5, pp. 1221–1239.

Institute for Government (2012) *The Development of Quasi-Markets in Education*, London, Institute for Government.

IoP (2018) *Why Not Physics? A Snapshot of Girls' Uptake at A-Level*, London, Institute of Physics.

IRIS Connect (2018) *Professional Learning for Teachers* [Online]. Available at https://www.irisconnect.com/uk/what-we-do/teacher-professional-learning/ (Accessed 21 September 2018).

Isaacs, T. (2010) 'Educational Assessment in England', *Assessment in Education: Principles, Policy & Practice*, vol. 17, no. 3, pp. 315–334.

Jarrett, T., Long, R. and Foster, D. (2016) *School funding: Pupil Premium*, House of Commons Briefing Paper, London.

Jarvis, P. and Doherty, J. (2016) 'Continuing Professional Development', in Jarvis, P., George, J., Holland, W., and Doherty, J. (eds), *The Complete Companion for Teaching and Leading Practice in the Early Years*, London, Routledge, pp. 218–240.

Jaume, D. and Willén, A. (2017) *The Long-run Effects of Teacher Strikes: Evidence from Argentina*, Working Paper Series, CEDLAS, Universidad Nacional de La Plata.

Jenkins, E. W. (2000) 'The Impact of the National Curriculum on Secondary School Science Teaching in England and Wales', *International Journal of Science Education*, vol. 22, no. 3, pp. 325–336.

Jensen, E. (2013) 'How Poverty Affects Classroom Engagement', *Educational Leadership*, vol. 70, no. 8, pp. 24–30.

Jerrim, J. (2013) 'The Reliability of Trends over Time in International Education Test Scores: Is the Performance of England's Secondary School Pupils Really in Relative Decline?', *Journal of Social Policy*, vol. 42, no. 02, pp. 259–279.

Jerrim, J. (2017) *Global Gaps: Comparing Socio-Economic Gaps in the Performance of Highly Able UK Pupils Internationally*, London, Sutton Trust.

Jerrim, J. and Shure, N. (2016) *Achievement of 15-Year-Olds in England: PISA 2015 National Report*, London, UCL Institute of Education.

Johnson, B. and Gray, R. (2010) 'A History of Philosophical and Theoretical Issues for Mixed Methods Research', in Tashakkori, A. and Teddlie, C. (eds), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, California, Sage Publications, pp. 69–94.

Johnson, D. W. and Johnson, R. T. (2014) 'Cooperative Learning in 21st Century', *Anales de Psicología*, vol. 30, no. 3.

Johnson, R. B., Onwuegbuzie, A. J. and Turner, L. A. (2007) 'Toward a Definition of Mixed Methods Research', *Journal of Mixed Methods Research*, vol. 1, no. 2, pp. 112–133.

Jones, P. (2003) *Introducing Social Theory*, Cambridge, UK, Polity Press.

Julnes, G. (2018) 'Ethical Issues in Evaluation', in Frey, B. B. (ed), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, California, Sage Publications, pp. 612–616.

Jung, J.-H. and Pinar, W. (2015) 'Conceptions of Curriculum', in Wyse, D., Hayward, L., and Pandya, J. (eds), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, London, SAGE Publications, pp. 29–46.

Kane, T. J., Taylor, E. S., Tayler, J. H. and Wooten, A. L. (2011) 'Identifying Effective Classroom Practices Using Student Achievement Data', *The Journal of Human Resources*, vol. 46, no. 3, pp. 587–613.

Kelly, A. and Downey, C. (2011) 'Professional Attitudes to the Use of Data in England', *School Effectiveness and School Improvement*, vol. 22, no. 4, pp. 415–437.

Kennedy, M. M. (2010) 'Attribution Error and the Quest for Teacher Quality', *Educational Researcher*, vol. 39, no. 8, pp. 591–598.

Kentli, F. D. (2009) 'Comparison of Hidden Curriculum Theories', *European Journal of Educational Studies*, vol. 1, no. 2, pp. 83–88.

Khader, F. (2012) 'Teachers' Pedagogical Beliefs and Actual Classroom Practices in Social Studies Instruction', *American International Journal of Contemporary Research*, vol. 2, no. 1, pp. 73–92.

Khalilzadeh, J. and Tasci, A. D. (2017) 'Large Sample Size, Significance Level, and the Effect Size: Solutions to Perils of Using Big Data for Academic Research', *Tourism Management*, vol. 62, pp. 89–96.

Kind, V. (2009a) 'Pedagogical Content Knowledge in Science Education: Perspectives and Potential for Progress', *Studies in Science Education*, vol. 45, no. 2, pp. 169–204.

Kind, V. (2009b) 'A Conflict in Your Head: An Exploration of Trainee Science Teachers' Subject Matter Knowledge Development and Its Impact on Teacher Self-Confidence', *International Journal of Science Education*, vol. 31, no. 11, pp. 1529–1562.

King, C. (2001) 'The Response of Teachers to New Subject Areas in a National Science Curriculum: The Case of the Earth Science Component', *Science Education*, vol. 85, no. 6, pp. 636–664.

Kirby, P. and Cullinane, C. (2016) *Class Differences: Ethnicity and Disadvantage*, Research Brief, London, Sutton Trust.

Ko, J., Sammons, P. and Bakkum, L. (2014) *Effective Teaching: A Review of Research and Evidence*, Reading, CfBT.

Korthagen, F. (2016) 'Inconvenient Truths About Teacher Learning: Towards Professional Development 3.0', *Teachers and Teaching*, pp. 1–19.

Krathwohl, D. (2002) 'A Revision of Bloom's Taxonomy: An Overview', *Theory into Practice*, vol. 41, no. 4, pp. 212–218.

Kvale, S. (2007) *Doing Interviews*, London, Sage Publications, Ltd.

Lamb, M. (2016) *DIRT in the classroom* [Online]. Available at https://www.tes.com/blog/dirt-classroom (Accessed 23 April 2018).

Lasky, S. (2005) 'A Sociocultural Approach to Understanding Teacher Identity, Agency and Professional Vulnerability in a Context of Secondary School Reform', *Teaching and Teacher Education*, vol. 21, no. 8, pp. 899–916.

Le Maistre, C. and Paré, A. (2010) 'Whatever It Takes: How Beginning Teachers Learn to Survive', *Teaching and Teacher Education*, vol. 26, no. 3, pp. 559–564.

Leander, K. M. and Osborne, M. D. (2008) 'Complex Positioning: Teachers as Agents of Curricular and Pedagogical Reform', *Journal of Curriculum Studies*, vol. 40, no. 1, pp. 23–46.

Leckie, G. and Goldstein, H. (2017) 'The Evolution of School League Tables in England 1992-2016: "Contextual Value-Added", "Expected Progress" and "Progress 8"', *British Educational Research Journal*, vol. 43, no. 3, pp. 193–212.

Legard, R., Keegan, J. and Ward, K. (2003) 'In-Depth Interviews', in *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, pp. 138–169.

Leithwood, K. (2016) 'Department-Head Leadership for School Improvement: Leadership and Policy in Schools', *Leadership and Policy in Schools*, vol. 15, no. 2, pp. 117–140.

Leithwood, K., Harris, A. and Hopkins, D. (2008) 'Seven Strong Claims About Successful School Leadership', *School Leadership & Management*, vol. 28, no. 1, pp. 27–42.

Leithwood, K. and Jantzi, D. (2006) 'Transformational School Leadership for Large-Scale Reform: Effects on Students, Teachers, and Their Classroom Practices: School Effectiveness and School Improvement: Vol 17, No 2', *School Effectiveness and School Improvement*, vol. 17, no. 2, pp. 201–227.

Levin, B. B., He, Y. and Allen, M. H. (2013) 'Teacher Beliefs in Action: A Cross-Sectional, Longitudinal Follow-Up Study of Teachers' Personal Practical Theories', *The Teacher Educator*, vol. 48, no. 3, pp. 201–217.

Levine, T. H. and Marcus, A. S. (2010) 'How the Structure and Focus of Teachers' Collaborative Activities Facilitate and Constrain Teacher Learning.', *Teaching and Teacher Education*, vol. 26, no. 3, pp. 389–398.

Lieberman, A. (2000) 'Networks as Learning Communities: Shaping the Future of Teacher Development', *Journal of Teacher Education*, vol. 51, no. 3, pp. 221–227.

Lincoln, Y. S., Lynham, S. and Guba, E. G. (2018) 'Paradigmatic Controversies, Contradictions, and Emerging Confluences, Revisited', in Denzin, N. K. and Lincoln, Y. S. (eds), *The SAGE Handbook of Qualitative Research*, Handbook of Qualitative Research, 5th edn, London, SAGE Publications, pp. 108–150.

Lingard, B. (2011) 'Policy as Numbers: Ac/counting for Educational Research', *The Australian Educational Researcher*, vol. 38, no. 4, pp. 355–382.

Lipsky, M. (2010) *Street-Level Bureaucracy, 30th Ann. Ed.: Dilemmas of the Individual in Public Service*, New York, Russell Sage Foundation.

Long, R. (2015) *Academies under the Labour Government*, London, House of Commons Library.

Long, R. (2017) *GCSE, AS and A Level Reform (England)*, Briefing Paper, London, House of Commons Library.

Long, R. and Bolton, P. (2017) *English Baccalaureate*, Briefing Paper, House of Commons Library.

Long, R., Foster, D. and Roberts, N. (2018) *Grammar schools in England*, Briefing Paper, House of Commons Library.

Louis, K. S. and Barton, A. C. (2002) 'Tales from the Science Education Crypt: A Critical Reflection of Positionality, Subjectivity, and Reflexivity in Research', *Forum: Qualitative Social Research*, vol. 3, no. 3.

Lourenço, O. (2012) 'Piaget and Vygotsky: Many Resemblances, and a Crucial Difference', *New Ideas in Psychology*, vol. 30, no. 3, pp. 281–295.

Loyens, K. and Maesschalck, J. (2010) 'Toward a Theoretical Framework for Ethical Decision Making of Street-Level Bureaucracy: Existing Models Reconsidered', *Administration & Society*, vol. 42, no. 1, pp. 66–100.

Lumby, J. (2013) 'Distributed Leadership: The Uses and Abuses of Power', *Educational Management Administration & Leadership*, vol. 41, no. 5, pp. 581–597.

Maass, K. and Engeln, K. (2018) 'Impact of Professional Development Involving Modelling on Teachers and Their Teaching', *ZDM*, vol. 50, no. 1–2, pp. 273–285.

MacBeath, J., Galton, M., Edwards, J., Page, C. and Steward, S. (2004) *A Life in Secondary Teaching: Finding Time for Learning: A Report commissioned by the National Union of Teachers concerning the workloads in Secondary Schools*, Cambridge, University of Cambridge.

Machin, S. and McNally, S. (2008) 'The Literacy Hour', *Journal of Public Economics*, vol. 92, no. 5–6, pp. 1441–1462.

Macleod, S., Sharp, C., Bernadinelli, D., Skipp, A and Higgins S (2015) *Supporting the Attainment of Disadvantaged Pupils: Articulating Success and Good Practice*, Research Report, London, DfE.

Magliaro, S. G., Lockee, B. B. and Burton, J. K. (2005) 'Direct Instruction Revisited: A Key Model for Instructional Technology', *Educational Technology Research and Development*, vol. 53, no. 4, pp. 41–55.

Magnusson, S., Krajcik, J. and Borko, H. (2002) 'Nature, Sources, and Development of Pedagogical Content Knowledge for Science Teaching', in Gess-Newsome, J. and Lederman, N. G. (eds), *Examining Pedagogical Content Knowledge*, Dordrecht, Kluwer Academic Publishers, vol. 6, pp. 95–132.

Maguire, M., Ball, S. J. and MacRae, S. (2001) '"In All Our Interests": Internal Marketing at Northwark Park School', *British Journal of Sociology of Education*, vol. 22, no. 1, pp. 35–50.

Maguire, M., Braun, A. and Ball, S. (2015) '"Where You Stand Depends on Where You Sit": The Social Construction of Policy Enactments in the (English) Secondary School', *Discourse: Studies in the Cultural Politics of Education*, vol. 36, no. 4, pp. 485–499.

Mansell, W. (2013) 'Misleading the Public Understanding of Assessment: Wilful or Wrongful Interpretation by Government and Media', *Oxford Review of Education*, vol. 39, no. 1, pp. 128–138.

Mansell, W. (2016) *Academies: Autonomy, Accountability, Quality and Evidence*, CPRT Research Survey, York, University of York.

Mansour, N. (2013) 'Consistencies and Inconsistencies Between Science Teachers' Beliefs and Practices', *International Journal of Science Education*, vol. 35, no. 7, pp. 1230–1275.

Marsden, D. (2015) *Teachers and Performance Pay in 2014: First Results of a Survey*, CEP Discussion Paper, London, Centre for Economic Performance, London School of Economics and Political Sciences.

Martin, M. O., Mullis, I. V., Foy, P. and Arora, A. (2011) 'Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales', *Methods and Procedures in TIMSS and PIRLS*, pp. 1–11.

Martin, M. O., Mullis, I. V. S., Hooper, M., Yin, L., Foy, P. and Palazzo, L. (2015) 'Creating and Interpreting the TIMSS 2015 Context Questionnaire Scales', *Methods and Procedures in TIMSS*.

Marzano, R. J. (2010) *Designing & Teaching Learning Goals & Objectives*, Bloomington, Indiana, Solution Tree Press.

Mason-Bish, H. (2019) 'The Elite Delusion: Reflexivity, Identity and Positionality in Qualitative Research', *Qualitative Research*, vol. 19, no. 3, pp. 263–276.

Matusov, E. and Hayes, R. (2000) 'Sociocultural Critique of Piaget and Vygotsky', *New Ideas in Psychology*, vol. 18, no. 2-3, pp. 215–239.

Maxwell, J. A. (2018) 'Collecting Qualitative Data: A Realist Approach', in *The SAGE Handbook of Qualitative Data Collection*, London, SAGE Publications, pp. 19–31.

Maxwell, J. A. and Mittapalli, K. (2010) 'Realism as a Stance for Mixed Methods Research', in *Handbook of Mixed Methods in Social & Behavioral Research*, pp. 145–168.

Maynard-Moody, S. and Musheno, M. (2012) 'Social Equities and Inequities in Practice: Street-Level Workers as Agents and Pragmatists', *Public Administration Review*, vol. 72, no. s1, pp. S16–S23.

McAnulla, S. (2002) 'Structure and Agency', in Marsh D and Stoker G (eds), *Theory and Methods in Political Science*, Political Analysis, 2nd edn, Basingstoke, Palgrave Macmillan, pp. 271–291.

McCulloch, G. (2016) 'History of the Curriculum', in Wyse, D., Hayward, L., and Pandya, J. (eds), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, London, SAGE Publications, pp. 47–62.

McKnight, L. and Whitburn, B. (2020) 'Seven Reasons to Question the Hegemony of Visible Learning', *Discourse: Studies in the Cultural Politics of Education*, Routledge, vol. 41, no. 1, pp. 32–44.

McLeod, S. A. (2014) *Lev Vygotsky* [Online]. Available at www.simplypsychology.org/vygotsky.html (Accessed 3 December 2018).

McLeod, S. A. (2017) *Behaviorist Approach*. [Online]. Available at https://www.simplypsychology.org/behaviorism.html (Accessed 17 October 2017).

Mcmaster, N. (2017) 'Who Studies STEM Subjects at A-Level and Degree in England? An Investigation into the Intersections Between Students' Family Background, Gender and Ethnicity in Determining Choice', *British Educational Research Journal*, vol. 43, no. 3, pp. 528–553.

