

IMPROVEMENT IN WHAT?

**How understanding of school improvement in England evolved
between 1970 and 2020**

David Potter

EdD

2025

IMPROVEMENT IN WHAT?

How understanding of school improvement in England evolved between 1970 and 2020

A thesis submitted to The University of Gloucestershire in accordance with the requirements of the degree of Doctor of Education (EdD) in the School of Education and Science.

David Potter

Submitted January 2025

Word count: 57,723

Abstract

This research investigates three aspects of school improvement, a movement that dominated schooling in England in the last quarter of the twentieth century and the early years of the twenty-first: its history (origins and evolution), how it has worked, and its impact.

Two main methods are used: a survey of scholarly and grey literature and semi-structured interviews with over 30 distinguished leaders, policymakers and researchers in the field. There is an element of autoethnography, using the author's lengthy experience of school improvement research and practice to triangulate the evidence gathered.

The thesis attempts a thorough analysis of the factors contributing to the rise of improvement, classifies them into those from within education and those from the febrile socio-econo-political environment of the time, and shows how they worked together with the new knowledge of school effectiveness to drive