Mercer, J. (2010) 'Emotional Beliefs', *International Organization*, vol. 64, no. 1, pp. 1–31.

Mercer, N. and Howe, C. (2012) 'Explaining the Dialogic Processes of Teaching and Learning: The Value and Potential of Sociocultural Theory', *Learning, Culture and Social Interaction*, vol. 1, no. 1, pp. 12–21.

Meyer, H.-D. and Benavot, A. (eds.) (2013) *Pisa, Power, and Policy: The Emergence of Global Educational Governance.*, Oxford, Symposium Books, vol. Oxford Studies in Comparative Education.

Millar, R. (2011) 'Reviewing the National Curriculum for Science: Opportunities and Challenges', *Curriculum Journal*, vol. 22, no. 2, pp. 167–185.

Millar, R. and Osborne, J. (1998) *Beyond 2000: Science Education for the Future: A Report with Ten Recommendations*, London, King's College London, School of Education.

Miller, A. D., Ramirez, E. M. and Murdock, T. B. (2017) 'The Influence of Teachers' Self-Efficacy on Perceptions: Perceived Teacher Competence and Respect and Student Effort and Achievement', *Teaching and Teacher Education*, vol. 64, pp. 260–269.

Miller, K., Lasry, N., Chu, K. and Mazur, E. (2013) 'Role of Physics Lecture Demonstrations in Conceptual Learning', *Physical Review Special Topics - Physics Education Research*, vol. 9, no. 2.

Miller, P. (2011) 'Free Choice, Free Schools and the Academisation of Education in England', *Research in Comparative and International Education*, vol. 6, no. 2, pp. 170–182.

Mingers, J. (2006) *Realising Systems Thinking: Knowledge and Action in Management Science*, New York, Springer Science & Business Media.

Mons, N. (2009) *Theoretical and Real Effects of Standardised Assessment*, EACEA, National Testing of Pupils in Europe: Objectives, Organisation and Use of Results, Brussels, Eurydice Network.

Moolenaar, N. M., Daly, A. J. and Sleegers, P. J. (2010) 'Occupying the principal position: Examining relationships between transformational leadership, social network position, and schools' innovative climate', *Educational administration quarterly*, vol. 46, no. 5, pp. 623–670.

Moore, A. and Clarke, M. (2016) '"Cruel Optimism": Teacher Attachment to Professionalism in an Era of Performativity', *Journal of Education Policy*, vol. 31, no. 5, pp. 666–677.

Morrison, K. (2006) 'Sensitive Educational Research in Small States and Territories: The Case of Macau', *Compare: A Journal of Comparative and International Education*, vol. 36, no. 2, pp. 249–264.

Morrisroe, J. (2014) *Literacy Changes Lives: A New Perspective on Health, Employment and Crime.*, London, National Literacy Trust.

Morse, J. M. (2012) 'The Implications of Interview Type and Structure in Mixed-Method Designs', in Gubrium, J., Holstein, J., Marvasti, A., and McKinney, K. (eds), *The SAGE Handbook of Interview Research: The Complexity of the Craft*, California, SAGE Publications, pp. 193–205.

Muijs, D. (2011) *Doing Quantitative Research with SPSS*, London: Sage.

Muijs, D. and Harris, A. (2007) 'Teacher Leadership in (In)action: Three Case Studies of Contrasting Schools', *Educational Management Administration & Leadership*, vol. 35, no. 1, pp. 111–134.

Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H. and Earl, L. (2014) 'State of the Art – Teacher Effectiveness and Professional Learning', *School Effectiveness and School Improvement*, vol. 25, no. 2, pp. 231–256.

Muijs, D. and Reynolds, D. (2000) 'School Effectiveness and Teacher Effectiveness in Mathematics: Some Preliminary Findings from the Evaluation of the Mathematics Enhancement Programme (Primary)', *School Effectiveness and School Improvement*, vol. 11, no. 3, pp. 273–303.

Muijs, D. and Reynolds, D. (2002) 'Teachers' Beliefs and Behaviors: What Really Matters?', *The Journal of Classroom Interaction*, vol. 37, no. 2, pp. 3–15.

Muijs, D. and Reynolds, D. (2017) *Effective Teaching: Evidence and Practice*, 4th edn, London, Sage.

Mulhall, P., Berry, A. and Loughran, J. (2003) 'Frameworks for Representing Science Teachers' Pedagogical Content Knowledge', *Asia-Pacific Forum on Science Learning and Teaching*, The Education University of Hong Kong, Department of Science and Environmental Studies, vol. 4, pp. 1–25.

Mullis, I. V. S. and Martin, M. O. (eds.) (2013) *TIMSS 2015 Assessment Frameworks*, Boston, International Association for the Evaluation of Educational Achievement TIMSS & PIRLS International Study Center.

Murphy, C., Smith, G., Varley, J. and Razi, O. (2015) 'Changing Practice: An Evaluation of the Impact of a Nature of Science Inquiry-Based Professional Development Programme on Primary Teachers', Boylan, M. (ed), *Cogent Education*, vol. 2, no. 1.

Murphy, R. (2013) 'Testing Teachers: What Works Best for Teacher Evaluation and Appraisal', Sutton Trust.

NASUWT (2016) *The Use of Pupil Performance Data in Target Setting and in the Evaluation of the Effectiveness and Capability of Teachers (England)*, Birmingham, NASUWT.

National Careers Service (2019) *Primary School Teacher | Explore Careers* [Online]. Available at https://nationalcareers.service.gov.uk/job-profiles/primary-school-teacher (Accessed 1 July 2019).

National College for School Leadership (2013) 'Implementing Curriculum and Qualifications Reform: Proposals from the 2013 Fellowship Commission', NCTL.

National Statistics (2017) *School workforce in England: November 2016* [Online]. Available at https://www.gov.uk/government/statistics/school-workforce-in-england-november-2016 (Accessed 21 May 2018).

Naylor, S. (2013) 'The New National Curriculum Programme of Study for Primary Science', *Education in Science*, no. 254, pp. 12–13.

Nespor, J. (1987) 'The Role of Beliefs in the Practice of Teaching', *Journal of curriculum studies*, vol. 19, no. 4, pp. 317–328.

Niglas, K. (2010) 'The Multidimensional Model of Research Methodology: An Integrated Set of Continua', in *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, California, SAGE Publications, pp. 215–236.

Nilsson, P. (2014) 'When Teaching Makes a Difference: Developing Science Teachers' Pedagogical Content Knowledge Through Learning Study', *International Journal of Science Education*, vol. 36, no. 11, pp. 1794–1814.

Noden, P. (2000) 'Rediscovering the Impact of Marketisation: Dimensions of Social Segregation in England's Secondary Schools, 1994-99', *British Journal of Sociology of Education*, vol. 21, no. 3, pp. 371–390.

Noden, P., Great Britain and Department for Education and Skills (2007) *Evaluation of the Two Year Key Stage 3 Project*, Nottingham, DfES Publications.

Noyes, A. (2010) 'Researching with Large Datasets Learning to Think Big When Small Is Beautiful', in *The Routledge Doctoral Student's Companion: Getting to Grips with Research in Education and the Social Sciences*, London; New York, Routledge, pp. 213–220.

Oates, T. (2011) 'Could Do Better: Using International Comparisons to Refine the National Curriculum in England', *The Curriculum Journal*, vol. 22, no. 2, pp. 121–150.

OECD (2000) *Science, Technology and Innovation in the New Economy*, Paris, Organisation for Economic Co-operation and Development.

OECD (2005) 'Teachers Matter Education and Training Policy Attracting, Developing and Retaining Effective Teachers Overview', OECD Publishing.

OECD (ed.) (2010) *Learning for Jobs: Synthesis Report*, OECD Reviews of Vocational Education and Training, Paris, OECD.

OECD (2012) *TALIS Teacher Questionnaire MS-12-01* [Online]. Available at https://www.mecd.gob.es/inee/dam/jcr:4198b3b5-d7bc-410e-96c7-4fb97d5490f9/talis-2008-teacher-questionnaire.pdf (Accessed 29 May 2018).

OECD (2013a) *Teaching and Learning International Survey TALIS 2013: Conceptual Framework*, Paris, OECD.

OECD (2013b) 'Teaching and Learning International Survey (TALIS) Teacher Questionnaire', OECD.

OECD (2018) *Education at a Glance 2018: OECD Indicators*, Paris, OECD Publishing.

Ofqual (2009) *GCSE Subject Criteria for Science*, Coventry, Ofqual.

Ofqual (2012) *English GCSE 2012*, London, Ofqual.

Ofqual (2013) *Review of Controlled Assessment in GCSEs*, London, Ofqual.

Ofqual (2014) *Ofqual Consultation on Setting the Grade Standards of new GCSEs in England.*, London, Ofqual.

Ofqual (2015a) *Assessment of Practical Work in GCSE Science Regulatory Impact Assessment*, London, Ofqual.

Ofqual (2015b) *Assessment of Practical Work in New Science GCSEs – Summary*, London, Ofqual.

Ofqual (2015c) *Ofqual Investigation into OCR's Summer 2014 GCSE and A Level Marking Issues – Findings*, London, Ofqual.

Ofqual (2015d) *GCSE Subject Level Conditions and Requirements for Combined Science*, London, Ofqual [Online]. (Accessed 9 June 2020).

Ofqual (2016) *Postcard Grading New GCSEs* [Online]. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/537147/Postcard_-_Grading_New_GCSEs.pdf (Accessed 18 November 2016).

OfSTED (2008) *Assessment for Learning: The Impact of National Strategy Support*, London, Ofsted.

Ofsted (2013) *Maintaining Curiosity - A Survey into Science Education in Schools*, London, Ofsted.

Ofsted (2019) 'The Education Inspection Framework', Ofsted.

Oliveira, A. W. (2010) 'Improving Teacher Questioning in Science Inquiry Discussions Through Professional Development', *Journal of Research in Science Teaching*, vol. 47, no. 4, pp. 422–453.

Olson, J. F., Martin, M. O. and Mullis, I. V. (2008) *TIMSS 2007 Technical Report*, Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College.

Opfer, V. D. and Pedder, D. (2011) 'Conceptualizing Teacher Professional Learning', *Review of Educational Research*, vol. 81, no. 3, pp. 376–407.

Opposs, D. (2016) 'Whatever Happened to School-Based Assessment in England's GCSEs and A-Levels?', *Perspectives in Education*, vol. 34, no. 4, pp. 52–61.

O'Reilly, M. and Parker, N. (2013) '"Unsatisfactory Saturation": A Critical Exploration of the Notion of Saturated Sample Sizes in Qualitative Research', *Qualitative research*, vol. 13, no. 2, pp. 190–197.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R. and Duschl, R. (2003) 'What "Ideas-About-Science" Should Be Taught in School Science? A Delphi Study of the Expert Community', *Journal of Research in Science Teaching*, vol. 40, no. 7, pp. 692–720.

Osborne, J. and Dillon, J. (eds.) (2010) *Good Practice in Science Teaching: What Research Has To Say*, 2nd edn, Maidenhead, Berks., McGraw-Hill Education.

Osbourne, C. (2011) *GCSE Science Examinations 2008, 2009 and 2010: Royal Society Summary Report*, London, Royal Society.

Ozga, J. (2013) 'Accountability as a Policy Technology: Accounting for Education Performance in Europe', *International Review of Administrative Sciences*, vol. 79, no. 2, pp. 292–309.

Pajares, M. F. (1992) 'Teachers' Beliefs and Educational Research: Cleaning up a Messy Construct', *Review of Educational Research*, vol. 62, no. 3, pp. 307–332.

Panadero, E. and Brown, G. T. L. (2017) 'Teachers' Reasons for Using Peer Assessment: Positive Experience Predicts Use', *European Journal of Psychology of Education*, vol. 32, no. 1, pp. 133–156.

Panitz, T. (1999) 'Collaborative versus Cooperative Learning: A Comparison of the Two Concepts Which Will Help Us Understand the Underlying Nature of Interactive Learning.', ERIC.

Parameshwaran, M. and Thomson, D. (2015) 'The Impact of Accountability Reforms on the Key Stage 4 Curriculum: How Have Changes to School and College Performance Tables Affected Pupil Access to Qualifications and Subjects in Secondary Schools in England?', *London Review of Education*, vol. 13, no. 2, pp. 157–173.

Park, J. (2017) 'Temperature, Test Scores, and Human Capital Production.', Doctoral Dissertation, Cambridge, Mass., Harvard University [Online]. Available at http://scholar. harvard. edu/files/jisungpark/files/temperature_test_scores_and_human_capital_production_-_j_park_-_2-26-17. pdf.

Park, S., Jang, J.-Y., Chen, Y.-C. and Jung, J. (2011) 'Is Pedagogical Content Knowledge (PCK) Necessary for Reformed Science Teaching? Evidence from an Empirical Study', *Research in Science Education*, vol. 41, no. 2, pp. 245–260.

Parsons, C. (2016) 'Ethnicity, Gender, Deprivation and Low Educational Attainment in England: Political Arithmetic, Ideological Stances and the Deficient Society', *Education, Citizenship and Social Justice*, vol. 11, no. 2, pp. 160–183.

Pavlov, I. P. (1927) *Conditional Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*, 1st edn, London, Oxford University Press.

Pawson, R. (2006) *Evidence-Based Policy: A Realist Perspective*, London, Sage Publications.

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*, London, Sage Publications.

Pearsall, G. (2018) *Fast and Effective Assessment: How to Reduce Your Workload and Improve Student Learning*, Virginia, ASCD.

Pearson (2018) *BTEC Works for Science* [Online]. Available at https://qualifications.pearson.com/content/dam/pdf/btec-works/leaflets/BTEC-Science.pdf (Accessed 1 November 2018).

Pearson (2019) *GCSE Combined Science Specification* [Online]. Available at https://qualifications.pearson.com/content/dam/pdf/GCSE/Science/2016/Specification/GCSE_CombinedScience_Spec.pdf (Accessed 9 June 2020).

Pedder, D. and Opfer, V. D. (2011) 'Are We Realising the Full Potential of Teachers' Professional Learning in Schools in England? Policy Issues and Recommendations from a National Study', *Professional Development in Education*, vol. 37, no. 5, pp. 741–758.

Perera, L. D. H. (2014) 'Parents' Attitudes Towards Science and their Children's Science Achievement', *International Journal of Science Education*, vol. 36, no. 18, pp. 3021–3041.

Perry, T. (2016) 'English Value-Added Measures: Examining the Limitations of School Performance Measurement', *British Educational Research Journal*, vol. 42, no. 6, pp. 1056–1080.

Piaget, J. (1952) *The Origin of Intelligence in Children*, New York, International University Press.

Piaget, J. (1977) 'Problems of Equilibration', in Appel, M. H. and Goldberg, L. S. (eds), *Topics in Cognitive Development: Equilibration: Theory, Research, and Application*, Boston, Springer, pp. 3–13.

PiXL.org.uk (2017) *The PiXL Club - Home* [Online]. Available at https://www.pixl.org.uk/ (Accessed 17 May 2018).

Poland, B. D. (2003) 'Transcription Quality', in *Inside Interviewing: New Lenses, New Concerns*, California, SAGE Publications, pp. 267–287.

Power, S. (1992) 'Researching the Impact of Education Policy: Difficulties and Discontinuities', *Journal of Education Policy*, vol. 7, no. 5, pp. 493–500.

Power, S. and Whitty, G. (1996) 'Teaching New Subjects? The Hidden Curriculum of Marketised Education Systems', *Melbourne Studies in Education*, vol. 37, no. 2, pp. 1–21.

Priestley, M., Biesta, G. and Robinson, S. (2013) 'Teachers as Agents of Change: Teacher Agency and Emerging Models of Curriculum', *Reinventing the Curriculum: New Trends in Curriculum Policy and Practice*, pp. 187–206.

Pritchard, A. (2017) *Ways of Learning: Learning Theories for the Classroom*, 4th edn, Abingdon, Routledge.

Pyhältö, K., Pietarinen, J. and Soini, T. (2014) 'Comprehensive School Teachers' Professional Agency in Large-Scale Educational Change', *Journal of educational change*, vol. 15, no. 3, pp. 303–325.

Pyhältö, K., Pietarinen, J. and Soini, T. (2015) 'Teachers' Professional Agency and Learning – from Adaption to Active Modification in the Teacher Community', *Teachers and Teaching*, vol. 21, no. 7, pp. 811–830.

QCA (2005a) *A Review of GCE and GCSE Coursework Arrangements*, Qualifications and Curriculum Authority [Online]. Available at http://dera.ioe.ac.uk/5746/1/qca-05-1845-coursework-report.pdf (Accessed 17 May 2018).

QCA (2005b) *Science: Changes to the Curriculum from 2006 for Key Stage 4.*, London, Qualifications and Curriculum Authority.

QCA (2006) *A Review of GCSE Coursework* [Online]. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/605708/0106_QCA_courseworkreport.pdf (Accessed 25 October 2018).

QCA (2007) 'Science Programme of Study for Key Stage 4', QCA [Online]. Available at https://www.stem.org.uk/system/files/elibrary-resources/legacy_files_migrated/11571-QCA-07-3345-p_Science_KS4_tcm8-1799.pdf (Accessed 29 June 2020).

QCA (2010) *The National Curriculum: level descriptions for subjects.*, Coventry, QCDA.

Quinton, S. and Smallbone, T. (2010) 'Feeding Forward: Using Feedback to Promote Student Reflection and Learning – a Teaching Model', *Innovations in Education and Teaching International*, vol. 47, no. 1, pp. 125–135.

Raviv, A., Cohen, S. and Aflalo, E. (2019) 'How Should Students Learn in the School Science Laboratory? The Benefits of Cooperative Learning', *Research in Science Education*, vol. 49, no. 2, pp. 331–345.

Reichertz, J. (2014) 'Induction, Deduction, Abduction', in *The SAGE Handbook of Qualitative Data Analysis*, London, SAGE Publications, pp. 123–135.

Reinholz, D. (2016) 'The Assessment Cycle: A Model for Learning Through Peer Assessment', *Assessment & Evaluation in Higher Education*, vol. 41, no. 2, pp. 301–315.

Ritchie, J., Lewis, J., Elam, G., Tennant, R. and Rahim, N. (2013) 'Designing and Selecting Samples', in *Qualitative Research Practice. a Guide for Social Science Students and Researchers*, pp. 111–146.

Roberts, N. (2016a) *FAQs: Academies and Free Schools*, Briefing Paper, London, House of Commons Library.

Roberts, N. (2016b) *The School Curriculum and SATs in England: Reforms Since 2010.*, Briefing Paper, London, House of Commons Library.

Roberts, N. (2018) *The School Curriculum in England*, Briefing Paper, London, House of Commons Library.

Roberts, N. and Bolton, P. (2016) *Changes to School Accountability and 'League Tables' in England in 2016*, Briefing Paper, London, House of Commons Library.

Robertson, S. (2015) 'What Teachers Need to Know About the "Global Education Reform Movement', in Little, G. (ed), *Global Education 'Reform': Building Resistance and Solidarity*, Croydon, Manifesto Press, pp. 10–17.

Robinson, C. and Sebba, J. (2010) 'Personalising Learning Through the Use of Technology', *Computers & Education*, Learning in Digital Worlds: Selected Contributions from the CAL 09 Conference, vol. 54, no. 3, pp. 767–775.

Rømer, T. A. (2019) 'A Critique of John Hattie's Theory of Visible Learning', *Educational Philosophy and Theory*, Routledge, vol. 51, no. 6, pp. 587–598.

Rose, J. (2009) *Independent Review of the Primary Curriculum*, Nottingham, DCSF.

Rosenshine, B. (2012) 'Principles of Instruction: Research-Based Strategies That All Teachers Should Know.', *American Educator*, vol. 36, no. 1, p. 12.

Rosenthal, L. (2004) 'Do School Inspections Improve School Quality? Ofsted Inspections and School Examination Results in the UK', *Economics of Education Review*, vol. 23, no. 2, pp. 143–151.

Roulston, K. (2010) *Reflective Interviewing: A Guide to Theory and Practice*, London, SAGE Publications.

Roulston, K. (2014) 'Analysing Interviews', in *The SAGE Handbook of Qualitative Data Analysis*, London, SAGE Publications, pp. 297–312.

Roulston, K. and Choi, M. (2017) 'Qualitative Interviews', in *The SAGE Handbook of Qualitative Data Collection*, London, SAGE Publications, pp. 233–249.

Rowe, M. B. (1986) 'Wait Time: Slowing down May Be a Way of Speeding Up!', *Journal of Teacher Education*, vol. 37, no. 1, pp. 43–50.

Royal Society (2010) *'State of the Nation' Report on 5-14 Science and Mathematics Education*, State of the Nation, London, Royal Society.

Royal Society (2017) *Royal Society Strategic Plan 2017-2022* [Online]. Available at https://royalsociety.org/-/media/about-us/governance/royal-society-strategic-plan-2017-2022.pdf?la=en-GB&hash=18CF2D1D28569B17459C39B82A2F92F2 (Accessed 9 June 2020).

Royal Society of Chemistry (2014) *Education Resources* [Online]. Available at http://www.rsc.org/resources-tools/education-resources/ (Accessed 12 September 2018).

Rubie-Davies, C., Hattie, J. and Hamilton, R. (2006) 'Expecting the Best for Students: Teacher Expectations and Academic Outcomes', *British Journal of Educational Psychology*, vol. 76, no. 3, pp. 429–444.

Rubin, A. (1985) 'Significance Testing with Population Data', *Social Service Review*, vol. 59, no. 3, pp. 518–520.

Rudd, T. and Goodson, I. (2017) 'Negotiating Neoliberal Education: A Positional Issue and an Oppositional Stance', in Rudd, T. and Goodson, I. (eds), *Negotiating Neoliberalism: Developing Alternative Educational Visions*, Studies in Professional Life and Work, Rotterdam, Sense Publishers, pp. 1–11.

Ryan, G. W. and Bernard, H. R. (2003) 'Techniques to Identify Themes', *Field Methods*, vol. 15, no. 1, pp. 85–109.

Ryder, J. (2015) 'Being Professional: Accountability and Authority in Teachers' Responses to Science Curriculum Reform', *Studies in Science Education*, vol. 51, no. 1, pp. 87–120.

Ryder, J. and Banner, I. (2011) 'Multiple Aims in the Development of a Major Reform of the National Curriculum for Science in England', *International Journal of Science Education*, vol. 33, no. 5, pp. 709–725.

Ryder, J. and Banner, I. (2013) 'School Teachers' Experiences of Science Curriculum Reform', *International Journal of Science Education*, vol. 35, no. 3, pp. 490–514.

Ryder, J., Lidar, M., Lundqvist, E. and Östman, L. (2018) 'Expressions of Agency Within Complex Policy Structures: Science Teachers' Experiences of Education Policy Reforms in Sweden', *International Journal of Science Education*, vol. 40, no. 5, pp. 538–563.

Sachs, J. (2016) 'Teacher Professionalism: Why Are We Still Talking About It?', *Teachers and Teaching*, vol. 22, no. 4, pp. 413–425.

Sadler, P. M. and Good, E. (2006) 'The Impact of Self-and Peer-Grading on Student Learning', *Educational Assessment*, vol. 11, no. 1, pp. 1–31.

Sahlberg, P. (2006) 'Education Reform for Raising Economic Competitiveness', *Journal of Educational Change*, vol. 7, no. 4, pp. 259–287.

Sahlberg, P. (2010) 'Rethinking Accountability in a Knowledge Society', *Journal of Educational Change*, vol. 11, no. 1, pp. 45–61.

Sahlberg, P. (2016) 'Professional Autonomy, Trust and Collaboration in Educators' Work', New College, Oxford, Philosophy of Education Society of Great Britain.

Sammons, P., Gu, Q., Day, C. and Ko, J. (2011) 'Exploring the Impact of School Leadership on Pupil Outcomes: Results from a Study of Academically Improved and Effective Schools in England', *International Journal of Educational Management*, vol. 25, no. 1, pp. 83–101.

Sammons, P., Melhuish, E., Sylva, K., Siraj, I., Toth, K., Smees, R. and Taggart, B. (2014) *Influences on Students' GCSE Attainment and Progress at Age 16*, Effective Pre-School, Primary & Secondary Education Project (EPPSE), London, Institute of Education.

Sandholtz, J. H. and Ringstaff, C. (2014) 'Inspiring Instructional Change in Elementary School Science: The Relationship Between Enhanced Self-efficacy and Teacher Practices', *Journal of Science Teacher Education*, vol. 25, no. 6, pp. 729–751.

Sayer, A. (2010) *Method in Social Science: A Realist Approach*, 2nd edn, London; New York, Routledge.

Scheuermann, J. A. (2018) 'Group vs. Collaborative Learning: Knowing the Difference Makes a Difference', *Faculty Focus | Higher Ed Teaching & Learning* [Online]. Available at https://www.facultyfocus.com/articles/course-design-ideas/group-vs-collaborative-learning-knowing-difference-makes-difference/ (Accessed 4 December 2018).

Schreier, M. (2018) 'Sampling and Generalization', in Flick, U. (ed), *The SAGE Handbook of Qualitative Data Collection*, London, SAGE Publications, pp. 84–97.

Schunk, D. H. (2012) *Learning Theories an Educational Perspective*, 6th edn, Boston, MA, Pearson.

Schwerdt, G. and Wuppermann, A. C. (2011) 'Is Traditional Teaching Really All That Bad? A Within-Student Between-Subject Approach', *Economics of Education Review*, vol. 30, no. 2, pp. 365–379.

Scoppio, G. (2002) 'Common Trends of Standardisation, Accountability, Devolution and Choice in the Educational Policies of England, UK, California, USA, and Ontario, Canada', *Current Issues in Comparative Education*, vol. 2, no. 2, pp. 130–141.

SCORE (2013a) *Resourcing Practical Science at Secondary Level*, [Online]. Available at http://www.score-education.org/media/11805/score%20resourcing%20secondary.pdf (Accessed 14 December 2018).

SCORE (2013b) 'Guidelines for the Content of Key Stage 4 Qualifications', [Online]. Available at http://score-education.org/media/12525/ks4%20guidelines%20final%20version.pdf (Accessed 26 June 2020).

SCORE (2013c) 'Draft National Curriculum for Science SCORE Response to the Department for Education's Revised Draft Science Curriculum for Key Stages 1–3.', [Online]. Available at https://www.rsb.org.uk/images/201308%20score%20covering%20response%20to%20nc%20final.pdf (Accessed 26 June 2020).

SCORE (2013d) 'Reformed GCSE Subject Content: SCORE's Response to the Department for Education Consultation', [Online]. (Accessed 26 June 2020).

Scott, D. (2005) 'Critical Realism and Empirical Research Methods in Education', *Journal of Philosophy of Education*, vol. 39, no. 4, pp. 633–646.

Sharma, M. (2018) 'Seeping Deficit Thinking Assumptions Maintain the Neoliberal Education Agenda: Exploring Three Conceptual Frameworks of Deficit Thinking in Inner-City Schools', *Education and Urban Society*, vol. 50, no. 2, pp. 136–154.

Sharp, C., Macleod, S., Bernardinelli, D., Skipp, A. and Higgins, S. E. (2015) 'Supporting the Attainment of Disadvantaged Pupils: Briefing for School Leaders', Department for Education.

Shaw, I., Newton, D. P., Aitkin, M. and Darnell, R. (2003) 'Do Ofsted Inspections of Secondary Schools Make a Difference to GCSE Results?', *British Educational Research Journal*, vol. 29, no. 1, pp. 63–75.

Shepherd, J. (2010) 'Students sat "easy" GCSE science exams', *The Guardian*, 27th August [Online]. Available at http://www.theguardian.com/education/2010/aug/27/science-gcses-too-easy (Accessed 4 September 2017).

Shirazi, S. (2017) 'Student Experience of School Science', *International Journal of Science Education*, vol. 39, no. 14, pp. 1891–1912.

Shulman, L. S. (1986) 'Those Who Understand: Knowledge Growth in Teaching', *Educational Researcher*, vol. 15, no. 2, pp. 4–14.

Shute, V. J. (2008) 'Focus on Formative Feedback', *Review of Educational Research*, vol. 78, no. 1, pp. 153–189.

Siemens (2018) *Schemes of Work Key Stage 2* [Online]. Available at https://www.siemens.co.uk/education/en/teachers/teaching-resources/schemes-of-work-ks2.htm (Accessed 12 September 2018).

Simon, S. (2012) 'Effective Continued Professional Development in Science Education', in Bolte, C., Holdbrook, J., and Rauch, F. (eds), *Inquiry-Based Science Education in Europe: Reflections from the PROFILES Project*, Berlin, Freie Universität, pp. 17–24.

Sims, S. (2017) 'What Happens When You Pay Shortage-Subject Teachers More Money? Simulation the Effect of Early-Career Salary Supplements on Teacher Supply in England.', *Education Datalab*, p. 34.

Sinclair, J., Ironside, M. and Seifert, R. (1996) 'Classroom Struggle? Market Oriented Education Reforms and their Impact on the Teacher Labour Process', *Work, Employment and Society*, vol. 10, no. 4, pp. 641–661.

Skinner, B. F. (1938) *The Behavior of Organisms: An Experimental Analysis.*, New York, D. Appleton-Century Company.

Skinner, B. F. (1953) *Science and Human Behavior*, New York, Simon and Schuster.

Skinner, B. F. (1973) 'The Free and Happy Student', *The Phi Delta Kappan*, vol. 55, no. 1, pp. 13–16.

Skott, J. (2015) 'The Promises, Problems and Prospects of Research on Teacher's Beliefs', in Fives, H. and Gill, M. G. (eds), *International Handbook of Research on Teachers' Beliefs*, Abingdon, England, Routledge, pp. 13–30.

Slavin, R. E. (2015) 'Cooperative Learning in Elementary Schools', *Education 3-13*, vol. 43, no. 1, pp. 5–14.

Smart, J. B. and Marshall, J. C. (2013) 'Interactions Between Classroom Discourse, Teacher Questioning, and Student Cognitive Engagement in Middle School Science', *Journal of Science Teacher Education*, vol. 24, no. 2, pp. 249–267.

Smith, E. (2008a) *Using Secondary Data in Educational and Social Research*, Maidenhead, Berks., Open University Press.

Smith, E. (2008b) 'Pitfalls and Promises: The Use of Secondary Data Analysis in Educational Research', *British Journal of Educational Studies*, vol. 56, no. 3, pp. 323–339.

Smith, E. (2018) *Key Issues in Education and Social Justice*, 2nd edn, London, SAGE Publications.

Smith, F., Hardman, F., Wall, K. and Mroz, M. (2004) 'Interactive Whole Class Teaching in the National Literacy and Numeracy Strategies', *British Educational Research Journal*, vol. 30, no. 3, pp. 395–411.

Stahl, G. (2016) 'White Working-Class Male Narratives of "loyalty to Self" in Discourses of Aspiration', *British Journal of Sociology of Education*, vol. 37, no. 5, pp. 663–683.

Standards and Testing Agency (2016) '2014 Science Sampling Methodology Outcomes Paper', Standards and Testing Agency.

Standards and Testing Agency (2017a) '2016 Science Sampling Methodology Outcomes Paper', Standards and Testing Agency.

Standards and Testing Agency (2017b) '2017 to 2018 Teacher Assessment Frameworks at the End of Key Stage 2', Standards and Testing Agency.

Steers, J. (2014) 'Reforming the School Curriculum and Assessment in England to Match the Best in the World – A Cautionary Tale', *International Journal of Art & Design Education*, vol. 33, no. 1, pp. 6–18.

STEM Learning (2018) *What we do | STEM* [Online]. Available at https://www.stem.org.uk/what-we-do (Accessed 18 October 2016).

Stevenson, H. and Wood, P. (2014) 'Markets, Managerialism and Teachers' Work: The Invisible Hand of High Stakes Testing in England', *International Education Journal: Comparative Perspectives*, vol. 12, no. 2.

Stewart, M. (2012) 'Understanding Learning: Theories and Critique', *University Teaching in Focus: A Learning-Centred Approach*, pp. 3–20.

Stobart, G. (2008) *Testing Times: The Uses and Abuses of Assessment.*, London, Routledge.

Stokes, M. (2016) 'NESTA: Making the Most of Data in Schools', NESTA.

Strand, L., Malmberg, L. and Hall, J. (2015) *English as an Additional Language (EAL) and Educational Achievement in England: An Analysis of the National Pupil Database*, Oxford, University of Oxford.

Strand, S. (2010) 'Do Some Schools Narrow the Gap? Differential School Effectiveness by Ethnicity, Gender, Poverty, and Prior Achievement', *School Effectiveness and School Improvement*, vol. 21, no. 3, pp. 289–314.

Strand, S. (2012) 'The White British–Black Caribbean Achievement Gap: Tests, Tiers and Teacher Expectations', *British Educational Research Journal*, vol. 38, no. 1, pp. 75–101.

Strand, S. (2014a) 'Ethnicity, Gender, Social Class and Achievement Gaps at Age 16: Intersectionality and "getting it" for the White Working Class', *Research Papers in Education*, vol. 29, no. 2, pp. 131–171.

Strand, S. (2014b) 'School Effects and Ethnic, Gender and Socio-Economic Gaps in Educational Achievement at Age 11', *Oxford Review of Education*, vol. 40, no. 2, pp. 223–245.

Strand, S. (2015) *Ethnicity, Deprivation and Educational Achievement at Age 16 in England: Trends Over Time*, London, DfE.

Suter, W. (2012) *Introduction to Educational Research: A Critical Thinking Approach*, California, SAGE Publications.

Sutton Trust (2011) *Improving the Impact of Teachers on Pupil Achievement in the UK – Interim Findings*, [Online]. Available at https://www.suttontrust.com/wp-content/uploads/2011/09/2teachers-impact-report-final.pdf (Accessed 20 March 2019).

Swann, M., McIntyre, D., Pell, T., Hargreaves, L. and Cunningham, M. (2010) 'Teachers' Conceptions of Teacher Professionalism in England in 2003 and 2006', *British Educational Research Journal*, vol. 36, no. 4, pp. 549–571.

Taylor, C. (2003) 'The Geography of Choice and Diversity in the "New" Secondary Education Market of England', *Area*, vol. 33, no. 4, pp. 368–381.

Taylor, I. (2007) 'Discretion and Control in Education: The Teacher as Street-level Bureaucrat', *Educational Management Administration & Leadership*, vol. 35, no. 4, pp. 555–572.

Taylor, J. (2018) *How Bad Does a Teacher Recruitment Crisis Have to Be Before the DfE Calls It a Crisis?* [Online]. Available at https://neu.org.uk/blog/how-bad-does-teacher-recruitment-crisis-have-be-dfe-calls-it-crisis (Accessed 20 March 2019).

Taylor, R. (2013) 'Early Entry to GCSE', Centre for Education Research and Policy.

TES (2018) *TES Primary Teaching Resources* [Online]. Available at https://www.tes.com/teaching-resources/hub/primary (Accessed 12 September 2018).

The Children, Schools and Families Committee (2010) *School Accountability: First Report of Session 2009–10 Volume I*, London, House of Commons Children, Schools and Families Committee.

Thomson, D. (2015) 'More Questions Than Answers?', *FFT Education Datalab* [Online]. Available at https://ffteducationdatalab.org.uk/2015/10/more-questions-than-answers/ (Accessed 7 November 2019).

Thomson, D. (2019) 'How Reliable Are Key Stage 2 Tests?', *FFT Education Datalab* [Online]. Available at https://ffteducationdatalab.org.uk/2019/04/how-reliable-are-key-stage-2-tests/.

Thomson, D. and Plaister, N. (2019) 'The 2019 Income Deprivation Affecting Children Index', *FFT Education Datalab* [Online]. Available at https://ffteducationdatalab.org.uk/2019/09/the-2019-income-deprivation-children-index/ (Accessed 2 January 2020).

Thoonen, E. E., Sleegers, P. J., Oort, F. J., Peetsma, T. T. and Geijsel, F. P. (2011) 'How to Improve Teaching Practices: The Role of Teacher Motivation, Organizational Factors, and Leadership Practices', *Educational Administration Quarterly*, vol. 47, no. 3, pp. 496–536.

Thorndike, E. L. (1906) *The Principles of Teaching: Based on Psychology*, London; New York, Routledge.

Thorndike, E. L. (1911) *Animal Intelligence: Experimental Studies*, New York, Macmillan.

Tikly, L. (2015) 'What Works, for Whom, and in What Circumstances? Towards a Critical Realist Understanding of Learning in International and Comparative Education', *International Journal of Educational Development*, vol. 40, pp. 237–249.

Tomei, A., Dawson, E. and Dillon, J. (2013) *A Study of Science, Technology, Engineering and Mathematics Education in the United Kingdom*, Consultant Report Securing Australia's Future STEM: Country Comparisons, Melbourne, Australian Council of Learned Academies.

Tomei, A., Dillon, J. and Dawson, E. (2015) 'United Kingdom: An Example of the Impact of High Stakes Accountability Regimes on STEM Education', in Freeman, B., Marginson, S., and Tytler, R. (eds), *The Age of Stem: Educational Policy and Practice Across the World in Science, Technology, Engineering and Mathematics*, London; New York, Routledge, pp. 161–177.

Toom, A., Pyhältö, K. and Rust, F. O. (2015) 'Teachers' Professional Agency in Contradictory Times', *Teachers and Teaching*, vol. 21, no. 6, pp. 615–623.

Toplis, R. and Allen, M. (2012) '"I Do and I Understand?" Practical Work and Laboratory Use in United Kingdom Schools.', *Eurasia Journal of Mathematics, Science & Technology Education*, vol. 8, no. 1.

Topping, K. J. (2013) 'Peers as a Source of Formative and Summative Assessment', in McMillan, J. H. (ed), *SAGE Handbook of Research on Classroom Assessment*, SAGE Publications, pp. 395–412.

Treagust, D. (2013) 'General Instructional Methods and Strategies', in Abell, S. K., Appleton, K., and Hanuscin, D. L. (eds), *Handbook of Research on Science Education*, London; New York, Routledge, pp. 373–391.

Tsay, M. and Brady, M. (2010) 'A Case Study of Cooperative Learning and Communication Pedagogy: Does Working in Teams Make a Difference?', *Journal of the Scholarship of Teaching and Learning*, vol. 10, no. 2, pp. 78–89.

Tschannen-Moran, M. and Hoy, A. W. (2001) 'Teacher Efficacy: Capturing an Elusive Construct', *Teaching and Teacher Education*, vol. 17, no. 7, pp. 783–805.

Tschannen-Moran, M. and Hoy, A. W. (2007) 'The Differential Antecedents of Self-Efficacy Beliefs of Novice and Experienced Teachers', *Teaching and Teacher Education*, vol. 23, no. 6, pp. 944–956.

Tummers, L. and Bekkers, V. (2014) 'Policy Implementation, Street-level Bureaucracy, and the Importance of Discretion', *Public Management Review*, vol. 16, no. 4, pp. 527–547.

Turner, C. (2017) 'Teachers Union Threatens to Boycott Primary School Tests Weeks After Government "Caved in" Over SATS', *The Telegraph*, 12th April [Online]. Available at https://www.telegraph.co.uk/education/2017/04/12/teachers-threaten-boycott-primary-school-tests/ (Accessed 20 June 2018).

Tymms, P. (2004) 'Are Standards Rising in English Primary Schools?', *British Educational Research Journal*, vol. 30, no. 4, pp. 477–494.

UK Data Service (2019) *UK Data Service › Series* [Online]. Available at https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000108#!/faqs (Accessed 22 November 2019).

UNESCO (2018) *Global Education Monitoring Report 2017/18: Accountability in Education Meeting Our Commitments.*, Paris, UNESCO.

University of Hertfordshire (2018) *What is PSQM?* [Online]. Available at http://www.psqm.org.uk/what-is-psqm (Accessed 15 May 2018).

Vaismoradi, M., Turunen, H. and Bondas, T. (2013) 'Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study', *Nursing & Health Sciences*, vol. 15, no. 3, pp. 398–405.

Vehovar, V., Toepoel, V. and Steinmetz, S. (2016) 'Non-probability Sampling', in *The SAGE Handbook of Survey Methodology*, Handbook, London, SAGE Publications, pp. 329–343.

Verger, A., Steiner-Khamsi, G. and Lubienski, C. (2017) 'The Emerging Global Education Industry: Analysing Market-Making in Education Through Market Sociology', *Globalisation, Societies and Education*, vol. 15, no. 3, pp. 325–340.

Vermunt, J. D., Vrikki, M., van Halem, N., Warwick, P. and Mercer, N. (2019) 'The Impact of Lesson Study Professional Development on the Quality of Teacher Learning', *Teaching and Teacher Education*, vol. 81, pp. 61–73.

Vieluf, S., Kaplan, D., Klieme, E. and Bayer, S. (2012) 'Teaching Practices and Pedagogical Innovation: Evidence from TALIS', OECD Publishing.

Vignoles, A., Sibieta, L. and Allen, R. (2018) *The Characteristics of and Earnings and Outcomes for Physics Teachers*, IFS Briefing Note, London, Gatsby Foundation and Institute for Fiscal Studies.

Vila, L. E. (2000) 'The Non-Monetary Benefits of Education', *European Journal of Education*, vol. 35, no. 1, pp. 21–32.

Villanueva, M., Taylor, J., Therrien, W. and Hand, B. (2012) 'Science Education for Students with Special Needs', *Studies in Science Education*, vol. 48, no. 2, pp. 187–215.

Voerman, L., Korthagen, F. A. J., Meijer, P. C. and Simons, R. J. (2014) 'Feedback Revisited: Adding Perspectives Based on Positive Psychology. Implications for Theory and Classroom Practice', *Teaching and Teacher Education*, vol. 43, pp. 91–98.

Voerman, L., Meijer, P. C., Korthagen, F. A. J. and Simons, R. J. (2012) 'Types and Frequencies of Feedback Interventions in Classroom Interaction in Secondary Education', *Teaching and Teacher Education*, vol. 28, no. 8, pp. 1107–1115.

Vosniadou, S. (2001) *How Children Learn*, Education Practices Series 7, Brussels, International Academy of Education.

Vygotsky, L. S. (1978) *Mind in Society: The Development of Higher Psychological Processes*, Cambridge, Mass., Harvard University Press.

Wallace, C. (2014) 'Overview of the Role of Teacher Beliefs in Science Education', in Evans, R., Luft, J., Czerniak, C., and Pea, C. (eds), *The Role of Science Teachers' Beliefs in International Classrooms: From Teacher Actions to Student Learning*, Sense Publishers, pp. 17–31.

Wallace, C. S. and Priestley, M. R. (2017) 'Secondary Science Teachers as Curriculum Makers: Mapping and Designing Scotland's New Curriculum for Excellence', *Journal of Research in Science Teaching*, vol. 54, no. 3, pp. 324–349.

Wang, V. C. (2012) 'Understanding and Promoting Learning Theories', *International Forum of Teaching and Studies*, vol. 8, no. 2, pp. 5–11.

Waring, M. (1979) 'Background to Nuffield Science', *History of Education*, vol. 8, no. 3, pp. 223–237.

Waters-Adams, S. (2006) 'The Relationship Between Understanding of the Nature of Science and Practice: The Influence of Teachers' Beliefs About Education, Teaching and Learning', *International Journal of Science Education*, vol. 28, no. 8, pp. 919–944.

Watson, J. B. (1913) 'Psychology as the Behaviorist Views It.', *Psychological Review*, vol. 20, no. 2, p. 158.

Webb, R. and Vulliamy, G. (2006) *Coming Full Circle: The Impact of New Labour's Education Policies on Primary School Teachers' Work*, London, Association of Teachers and Lecturers.

Weinstein, M., Blades, D. and Gleason, S. C. (2016) 'Questioning Power: Deframing the STEM Discourse', *Canadian Journal of Science, Mathematics and Technology Education*, vol. 16, no. 2, pp. 201–212.

Weißenrieder, J., Roesken-Winter, B., Schueler, S., Binner, E. and Blömeke, S. (2015) 'Scaling CPD Through Professional Learning Communities: Development of Teachers' Self-Efficacy in Relation to Collaboration', *ZDM*, vol. 47, no. 1, pp. 27–38.

Wellcome Trust (2011) 'Wellcome Trust Primary Science Survey Report', Wellcome Trust, London.

Wellcome Trust (2013) 'Reform of the National Curriculum in England Response by the Wellcome Trust', [Online]. Available at https://wellcome.ac.uk/sites/default/files/wtp052330.pdf (Accessed 11 May 2020).

Wellcome Trust (2014) 'Primary Science is it Missing Out: Recommendations for Reviving Primary Science', Wellcome Trust, London.

Wellcome Trust (2017) *State of the Nation Report of UK Primary Science Education*, Leicester, CFE Research.

West, A., Mattei, P. and Roberts, J. (2011) 'Accountability and Sanctions in English Schools', *British Journal of Educational Studies*, vol. 59, no. 1, pp. 41–62.

Westbrook, J. (2013) *Pedagogy, Curriculum, Teaching Practices and Teacher Education in Developing Countries. Final Report. Education Rigorous Literature Review*, London, EPPI-Centre, Social Science Research Unit, Institute of Education.

Wheater, R., Ager, R. and Burge, B. (2014) 'Achievement of 15-Year-Olds in England: PISA 2012 National Report (OECD Programme for International Student Assessment)', p. 176.

Wheatley, K. F. (2005) 'The Case for Reconceptualizing Teacher Efficacy Research', *Teaching and Teacher Education*, vol. 21, no. 7, pp. 747–766.

Whetton, C. (2009) 'A Brief History of a Testing Time: National Curriculum Assessment in England 1989–2008', *Educational Research*, vol. 51, no. 2, pp. 137–159.

White, E., Dickerson, C. and Mackintosh, J. (2015) *Impact of Royal Society of Chemistry Bursary Funded Primary Science Quality Mark on Primary Science Teaching*, University of Hertfordshire.

White, J. (2016) 'Aims of Education in a Democracy', in Wyse, D., Hayward, L., and Pandya, J. (eds), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, London, SAGE Publications, pp. 63–78.

Whittaker, F. (2019) *DfE Funding Boss: 2019-20 School Budgets Are 'Not Generous'* [Online]. Available at https://schoolsweek.co.uk/next-years-school-budgets-are-not-generous-says-dfe-funding-chief/ (Accessed 23 April 2019).

Whitty, G. (2006) 'Teacher Professionalism in a New Era', *General Teaching Council for Northern Ireland Annual Lecture*, Belfast, p. 17.

Whitty, G. and Edwards, T. (1998) 'School Choice Policies in England and the United States: An Exploration of Their Origins and Significance', *Comparative Education*, vol. 34, no. 2, pp. 211–227.

Whitty, G. and Power, S. (2000) 'Marketization and Privatization in Mass Education Systems', *International Journal of Educational Development*, vol. 20, no. 2, pp. 93–107.

Whitty, G., Power, S. and Edwards, T. (1998) 'The Assisted Places Scheme: Its Impact and Its Role in Privatization and Marketization', *Journal of Education Policy*, vol. 13, no. 2, pp. 237–250.

Wiliam (2017) *Embedded Formative Assessment: (Strategies for Classroom Assessment That Drives Student Engagement and Learning).*, 2nd edn, Bloomington, Indiana, Solution Tree.

Wiliam, D. (2011) *Embedded Formative Assessment*, Bloomington, Indiana, Solution Tree Press.

Wiliam, D. (2017) *Collaborative Learning* [Online]. Available at https://education.gov.scot/improvement/learning-resources/Dylan%20Wiliam%20-%20Collaborative%20learning (Accessed 4 December 2018).

Wilkins, A. (2012) 'School Choice and the Commodification of Education: A Visual Approach to School Brochures and Websites', *Critical Social Policy*, vol. 32, no. 1, pp. 69–86.

Wilkins, C. (2011) 'Professionalism and the Post-Performative Teacher: New Teachers Reflect on Autonomy and Accountability in the English School System', *Professional Development in Education*, vol. 37, no. 3, pp. 389–409.

Wilkins, C. (2015) 'Education Reform in England: Quality and Equity in the Performative School', *International Journal of Inclusive Education*, vol. 19, no. 11, pp. 1143–1160.

Willig, C. (2014) 'Interpretation and Analysis', in *The SAGE Handbook of Qualitative Data Analysis*, London, SAGE Publications Ltd, pp. 136–150.

Wills, G. (2014) 'The Effects of Teacher Strike Activity on Student Learning in South African Primary Schools', *Economic Research Southern Africa Working paper*.

Wilshaw, M. (2016) *HMCI's Commentary: Science and Foreign Languages in Primary School* [Online]. Available at https://www.gov.uk/government/speeches/hmcis-monthly-commentary-may-2016 (Accessed 11 December 2016).

Wilson, J. (2014) *Closing the Gap with the New Primary National Curriculum*, Nottingham, National College for Teaching and Leadership.

Wilson, L., O. (2016) *Anderson and Krathwohl - Understanding the New Version of Bloom's Taxonomy* [Online]. Available at http://thesecondprinciple.com/wp-content/uploads/2014/01/Anderson-and-Krathwohl-revised-10-2016.pdf (Accessed 26 October 2017).

Winter, C. (2017) 'Curriculum Policy Reform in an Era of Technical Accountability: "fixing" Curriculum, Teachers and Students in English Schools', *Journal of Curriculum Studies*, vol. 49, no. 1, pp. 55-74.

Witzel, A. and Reiter, H. (2012) *The Problem-Centred Interview: Principles and Practice*, London, SAGE.

Wolf, A. (2011) *Review of Vocational Education – The Wolf Report*, London, DfE.

Wolfe, B. L. and Haveman, R. H. (2002) 'Social and Nonmarket Benefits from Education in an Advanced Economy', *Conference series-federal reserve bank of Boston*, Federal Reserve Bank of Boston; 1998, vol. 47, pp. 97–131.

Wolfe, R. E., Steinberg, A. and Hoffman, N. (2013) *Anytime, Anywhere: Student-Centered Learning for Schools and Teachers*, Cambridge, Mass., Harvard Education Press.

Wong, V. (2019) 'Authenticity, Transition and Mathematical Competence: An Exploration of the Values and Ideology Underpinning an Increase in the Amount of Mathematics in the Science Curriculum in England', *International Journal of Science Education*, vol. 41, no. 13, pp. 1805–1826.

Woodley, E. (2009) 'Practical Work in School Science-Why Is It Important', *School Science Review*, vol. 91, no. 335, pp. 49–51.

Woolley, C. M. (2009) 'Meeting the Mixed Methods Challenge of Integration in a Sociological Study of Structure and Agency', *Journal of Mixed Methods Research*, vol. 3, no. 1, pp. 7–25.

Wragg, T. (2011) *An Introduction to Classroom Observation (Classic Edition)*, London; New York, Routledge.

Wright, O. (2013) *Ofsted Inspectors Are Helping Schools to Pass Their Tests by Charging* [Online]. Available at http://www.independent.co.uk/news/education/education-news/ofsted-inspectors-are-helping-schools-to-pass-their-tests-by-charging-600-a-day-8530252.html (Accessed 21 September 2018).

Wyatt-Smith, C. M. and Looney, A. (2016) 'Professional Standards and the Assessment Work of Teachers', in *The SAGE Handbook of Curriculum, Pedagogy and Assessment: Two Volume Set*, London, SAGE Publications, pp. 805–820.

Zanotti, T. and DfE (2011) 'School and College Performance Tables Statement of Intent – 2011', DfE.

Zolle, O. (2006) 'Science GCSEs: The Choices for September', *Journal of Biological Education*, vol. 40, no. 2, pp. 63–67.

Zyngier, D. (2015) 'Culture and Pedagogy(ies): (What) Have We Learnt from and Since Alexander 2001', in Wyse, D., Hayward, L., and Pandya, J. (eds), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*, London, SAGE Publications, pp. 171–186.

Appendices

A1: Context
Appendix Table A: Curriculum subjects taught in each key stage

Subject	Primary KS1 & 2	Secondary KS3	Secondary KS4
ENGLISH	✓	✓	✓
MATHS	✓	✓	✓
SCIENCE	✓	✓	✓
History	✓	✓	✗
Geography	✓	✓	✗
Art & Design	✓	✓	✗
Physical Education	✓	✓	✓
Music	✓	✓	✗
Modern Foreign Languages	✓	✓	✗
Computing	✓	✓	✓
Design Technology	✓	✓	✗
Citizenship	✗	✓	✓

Appendix Table B: Summary of major science curriculum policy changes

Year	Policy Description
1988	17 Attainment Targets (ATs) – 409 statements in science for the programme of study (PoS) Weighting of attainment targets: Exploration of science (AT1) - 50% at KS1 - 45% at KS2 - 35% at KS3 - 30% at KS4 Knowledge and understanding of science (ATs 2-17) - 50% at KS1 - 55% at KS2 - 65% at KS3 - 70% at KS4 Grading: KS1 - levels 1 to 3 KS2 – levels 2 to 5 KS3 – levels 3 to 7 KS4 – levels 4 to 10 GCSE Exam board grades A to U
1989	Previous version was considered too large for manageable assessment. A complete revision of the statements of attainment was undertaken, reducing the ATs to five with a total of 178 statements of attainment. AT1 - Scientific investigation AT2 - Life and living processes AT3 - Earth and environment AT4 - Materials and their behaviour AT5 - Energy and its effects
1991	A reduction to four ATs. Weighting of attainment targets at KS1 and KS2: - 50% AT1 - 50% equally shared by ATs 2-4 Weighting of attainment targets at KS3 and KS4: equal weighting for all four attainment targets ATs 1-4. AT1 Scientific investigation; AT2 Life and living processes; AT3 Materials and their properties; AT4 Physical processes KS1 to KS3 given levels 1 to 10 Double science and single science introduced at GCSE, grades A-G
1995	This revision of the National Curriculum builds on the 1991 version and follows Sir Ron Dearing's 1993 review of the whole National Curriculum (Dearing, 1993). AT1 becomes 'Experimental and investigative science', and AT2 becomes 'Life processes and living things'. (DfES, 1991). The 10 levels of the ATs were reorganised to eight levels which run only to the end of KS3 but the new "exceptional performance" would draw on KS4 PoS KS1 to KS3 given levels 1 to 8 (EP) GCSE specifications graded with A*-G

1998	Beyond 2000 (Millar and Osborne, 1998) report published which helped to shape science reform.
1999	Science: The National Curriculum for England (1999) 'Importance of Science' statement introduced and the inclusion of 'Ideas and Evidence in Science' into 'Scientific Enquiry'.
2002	GCSEs in Applied Science introduced. Equivalent to two GCSEs, graded A*A* to GG. Modular in format, with three equally weighted units, 33% externally assessed 66% internally assessed coursework (DfES, 2003).
2004	KS1, KS2 and KS3 remain the same as they were in 1999. This version of the National Curriculum includes only changes to the KS4 Double Award Science (DfES and QCA, 2004). The removal of the structure of Sc1 to Sc4 for the PoS, reduction in prescription and the introduction of the ideas and associated terminology of 'How Science Works' (HSW). KS4 is restructured: How Science Works: 1. Data, evidence, theories and explanations 2. Practical and enquiry skills 3. Communication skills 4. Applications and implications of science Breadth of study 5. Organisms and health 6. Chemical and material behaviour 7. Energy, electricity and radiations 8. Environment, Earth and universe QCA statutory regulations specify that assessment for GCSE should: allow only one retake of assessment units with staged assessments with the better result counting towards the qualification. Reference to single or double science was replaced by GCSE Science and GCSE Additional Science or three separate sciences.
2006	Double award GCSE replaced by a wider range of new GCSEs in the sciences, proving greater choice, flexibility and freedom to combine GCSEs. Science GCSE Courses available: Entry Level Certificate Science, Astronomy, Additional Science, Electronics, Biology, Environmental, Chemistry, Geology, Physics, Psychology Applied Science (double award), Human Physiology & Health Additional Science BTEC First Applied Science, Additional Applied Science 21st Century Science Schools can no longer disapply students from science and there is a statutory entitlement to study at least two science GCSEs. From 2006 (QCA, 2005b) all GCSE specifications graded with A*-G: - must have a minimum of 25 per cent internal assessment and 25 per cent external assessment. - no requirement for a terminal, externally set assessment - examinations can be taken at the end of year 10 or year 11. - Assessments can be multiple choice
2007	A major overhaul of the National Curriculum at KS3 by QCA restructured science in line with other subjects. ATs and PoS for KS1 and KS2 remained the same as the 1999 version. The PoS for KS4 remained the same as the 2004 version with practical skills assessed through coursework accounting for 20% weighting of the final grade (QCA, 2006).

	AT1: How Science Works AT2: Organisms, their behaviour and the environment AT3: Materials, their properties and the Earth AT4: Energy, forces and space As well as its own AT, HSW wording is also present in AT2-4.
2008	Written standard assessment tests end for KS3.
2009	Modular GCSEs introduced for first teaching (Baird et al., 2019).
2010	Written Standard Assessment Tests in science end at KS2. Cohort sampling methods are used but Teacher Assessment data is still collated. Publication of the Importance of Teaching White Paper (DfE, 2010b) which proposed changes to exams taken at the end of the course.
2011	New primary curriculum introduced in response to the Rose Review (2009) focusing on cross-curricular teaching. Bew Report (2011) review of KS2 assessment arrangements and the role of SATs is published. GCSE specifications in science must allocate a weighting of 75 per cent to external assessment and a weighting of 25 per cent to controlled assessment in the overall scheme of assessment. Question papers in science must be targeted at either foundation or higher tier. (Ofqual, 2009) The Wolf Report (2011) recommendations reduced the numbers of non-GCSE qualifications eligible for inclusion in performance measures and ensured that no qualification counted for more than one GCSE in size, resulting in a decrease in the numbers of students entering Applied Science GCSE.
2014	A new more rigorous national curriculum is introduced for teaching at primary school, including teaching evolution in KS2 science. KS2 Teacher Assessment meeting/not meeting expected standards, attainment in science is only reported at the national level. For KS1 and KS2: AT1: Working scientifically AT2: Organisms, their behaviour and the environment AT3: Materials, their properties and the Earth AT4: Energy, forces and space For KS3 and KS4 the PoS is split in equal proportions between Biology, Chemistry and Physics, with the Working Scientifically strand taught through and clearly related to substantive science content (DfE, 2014c)
2016	New science national curriculum for first teaching at KS4 to be examined in 2018. GCSE grades 1 to 9, no internal assessments, terminal examinations at the end of two years consisting of six written examination papers, tiered foundation and higher. Practical skills are assessed throughout the terminal examination contributing 15% of the overall mark. More rigour is introduced including an increase to maths content, e.g. The physics curriculum includes 14 different 3-part equations which must be memorised and used, together with several other examples of equations (Ofqual, 2015d). PoS remains split in equal proportions between Biology, Chemistry and Physics, with the Working Scientifically strand taught through and clearly related to substantive science content (DfE, 2014c).

Appendix Table C: Teacher population characteristics

	Number of Teachers	% Male	% Female
All teachers	503,900		
All teachers full time equivalent	457,300		
Primary	222,400	14	86
Secondary	208,200	36	64
Special or Alternative provision	73,300		
Science Teachers	36,600		

Appendix Table D: Age Profile of Teacher Population by phase

Phase	Age	Percentage
Primary	- under 30 - between 30-50 - over 50	28% 55% 17%
Secondary	- under 30 - between 30-50 - over 50	23% 60% 17%

Appendix Table E: End of key stage assessments by year by core subject

Year	Exams Taken in KS2	Science Exams Taken in KS4
2007-08	R, W, M, S Levels	
2008-09	R, W, M, S Levels	
2009-10	R, W, M, S 5% sample	
2010-11	R, W, M, S 5% sample	GCSE grades Modular/GCSE Equivalents Allowed
2011-12	R, TA: W, M, S 5% sample	GCSE grades Modular/GCSE Equivalents Allowed
2012-13	R, SPaGV, TA: W, M, S 5% sample	GCSE grades Linear/GCSE Equivalents Allowed
2013-14	R, SPaGV, TA: W, M, S 5% sample	GCSE grades Linear/GCSE Equivalents Not allowed
2014-15	R, SPaGV, TA: W, M, S 5% sample	GCSE grades Linear/GCSE Equivalents Not allowed
2015-16	R, SPaGV, TA: W, M, S 5% sample	GCSE grades Linear/GCSE Equivalents Not allowed
2016-17	R, SPaGV, TA: W, M, S 5% sample	GCSE grades Linear/GCSE Equivalents Not allowed Eng. & Maths number grades
2017-18	R, SPaGV, TA: W, M, S 5% sample	New GCSE number grades Linear/GCSE Equivalents Not allowed

Key: R = Reading; W = Writing; M = Maths; S = Science; TA= Teacher Assessment; SPaGV = Spelling, Punctuation, Grammar & Vocabulary

A2: Theories of Learning

Behaviourism

Behaviourism emphasises the role of environmental factors in influencing behaviour (McLeod, 2017). For behaviourist like Watson (1913) we are born knowing nothing, and learning is then achieved through conditioning, where an external stimulus results automatically in an observable response, with no recourse to the internal mechanisms like thinking. The main approaches to the study of behaviour are defined as classical conditioning (Pavlov, 1927), connectionism (Thorndike, 1911) and operant conditioning (Skinner, 1938, 1953).

With classical conditioning, creating supportive settings and non-threatening activities in the classroom, and associating these with new contexts may help to reduce student anxieties and fears of the unfamiliar (Schunk, 2012). Positive emotional reactions can be created by associating learning activities with certain stimuli which encourage students to overcome their reluctance to engage with difficult tasks. For connectionism, successfully repeating the learnt behaviour reinforces patterns and connections; this features in classrooms as rote learning and recall of specific information (Thorndike, 1906) such as the times tables or the order of displacement reactions. However, learning through this mode is thought not to stimulate student's thinking skills. Important to Skinners' theory is that of developing a positive classroom climate by responding to student success – what they do right, rather than their failures -what they do wrong (Skinner, 1973,). However, where a teacher is responsible for a large class, positive reinforcement for every individual student may be difficult to achieve in the limited time of a lesson. Activities which involve IT and electronic devices which feedback to students on their responses to stimulus (problems, questions etc) is considered a means to provide the necessary reinforcement to support behaviour responses (Gredler, 2009).

Critics comment that behaviourism does not account for the human element of conscious thought involved in response to the environment (Schunk, 2012). Neither does it take into account student's individual differences or prior knowledge and understanding of the world around them (Westbrook, 2013). Furthermore, behavioural learning tends not to employ problem solving, reasoning and thinking skills this encourages a surface level of learning (Stewart, 2012).

Constructivism and Social Constructivism

With constructivism, unlike behaviourism, children are not the empty vessels waiting to be filled. They create their own meaning, actively making sense of the world through their own conception of reality (Vosniadou, 2001). This cognitive development progresses through the reorganisation of mental processes as a child matures and experiences the world around them through (de Corte, 2010). Through Stage Theory, Piaget (1952, 1977) conceived knowing and intelligence as a process rather than a static, unchanging entity where children's development progresses through a series of stages from sensorimotor to formal operational abstract thinking. Cognitive development theory describes the qualitative changes in reasoning but not how the specific learning of facts is achieved (Schunk, 2012).

Social constructivism (Vygotsky, 1978) arose out of the acknowledgment that social interactions mediate knowledge construction, it challenged the notion of fixed stages of development and introduced the idea that culture and context influence learning. Two main ideas underpin to Vygotsky's work namely the notion of the More Knowledgeable Other (MKO) and the Zone of Proximal Development (ZPD) (Vygotsky, 1978). The problem-solving activities within the ZPD can be defined as those which are almost within a student's grasp; tasks, ideas and concepts that are challenging but not yet mastered. The MKO is often a teacher, an older adult or even another student

with a better understanding, or higher ability level than the learner in a particular area (McLeod, 2014).

The work of Piaget and Vygotsky has been subject to critique and debate over time (Lourenço, 2012; Matusov and Hayes, 2000). One criticism of Piaget's constructivist theory of learning concerns the claim that cognitive development happens in discrete stages which are the same for all children, ignoring culture race or gender (Gray and MacBlain, 2015).   Among the critics of Vygotsky's theory of ZPD, Chaiklin (2003) suggests that the ZPD does not explain the process of development and does not give a clear account of a student's motivational influences in their learning. Despite the critiques, the legacy of the conceptualisation of learning from both theorists has implications for teacher's practice (Schunk, 2012).

A3: Methodology

Ethics Certificate of Approval

US

**UNIVERSITY
OF SUSSEX**

Social Sciences & Arts C-REC
c-recss@admin.susx.ac.uk

Certificate of Approval	
Reference Number	ER/MJW23/1
Title Of Project	Pupil Assessment, Teacher Perspectives
Principal Investigator (PI):	Marilyn Hall
Student	Marilyn Hall
Collaborators	
Duration Of Approval	2 years, 4 months, 26 days
Expected Start Date	01-Aug-2017
Date Of Approval	01-Aug-2017
Approval Expiry Date	27-Dec-2019
Approved By	Ana Pereira
Name of Authorised Signatory	Liz McDonnell
Date	01-Aug-2017

*NB. If the actual project start date is delayed beyond 12 months of the expected start date, this Certificate of Approval will lapse and the project will need to be reviewed again to take account of changed circumstances such as legislation, sponsor requirements and University procedures.

Please note and follow the requirements for approved submissions:

Amendments to protocol
* Any changes or amendments to approved protocols must be submitted to the C-REC for authorisation prior to implementation.

Feedback regarding the status and conduct of approved projects
* Any incidents with ethical implications that occur during the implementation of the project must be reported immediately to the Chair of the C-REC.

Feedback regarding any adverse(1) and unexpected events(2)
* Any adverse (undesirable and unintended) and unexpected events that occur during the implementation of the project must be reported to the Chair of the Social Sciences and Arts C-REC. In the event of a serious adverse event, research must be stopped immediately and the Chair alerted within 24 hours of the occurrence.

Monitoring of Approved studies
The University may undertake periodic monitoring of approved studies. Researchers will be requested to report on the outcomes of research activity in relation to approvals that were granted (full applications and amendments).

Research Standards
Failure to conduct University research in alignment with the Code of Practice for Research may be investigated under the Procedure for the Investigation of Allegations of Misconduct in Research or other appropriate internal mechanisms (3). Any queries can be addressed to the Research Governance Office: rgoffice@sussex.ac.uk

(1) An "adverse event" is one that occurs during the course of a research protocol that either causes physical or psychological harm, or increases the risk of physical or psychological harm, or results in a loss of privacy and/or confidentiality to research participant or others.

(2) An "unexpected event" is an occurrence or situation during the course of a research project that was a) harmful to a participant taking part in the research, or b) increased the probability of harm to participants taking part in the research.

(3) http://www.sussex.ac.uk/staff/research/rqi/policy/research-policy

Participant Information Sheet

Project title

Pupil Attainment, Teacher Perspectives: Exploring the Evidence for Reform in Science Education in England.

Purpose of the study

As you may know, in recent years there have been a number of significant changes to the content and assessment of the National Curriculum across the Key Stages. These include the withdrawal of the national curriculum tests at Key Stage 3 in 2008; the removal of the national curriculum tests in science for the majority of Key Stage 2 students in 2010 and the overhaul of the National Curriculum in 2014. This study aims to explore the relationship between the changes in the curriculum and its assessment; teachers' agency and professional practice and children's attainment in science.

Using data covering a ten-year period 2008 to 2018, the research will map the major education reforms against student attainment then use the outcome of the analysis as a starting point to explore teachers' perspectives. In addition to conducting on-line questionnaires sent to participants; I will be carrying out interviews with teachers from a small number of schools in order to investigate the extent to which the reforms influence teacher professional practice.

It is anticipated that the data collection phase of the study will run for approximately 6 months, followed by a period of analysis and writing up of the findings, leading to the publication of a doctoral thesis in 2020.

Why should you take part?

Questionnaires and interviews will be carried out with secondary science teachers and primary science coordinators, who like you are in a unique position to share experiences, understanding and actions in relation to the teaching, learning and assessment of this core curriculum subject. Your views are important in developing an understanding of how systematic changes in education over time affect teachers practice. By taking part in this research project you will be contributing to this body of knowledge, the outcomes of

which will then be shared with the wider academic community through publications and conferences.

Taking part in the research is entirely voluntary and it is up to you to decide whether or not to get involved. If you do decide to take part you will be given this information sheet to keep and be asked to sign a <u>consent form</u>, however you are still free to withdraw at any time and without giving a reason. Furthermore, you also have the right to withdraw your data up until the point at which it is no longer practical to do so.

What will happen to me if I take part?

On receipt of your consent form, you will be emailed a link to an on-line self-completion questionnaire, this consists of short-answer closed questions and should take approximately 20 minutes to complete. You are also asked to suggest suitable times for me to visit your school to conduct an interview. This will follow-up your responses to the questionnaire with longer open-ended questions, giving you the opportunity to share and develop your views in greater detail. The interview will take approximately 30-40 minutes and your responses will be audio-recorded.

There will be minimal risk of participants being identified through the information shared with me. All data, audio recordings and information collected from or about individual respondents will be anonymised, kept securely using password protected media and destroyed once anonymised transcripts have been made. Hand-written field-notes and transcripts will also anonymised and safely stored in a locked filing cabinet. Some people may find it difficult talking about aspects of their work which they find challenging, if you find any of the questions or discussions upsetting or intrusive, you will be encouraged to say so and we can change or end the conversation. Should you need further support you should contact: <u>https://www.educationsupportpartnership.org.uk</u>

The only circumstances in which complete confidentiality would not be guaranteed is if a participant discloses something that causes me concern that they or someone else might be at significant risk of harm. In that case, I would let the participant know about any actions that may be taken, except if discussing it with them is likely to cause further harm to those involved.

What should I do if I want to take part?

If you would like to take part in this study, please click this link to complete the <u>consent form</u>. You can print the form and a copy of the information sheet for your personal records.

What will happen to the results of the research study?

Ultimately, the results of this research will be published as part of my doctoral thesis through the School of Education and Social Work at Sussex University. An executive summary of the findings will be shared with the school, however should you wish to obtain a copy of the completed research it will be available through open-access publishing.

Who has approved this study?

The research has been approved by the Social Sciences & Cross-Schools Research Ethics Committee (C-REC) through the School of Education and Social Work ethical review process.

Contact for Further Information

Should you have any queries or require further information please contact me by email or by post.

Email: mjw23@sussex,ac.uk

Post:

Marilyn Hall

Essex House

University of Sussex

Falmer, East Sussex

BN1 9RH

However, if you have any concerns about the way in which the study has been conducted that you do not wish to discuss with me, please contact my supervisor Professor Gillian Hampden-Thompson at G.Hampden-Thompson@Sussex.ac.uk for further advice and information.

The University of Sussex has insurance in place to cover its legal liabilities in respect of this study.

Participant Consent Form

US

UNIVERSITY
OF SUSSEX

Project Title: Pupil Attainment, Teacher Perspectives: Exploring the Evidence for Reform in Science Education in England.

Project Approval Reference: ER/MJW23/1

I agree to take part in the above University of Sussex research project. I have had the project explained to me and I have read and understood the Information Sheet, which I may keep for records.:

Declaration	Please tick
I understand that agreeing to take part means that I am willing to be interviewed by the researcher	
I understand that agreeing to take part means that I am willing to respond to a self-completion questionnaire	
I understand that agreeing to take part means that I am willing to allow the interview to be audio taped	
I understand that any information I provide is confidential, and that no information that I disclose will lead to the identification of any individual in the reports on the project, either by the researcher or by any other party.	
I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and that I can withdraw at any stage of the project without being penalised or disadvantaged in any way.	
I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the Data Protection Act 1998.	

Name _____

School _____

Signature _____

Date _____

Interview Schedule

Project Title: Pupil Attainment, Teacher Perspectives: Exploring the Evidence for Reform in Science Education in England.

Researcher: R (Marilyn Hall)

Participant: P (code)

R/P	**Establish Rapport** Introduction: Shake hands and formally introduce myself.
R	Summary of experience: In addition to being a secondary national strategy consultant, I taught secondary science for almost 20 years, most recently as an assistant head in an East Sussex secondary school.
R	What the study is about: The study explores the interaction between reforms in UK science education, pupil attainment and teachers' practice. As well as conducting questionnaires and interviews with primary and secondary science teachers I will be carrying out secondary data analysis of student attainment data from the National Pupil Database.
R	**Purpose** Process: This will be semi-structured interview when I will be referring to the five themes in online questionnaire to follow-up on your responses.
R	Outcomes: I want to encourage you to explore and reflect on your day-to-day teaching practice. If I don't respond enthusiastically it's because having been a teacher, I'm trying to keep my personal feelings/thoughts out of the interview.
R	**Consent** Reminder: May I remind you that the interview will be recorded, your data will be anonymised and stored securely. Let me also reiterate that you are free to withdraw your consent at any time. You may also withdraw their data up to a point at which it is no longer practical to do so.
R	**Time Line** The interview should take about 30-40 minutes. Are you still willing to go ahead at this time?
R	**Transition:** Let me begin by asking you some questions about your day-to-day teaching
R	**Theme: Pedagogy and Practice** *Question 1: Looking back over the list of activities from this part of the questionnaire, you've indicated that some of the things that you do in class have changed. Can you tell me more about this?* *(Prompt: copy of questionnaire to be shown to participant as a reminder of their responses)* *Follow-up: What do you think has driven this change?*
R	**Theme: Relationships with colleagues and CPD** *Question 2a: Thinking back to the questions regarding CPD, meetings and joint inset. I would like to hear more about how you work with colleagues both inside and outside of your school.*

	Follow-up: Has this changed in any way? Why do you think this is? *Prompt: This can relate to any key stage, area of teaching or new curricula or assessments?* *Question 2b: Can you tell in what ways the CPD has made a difference?*
R	**Theme: Relationship between practice and government policy reforms** *Question 3: Thinking of the changes made to measuring and reporting student attainment, have any of these had an impact on what you do in the classroom?* *Prompt: e.g. levels, scaled scores, GCSE grade 9 to 1 etc.* *Follow-up: Can you give a more detailed description of how this affects your teaching practice?*
R	**Question 4:** *In the questionnaire, you indicated that that several of the policies would have a negative/positive impact. Overall, I'd like to hear your thoughts on whether changes in government education policy has changed the way you work?* *Prompt: Is this with respect to a particular year group or exam? Can you explain in a little more detail why you think this is?*
R	**Summarising** 　　　　Well it has been a pleasure finding out more about you. Let me briefly summarise the information that I have recorded during our interview. 　　　　Your teaching practice has _____. 　　　　This shown by _____. 　　　　You consider that your professional development is _____. 　　　　Your understanding and engagement with the new curriculum and accountability measures is _____. 　　　　You think that the impact of government policy on science teaching has 　　　　_____. 　　　　Have I summed this up correctly?
R	**Maintain Rapport** 　　　　I appreciate the time you have given to take part in this interview. I have covered all of the key points and have all the information required for the study. *Question 5: Finally, is there anything else you think would be helpful for me to know or that you feel would be relevant?*
	Closing 　　　　I will transcribe the interview to produce a written copy. This will be anonymised and stored securely. When all the interviews have been conducted, a full analysis of the data will be carried out to draw out the themes. I will then share an executive summary of the findings with you and all the participants and their schools. 　　　　Thank you again for your input.

Questionnaire Domains

Appendix Table F: Proposed seven domains of teacher practice

Pedagogy	Teaching and Learning Actions
Whole Class Teaching	• I present new topics to the class (lecture-style presentation)
	• At the beginning of the lesson I present a short summary of the previous lesson
	• Engage the whole class in discussions
	• Listen and take notes during presentation by teacher
	• Watch audio-visual presentations
	• Watch a science demonstration
	• Use the homework as a basis for class discussion
Differentiation	• I explicitly state learning goals
	• I give different work to the students that have difficulties learning and/or to those who can advance faster
	• I work 1-to-1 with individual students
	• Allow students to work at their own pace
	• Read from a science textbook in class
	• Work alone to answer textbook or worksheet questions
	• Students work in groups based upon their abilities.
Rote Learning	• I ask my students to remember every step in a procedure
	• Follow specific instructions in an activity or investigation
	• Writing definitions or other short writing assignments
Extending learning	• Pose open-ended questions
	• Help students see connections between science and other disciplines
	• Ask my students to write an essay requiring in-depth explanations or detailed reasoning
	• Read other (non-textbook) science-related materials in class
	• Design or implement their own investigation
	• Work on extended science investigations or projects (a week or more in duration)
	• Students hold a debate and argue for a particular point of view.

Formative Assessment	• I check my students' exercise books.
	• I ask questions to check whether or not the subject matter has been understood.
	• Make formal presentations to the rest of the class
	• Students evaluate and reflect upon their own work
	• Revising for an exam or preparing revision materials
	• Correct assignments and then give feedback to students
	• Have students correct their own homework in class
Summative Assessment	• I administer a test or quiz to assess student learning.
	• Take assessments, tests or exams
	• Prepare written science reports
Practical work	• Organise practical hands-on/laboratory science activities or investigations
	• Do hands-on/laboratory science activities or investigations

Questionnaire: Teacher Perspectives on Change in Science Education v3

Start of Block: About You

UNIVERSITY OF SUSSEX

Q1.1

Q1.2 **Project Title:** Pupil Attainment, Teacher Perspectives: Exploring the Evidence for Reform in Science Education in England.

The questionnaire consists of 5 sections and should take approximately 20 minutes.
Section 1: Biographical Section 2: Classroom Practice and Pedagogy Section 3: Curriculum & CPD Section 4: Understanding and applying attainment data
Section 5: Perspectives on Science Education Reform

Your responses to this questionnaire are completely confidential and will be anonymised. They will be used as the focus of the interview to follow-up the themes and form part of the data analysis.

Q1.3 Please enter your name.

_____

Q1.4 About You

This section is about you and your teaching experience, please select the most appropriate responses.
How would you describe your current role?

- Primary Science Coordinator (1)

- Head of Faculty (2)

- Head of Science (3)

- Head of Department (4)

- Head of Subject (5)

- Main Scale Teacher (6)

- Senior Leadership (7)

- Other (please state) (8) _____

Q1.5 What do you consider to be your subject specialism?

- Biology (1)

- Chemistry (2)

- Physics (3)

- Maths (4)

- Another STEM subject (5)

- Other (please state) (6) _____

Q1.6 Working Hours - How would you describe your role?

- Full Time (1)

- Part Time (2)

- If part-time, what of FTE does this equate to? e.g. 0.6fte (3 days per week) (3) _____

Q1.7 How old are you?

- aged between 22 - 29 (1)

- aged between 30 - 39 (2)

- aged between 40 - 49 (3)

- aged between 50 - 59 (4)

- aged 60+ (5)

- Rather not say (6)

Q1.8 How would you describe your gender?

- Female (1)

- Male (2)

- Gender neutral (3)

- Transgender (4)

- Rather not say (5)

Q1.9

Which key stages do you normally teach (tick all that apply)?

- Key Stage 1 (1)

- Key Stage 2 (2)

- Key Stage 3 (3)

- Key Stage 4 (4)

- Key Stage 5 (5)

Q1.10 Teaching experience - By the end of this academic year (Aug 2018), how many years will you have been teaching? Please round up to whole years.

_____

Q1.11 On average, how many hours of science teaching do you deliver each week? Please round up to whole hours.

_____

End of Block: About You

Start of Block: Classroom Practice and Pedagogy

Q2.1 Classroom Practice and Pedagogy
This next section asks about what happens in the classroom and is split into two parts.
Part A asks that you reflect on your teaching over a typical week this term, think about the kinds of activities that you have carried out with your students.
 Part B in the final column asks that you indicate whether you think this has changed over the past 2 years.

How often do you carry out the following activities in your classroom?

	Every or almost every lesson (1)	Often (2)	Sometimes (3)	Rarely (4)	Never (5)	N/A (6)	Yes, this has changed (7)
I present new topics to the class (lecture-style presentation). (1)	—	—	—	—	—	—	—
I explicitly state learning goals (2)	—	—	—	—	—	—	—
I review with the students the homework they have prepared. (3)	—	—	—	—	—	—	—
I give different work to the students that have difficulties learning and/or to those who can advance faster. (4)	—	—	—	—	—	—	—
I ask my students to suggest or to help plan classroom activities or topics. (5)	—	—	—	—	—	—	—

I ask my students to remember every step in a procedure. (6)	—	—	—	—	—	—	—
I check, by asking questions, whether or not the subject matter has been understood. (7)	—	—	—	—	—	—	—
At the beginning of the lesson I present a short summary of the previous lesson. (8)	—	—	—	—	—	—	—
I check my students' exercise books. (9)	—	—	—	—	—	—	—
I work 1-to-1 with individual students (10)	—	—	—	—	—	—	—
Engage the whole class in discussions (11)	—	—	—	—	—	—	—
Pose open-ended questions (12)	—	—	—	—	—	—	—
I administer a test or quiz to assess student learning. (13)	—	—	—	—	—	—	—
Organise practical hands-on/laboratory science activities or investigations (14)	—	—	—	—	—	—	—

Q2.2 Carrying on from exploring what you do in the classroom, this question is also split into two parts. For Part A reflect on the types of activities that students undertake over the course of a week. With part B in the final column please indicate whether you think this has changed over the past 2 years.

How often do students carry out the following activities in your classroom?

	Every or in almost every lesson (1)	Often (2)	Sometimes (3)	Rarely (4)	Never (5)	N/A (6)	Yes, this has changed (7)
Students work in small groups to come up with a joint solution to a problem or task. (2)	▬	▬	▬	▬	▬	▬	▬
Students work on projects that require at least one week to complete. (3)	▬	▬	▬	▬	▬	▬	▬
Students evaluate and reflect upon their own work (4)	▬	▬	▬	▬	▬	▬	▬
Students work individually with the textbook or worksheets to practice newly taught subject matter (5)	▬	▬	▬	▬	▬	▬	▬

Students hold a debate and argue for a particular point of view which may not be their own. (6)	▬	▬	▬	▬	▬	▬	▬
Students work in groups based upon their abilities. (7)	▬	▬	▬	▬	▬	▬	▬
Students make a product that will be used by someone else. (8)	▬	▬	▬	▬	▬	▬	▬

Q2.3 Thinking about how you use homework for all your classes in science, please select the most appropriate option. How often do you assign homework to your students?

▬ In all or almost all of my lessons (1)

▬ Often (2)

▬ Sometimes (3)

▬ Rarely (4)

▬ Never (5)

▬ N/A (6)

▬ Other (please state) (7) _____

Q2.4 How often do you assign the following kinds of homework?

	Always or almost always (1)	Sometimes (2)	Never or almost never (3)	N/A (4)
Doing problem/question sets (1)	▬	▬	▬	▬
Finding one or more applications of the content covered (2)	▬	▬	▬	▬
Reading from a textbook or supplementary material (3)	▬	▬	▬	▬
Writing definitions or other short writing assignments (4)	▬	▬	▬	▬
Working on projects (5)	▬	▬	▬	▬
Working on small investigations or gathering data (6)	▬	▬	▬	▬
Preparing reports (7)	▬	▬	▬	▬
Other (please state) (8)	▬	▬	▬	▬

Q2.5 How often do you do the following with the science homework?

	Always or almost always (1)	Sometimes (2)	Never or almost never (3)	N/A (4)
Monitor whether or not the homework was completed (1)	–	–	–	–
Correct assignments and then give feedback to students (2)	–	–	–	–
Have students correct their own homework in class (3)	–	–	–	–
Use the homework as a basis for class discussion (4)	–	–	–	–
Use the homework to contribute towards students' grades or marks (5)	–	–	–	–

Q3.1 Curriculum and CPD   This section asks about at your involvement in curriculum development and CPD.

Has your department developed new schemes of learning following the introduction of the new national curriculum?

- Yes (1)

- No (2)

Q3.2 How would you characterise your involvement in producing the new schemes of learning?

- At the forefront (1)

- Heavily involved (2)

- Somewhat involved e.g. writing or reviewing short units of work (3)

- No involvement (4)

- N/A (5)

Q3.3 Have you had professional development in any of the following in the last 12 months? Tick all that apply

- Science content (1)

- Science pedagogy (2)

- Science curriculum (3)

- Integrating IT into science (4)

- Improving students critical thinking skills (5)

- Science assessment (6)

- Addressing individual students' needs (7)

- New grades and/or scaled scores (8)

- Other (please state) (9) _____

Q3.4 How often do you have the following types of interactions with other teachers?

	Very Often (weekly) (1)	Often (monthly) (Q3.4_1)	Sometimes (termly) (Q3.4_2)	Never or almost never (Q3.4_3)	N/A (Q3.4_4)
Work as a group on curriculum developments (Q3.4_1)	▬	▬	▬	▬	▬
Carry out joint lesson planning (Q3.4_2)	▬	▬	▬	▬	▬
Discuss how to teach a particular topic (Q3.4_3)	▬	▬	▬	▬	▬
Carry out joint marking and assessment (Q3.4_4)	▬	▬	▬	▬	▬
Visit another classroom to learn more about teaching (Q3.4_5)	▬	▬	▬	▬	▬
Share teaching experiences (Q3.4_6)	▬	▬	▬	▬	▬
Work with teachers from other phase to ensure continuity in learning (Q3.4_7)	▬	▬	▬	▬	▬
Work together to try out new ideas (Q3.4_8)	▬	▬	▬	▬	▬
Other (please state) (Q3.4_9)	▬	▬	▬	▬	▬

Q3.5 Does your school participate in joint inset/CPD with other schools?

- Yes, often (once or twice a month) (1)

- Yes, sometimes (once or twice a term) (2)

- Yes, rarely (once a year) (3)

- Never (4)

- N/A (5)

- Other (please state) (6) _____

Q3.6 At your joint inset meetings, what are the main topics of discussion (tick all that apply)?

- Curriculum Developments (1)

- Transition Arrangements (2)

- Teaching, Learning & Pedagogy (3)

- Marking and Assessment (4)

- Exam Moderation (5)

- Other (please state) (6) _____

Q4.1 Understanding and Applying Student Attainment Data This question asks you to use the slider to rate your level of confidence on a number of initiatives and measures.

Please rate your confidence levels in the following?

	Little or no confidence	Confident	Highly confident
	0 10 20 30 40 50 60 70 80 90 100		
Using the new KS2 scaled scores? ()			
Using the new KS4 number grades? ()			
Using the new Progress 8 measure? ()			
Using the new attainment 8 measure? ()			
Assessing students in science at KS3? ()			
Using attainment data to monitor progress at KS3? ()			
Your understanding of the school's overall attainment targets? ()			
Assessing students in the new GCSE science specifications ()			
Assessing students in the new A-level science specifications ()			

Q5.1 Science Education Reform
Overall, what impact do you think the following reforms have or will have on what you do in the classroom?

	Extremely positive (1)	Somewhat positive (2)	Neither positive nor negative (3)	Somewhat negative (4)	Extremely negative (5)	N/A (6)
New style GCSEs (1)	-	-	-	-	-	-
New national curriculum (2)	-	-	-	-	-	-
New style A-levels (3)	-	-	-	-	-	-
Focus on literacy and numeracy at KS2 (4)	-	-	-	-	-	-
Changes to the KS2 assessments (5)	-	-	-	-	-	-
New floor targets (6)	-	-	-	-	-	-
Progress 8 (7)	-	-	-	-	-	-
Attainment 8 (8)	-	-	-	-	-	-
Compulsory double science (9)	-	-	-	-	-	-
EBacc (10)	-	-	-	-	-	-
KS2 Scale scores (11)	-	-	-	-	-	-

National Pupil Database

The National Pupil Database (NPD) collates individual student and school level attainment data. It is the foundation for evidence-based accountability measures and is said to support school improvement (DfE, 2015c). There are three tiers of data, the Tier 2 data, containing sensitive information or data which could identify individual children, was accessed through an application process. This required an outline of the project and its purpose; descriptions of how the data would be stored and disposed of a; a commitment to confidentiality over and above the university ethical code of conduct and the production of a current Disclosure and Barring Service certificate (DBS) from both myself and my supervisor (DfE, 2015c). The General Data Protection Regulation which came into force on 25th May 2018 (European Commission, 2018) has temporarily suspended any further release of NPD data to researchers (DfE, 2018f). It was understood that parents were not aware that their children's data was being shared (Defend Digital Me, 2018).

There are three annual data collection procedures, School Census, carried out in January, May and October. Each module of data contains information on individual pupils, their identifiers, characteristics, status and special educational needs, this can then be linked to the individual pupil's attainment outcomes in national curriculum assessments through their Unique Pupil Number (UPN). The data must be thoroughly disposed of after use (DfE, 2015c) so it ceases to be copied to backups, and secure file deletion software should be used so that unerase / undelete utilities cannot recover the data. All media on which NPD / linked data has been processed should be shredded, destroyed using commercial best practice, de-magnetised, or securely erased. In my research, the complete deletion of this key data is a major concern as it will affect how I can reference and locate specific data items in the future. However, as the NPD is a secondary data source and nationally available to all, under certain circumstances other researchers can apply to the DfE should they need to replicate or corroborate my findings.

Coding for Ethnicity:

The school census collects information about the ethnicity of pupils in schools in England. Ethnicity is collected for all pupils and records the ethnicity as stated by the parent / guardian or pupil. Ethnicity is a personal awareness of a common cultural identity and relates to how a person feels and not how others perceive them. It is a subjective decision as to which category a person places himself or herself in and does not infer any other characteristics such as religion or country of origin. Schools must not ascribe any ethnicity to a pupil. This information must come from the parent / guardian or pupil. Where the ethnicity has not yet been collected this is recorded as 'NOBT' (information not yet obtained). Where a parent / guardian or pupil declines to provide ethnicity data, code 'REFU' (refused) is recorded and returned.

The ethnicity codeset reflects categories used in the 2001 national population census, with additional categories for Travellers of Irish heritage, Sri Lankan other and pupils of Gypsy / Roma heritage. The codeset can be found at:

https://www.gov.uk/government/publications/common-basic-data-set-cbds-database

National Pupil Database and Data Sharing team

Department for Education

Schedule to the Agreement for the supply of NPD Data
Request No: DR170412.01

1	Name of Requester	University of Sussex, School of Education and Social Work, Department of Education
2	Contact name	Gillian Hampden-Thompson (Tutor) for Marilyn Hall (PhD Student)
3	Address	Essex House Falmer Brighton BN19RH
4	Telephone number	
5	Contact e-mail address	
6	DfE contact	Janet Crame
7	DfE contact e-mail address	NPD.requests@education.gov.uk
8	Commencement Date	23 August 2017
9	Licence End Date	31 August 2020
10	NPD Data	Data to be sent to: Marilyn Hall T4 KS2 Pupil and exam 2013/14 -2015/16 final; linked to T4 Spring Census for corresponding years, plus T2 listed below. T4 KS4 pupil and exam 2010/11 -2015/16 final, with T4 KS3 and T4 KS2 prior attainment; linked to T4 Spring Census for corresponding years, plus T2 listed below. Tier 4 KS5 student and exam 2012/13 - 2015/16 final with T4 KS4 and T4 KS2 prior attainment; linked to T4 Spring Census for corresponding years, plus T2 listed below. **Tier 2 fields** KS2_IDACI KS2_FSM KS2_SENELK

		KS2_SENPS
		KS2_SENA
		KS2_FLANG
		KS2_WIRI
		KS2_WIRT
		KS2_WROM
		KS2_WOTH
		KS2_MWBC
		KS2_MWBA
		KS2_MWAS
		KS2_MOTH
		KS2_AIND
		KS2_APKN
		KS2_ABAN
		KS2_AOTH
		KS2_BCRB
		KS2_BAFR
		KS2_BOTH
		KS2_CHNE
		KS2_OOTH
		KS2_UNCLA
		KS4_IDACI
		KS4_FSM
		KS4_SENK
		KS4_SENPS
		KS4_SENA
		KS4_FLANG
		KS4_WIRI
		KS4_WIRT
		KS4_WROM
		KS4_WOTH
		KS4_MWBC
		KS4_MWBA
		KS4_MWAS
		KS4_MOTH

		KS4_AIND
		KS4_APKN
		KS4_ABAN
		KS4_AOTH
		KS4_BCRB
		KS4_BAFR
		KS4_BOTH
		KS4_CHNE
		KS4_OOTH
		KS4_UNCLA
		KS5_FSM
		EthnicGroupMajor
		FSMeligible
		LanguageGroupMajor
		SENprovision
		IDACIScore
11	Expected timescale for delivery of NPD Data (from receipt of this signed Schedule)	2-3 weeks
12	Permitted Use	Research to support a PhD in Education. Focussing on the progression of pupils from Key Stage 2 to Key Stage 3 in Science in the light of changes to assessment practices. The study uses a two-stage combined methods approach, utilising secondary data analysis of large-scale national level data (NPD) and in-depth qualitative methods to establish the interconnection between performance and progression in science and subject curriculum and assessment reforms.
13	Permitted User(s)	Marilyn Hall Gillian Hampden-Thompson
14	Special conditions (if any)	**Publication** The Requester shall not reproduce the data or include secondary analysis of the data within any publication without the prior written consent of the Data Controller unless such reproduction / publication is included within the Permitted Use.

		Consent is only valid if given in writing, in advance of reproduction or publication of the data. You must inform us of your intent to publish your research and/or analysis 2 working days or, where that is not possible, as early as possible prior to publication being released to: NPD.REQUESTS@education.gov.uk Please include the reference DR170412.01 in the subject line.

This document is the Schedule to the Agreement for the Supply of NPD Data, a copy of which has been supplied to the Requester ("the Agreement"). In signing this Schedule, the parties are agreeing to the terms and conditions set out in the Agreement, including this Schedule.

NPD Codes

Appendix Table G: Key stage 2 teacher attainment level codes reported in the NPD

1 = Achieved Level 1

2 = Achieved Level 2

3 = Achieved Level 3

4 = Achieved Level 4

5 = Achieved Level 5

6 = Achieved Level 6

A = Absent

B = Working below the level assessed by the test

D = Disapplied

F = KS2 pupil not at end of KS2 and taking this subject in future years

L = Left

M = Missing

N = No test level awarded

P = Results for this subject found in previous year's dataset

Q = Malpractice

T = Working at the level of the tests but not able to access them

W = Working towards level 1

X = Lost

Y = DfE ineligible

Z = Ineligible

_X = Invalid Code entered absent, working below

Appendix Table H: National pupil database reference codes for the end of key stage attainment measures

	Measure	Dates	Code	Description	Response
KS2	KS2 Exam	2008-2009	KS2_SCILEV	National Curriculum level awarded for Science test.	See above for codes
	KS2 TA levels	2009/10 - 20014/15	KS2_SCILEVTA	National Curriculum level awarded for Science Teacher Assessment.	
	KS2 Working at expected standard	2014/15 to present	KS2_SCITAEXP	Working at the expected standard in science TA	
KS4	Entered and achieving at triple science	2007/08 - 2012/13	KS4_PASS_ABSCID	Entered Biology, Physics, Chemistry GCSEs or ASs and achieved equivalent of A*-B GCSE in Physics and Chemistry GCSEs or ASs.	0 = No 1 = Yes
		2013/14 only	KS4_PASS_ABSCID_PTQ		
		2014/15 -	PASS_ABSCID_PTQ_EE		
	Entered for triple science	2006/07 - 2012/13	KS4_ALLSCI	Entered all of Biology, Physics, Chemistry GCSE.	
		2013/14 only	KS4_EBALLSCI_PTQ		
		2014/15 -	EBALLSCI_PTQ_EE		
		2015/16 -	KS4_TRIPLESCI_E	Entered biology, physics and chemistry (EBacc qualifications only).	
	Achieving at two sciences	2007/08 - 2012/13	KS4_PASS_ABSCIB	Achieved A*-B in GCSE Science plus GCSE Additional Science or Applied Science or GCSE Land & Environment	
		2013/14 only	KS4_PASS_ABSCIB_PTQ		
		2014/15 -	KS4_PASS_ABSCIB_PTQ_EE		

Appendix Table I: NPD Reference codes for pupil characteristics KS4

NPD Alias	Field Reference	Old Alias	Years Populated	Description	Values
KS4_ACADYR	ACADYR	k4_ac	2001/02 -	Academic year.	
KS4_PupilMatchingRefAnonymous	s.PupilMatchingRefAnonymous	k4_pmr	2001/02 -	Pupil matching reference - Anonymous.	
KS4_UPN	UPN	k4_upn	2002/03 -	Unique Pupil Number.	
KS4_GENDER	GENDER	k4_gend	2001/02 -	Gender.	M = Male F = Female
KS4_FSM	FSM	fsm	2006/07 -	Is pupil known to be eligible for FSM?	0 = No 1 = Yes
KS4_FLANG	FLANG	flang	2006/07 -	Is English not the pupil's first language?	0 = No 1 = Yes
KS4_SENPS	SENPS	senps	2006/07 -	Does pupil have SEN - Action Plus or Statemented?	0 = No 1 = Yes
EthnicGroupMajor_[term][yy]	EthnicGroupMajor	ethgm_[yy]	2008/09 -	AOEG = Any Other Ethnic Group ASIA = Asian BLAC = Black CHIN = Chinese MIXD = Mixed UNCL = Unclassified WHIT = White	Spr

Appendix Table J: NPD Reference codes for pupil characteristics KS2

NPD Alias	Field Reference	Old Alias	Years Populated	Description	Values
KS2_ACADYR	ACADYR	k2_ac	1995/96 -	Academic year.	
KS2_PupilMatchingRefAnonymous	s.PupilMatchingRefAnonymous	k2_pmr	1995/96 -	Pupil matching reference - Anonymous.	
KS2_UPN	UPN	k2_upn	1995/96	Unique Pupil Number.	
KS2_GENDER	GENDER	k2_gend	1995/96 -	Gender.	M = Male F = Female
KS2_FSM	FSM	k2_fsm	2006/07-	Is pupil known to be eligible for FSM?	0 = No 1 = Yes
KS2_SENPS	SENPS	k2_senps	2006/07-2013/14	Does pupil have SEN - Action Plus or Statemented?	0 = No 1 = Yes
KS2_SEN	SEN	k2_sena	2013/14-2014/15	Does pupil have SEN - school action?	0 = No 1 = Yes
KS2_FLANG	FLANG	k2_flang	2006/07-2014/15	Is English not the pupil's first language?	0 = No 1 = Yes
EthnicGroupMajor_[term][yy]	EthnicGroupMajor	ethgm_[yy]	2008/09 -	AOEG = Any Other Ethnic Group ASIA = Asian BLAC = Black CHIN = Chinese MIXD = Mixed UNCL = Unclassified WHIT = White	Spr

A4: Schools, Pupils and their Characteristics: National Data

Appendix Table K: Primary and Secondary School population in England 2008 to 2018

Year	State funded primary schools	State funded secondary schools
2008	4,090,400	3,294,575
2009	4,077,350	3,278,130
2010	4,096,580	3,278,485
2011	4,137,755	3,262,635
2012	4,217,000	3,234,875
2013	4,309,580	3,210,120
2014	4,416,710	3,181,360
2015	4,510,310	3,184,730
2016	4,615,170	3,193,420
2017	4,689,660	3,223,090
2018	4,716,244	3,258,451
2019	4,727,090	3,327,970

Source: (DfE, 2019d)

Student population data can be found from the following government websites.

2008 and 2009: National Archive
https://webarchive.nationalarchives.gov.uk/20080610163225/http://www.dfes.gov.uk/rsgateway/DB/SFR/s000786/index.shtml

2010 to 2018:
https://www.gov.uk/government/collections/statistics-school-and-pupil-numbers

The individual tables give students numbers for the following characteristics FSM, Ethnicity and EAL
    SFR09/2008: Tab3, Tab4
    SFR08/2009: Table 3a, Table 4 and Table 5
    SFR09/2010: Table 3a, Table 4 and Table 5
    SFR12/2011: Table 4, Table 5,
    SFR10/2012: Table 4a, Table 5a
    SFR21/2013: Table 4a, Table 5a
    SFR15/2014: Table 4a, Table 5a
    SFR16/2015: Table 4a, Table 5a
    SFR20/2016: Table 4a, Table 5a
    SFR28/2017: Table 4a, Table 5a
    Schools_Pupils_and_their_Characteristics_2018_National_Tables: Table 3b, Table 4a, Table 5a.

The SEN data can be found at
https://www.gov.uk/government/collections/statistics-special-educational-needs-sen SEN_2018 National: Tables Table 1

A5: Student Attainment Data

Appendix Table L: % of students achieving teacher assessment level 4 and above and level 5 and above in science by year and by gender

	All		Male		Females	
	Level 4+	Level 5+	Level 4+	Level 5+	Level 4+	Level 5+
2008	85	38	84	38	86	38
2009	86	38	85	38	87	38
2010	85	37	84	37	86	36
2011	85	35	83	35	86	35
2012	87	36	85	36	88	36
2013	88	38	86	38	89	38
2014	88	39	87	38	90	39
2015	89	40	87	39	90	40
Mean	87	38	85	37	88	38

Sources: (DfE, 2015b) National Tables SFR30/2015: National Curriculum Assessments at KS2, 2015 Table 2b

Appendix Table M: % of Students achieving teacher assessment level 4 and above in science by characteristic*

Year	All	Boys	Girls	Non-FSM	FSM	Non-SEN	SEN	Non-EAL	EAL	White	Mixed	Asian	Black	Chinese	Other	Missing
2008	88.1	87.1	89.1	90.1	76.1	91.4	58.2	89.0	79.9	89.0	89.1	83.8	82.3	90.6	80.2	89.0
2009	87.7	87.1	88.4	89.6	76.4	90.9	58.3	88.7	79.9	89.1	88.6	83.5	82.0	90.7	80.6	89.1
2010	84.8	83.6	86.0	87.1	70.3	87.1	37.2	85.9	76.4	85.8	85.3	80.7	78.4	90.8	76.0	85.8
2011	84.9	83.5	86.3	87.1	70.2	86.8	25.6	86.0	76.4	85.5	84.9	80.7	78.6	90.1	75.2	85.5
2012	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2013	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2014	88.0	86.5	89.5	90.3	76.6	92.5	49.0	88.9	83.4	88.5	88.4	86.9	85.1	92.3	82.2	88.5
2015	88.7	87.2	90.3	90.8	77.9	92.1	52.7	89.6	84.5	89.2	89.3	88.1	86.5	93.3	83.8	89.2
Mean	87.2	85.8	88.4	89.3	74.9	90.2	51.1	88.2	80.8	87.9	87.6	84.3	82.5	91.4	80.1	87.9

Source: National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

*Note: As discussed in the methodology (chapter 5) end of KS2 Teacher assessment data for the academic years 2011/12 and 2012/13 is missing from the NPD analysis.

Appendix Table N: GCSE and equivalent entries of pupils at the end of key stage 4 2007/08 to 2017/18 in England

KS4 Year	Core	Additional	Biology	Chemistry	Physics
2007/08	516,600	355,600	77,100	68,900	67,900
2008/09	488,000	334,200	92,000	85,000	84,300
2009/10	406,413	288,776	115,888	113,286	112,272
2010/11	350,174	251,794	133,786	132,018	131,138
2011/12	311,231	244,704	149,022	147,464	146,492
2012/13	321,964	248,530	152,424	150,769	149,396
2013/14	352,902	271,202	132,952	129,644	128,740
2014/15	369,209	304,968	126,119	123,378	122,596
2015/16	394,301	349,038	132,064	130,245	129,443
2016/17	388,061	357,911	134,941	133,727	132,963
Mean	389,886	300,672	124,630	121,443	120,524
2017/18	739,406		159,516	156,224	154,768

2007/08: SFR02/2009 GCSE and equivalents in England 2007/08 (provisional) Table 13 found at
https://webarchive.nationalarchives.gov.uk/20120505061657/https://www.educatio n.gov.uk/researchandstatistics/statistics/allstatistics/a00195931/gcse-and-equivalent-results-in-england

2008/09: SFR01/2010 GCSE and equivalents in England 2008/09 (provisional) Table 12 found at
https://webarchive.nationalarchives.gov.uk/20120505061506/https://www.educatio n.gov.uk/researchandstatistics/statistics/allstatistics/a00196306/ks4-results-in-england-2008-09-(revised)

2009/10 to 2014/15: SFR01/2016: GCSE and equivalents in England 2014/15 Revised Subject Time Series Table found at
https://www.gov.uk/government/statistics/revised-gcse-and-equivalent-results-in-england-2014-to-2015

2015/16 to 2017/18: GCSE and equivalents in England 2014/15 Revised Subject Time Series Table found at
https://www.gov.uk/government/statistics/key-stage-4-and-multi-academy-trust-performance-2018-revised

Appendix Table O: % Students achieving 2 good science GCSEs by year 2011-2016 in All Maintained Schools in England by characteristic

Year	All	Boys	Girls	Non-FSM	FSM	Non-SEN	SEN	Non-EAL	EAL	White	Mixed	Asian	Black	Chinese	Other	Missing
2011	48.9	47.2	50.6	52.5	24.2	53.2	11.7	49.3	45.6	48.0	47.3	52.2	40.8	73.4	43.7	56.0
2012	49.7	48.0	51.4	53.4	25.2	53.8	12.3	50.1	46.8	48.7	49.0	53.8	42.2	74.9	47.5	55.9
2013	48.3	45.6	51.0	52.0	24.5	52.1	11.8	48.4	47.5	47.4	49.1	54.1	43.5	75.0	48.5	50.3
2014	47.2	44.1	50.4	50.4	26.0	50.7	12.4	47.0	48.7	48.5	49.5	55.2	45.8	74.7	50.2	30.4
2015	48.4	45.4	51.7	51.5	27.5	50.2	9.9	48.1	50.5	50.2	51.2	56.9	46.9	76.8	52.8	28.1
2016	51.1	47.8	54.6	54.0	30.8	52.4	8.6	50.7	53.8	53.4	53.9	60.4	50.4	80.4	54.4	26.1
Mean	48.8	46.4	51.6	52.3	26.3	52.0	11.6	48.9	49.0	49.3	50.1	55.6	45.0	75.8	49.8	41.7

Source: National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016

Appendix Table P: % of Students entered for triple science GCSE subjects and achieved A*-B in Chemistry or Physics 2011-2016 in All Maintained Schools in England by characteristic

Year	All	Boys	Girls	Non-FSM	FSM	Non-SEN	SEN	Non-EAL	EAL	White	Mixed	Asian	Black	Chinese	Other	Missing
2011	15.1	15.7	14.5	16.7	4.2	16.6	2.0	15.3	13.3	13.2	14.3	17.2	8.8	34.8	13.5	28.3
2012	17.0	17.5	16.5	18.9	5.0	18.6	2.5	17.2	15.6	15.2	15.7	19.8	10.6	40.1	16.0	29.6
2013	16.6	16.6	16.7	18.5	5.1	18.1	2.5	16.7	16.4	14.9	16.1	20.7	11.4	40.0	17.5	26.8
2014	14.0	13.6	14.4	15.5	4.3	15.2	2.3	14.1	13.7	13.8	14.7	18.2	10.0	35.5	14.4	13.0
2015	13.7	13.4	14.0	15.0	4.5	14.2	1.8	13.7	13.6	13.6	14.9	17.8	9.9	34.9	14.1	11.3
2016	14.3	13.8	14.8	15.7	5.0	14.7	1.3	14.2	14.9	14.3	15.4	20.0	11.2	38.9	15.4	9.6
Mean	15.1	15.1	15.2	16.7	4.7	16.2	2.2	15.2	14.6	14.2	15.2	19.0	10.4	37.4	15.2	20.1

Source: National Pupil Database, Spring Census 2008 to 2016 requested from the Department of Education in 2016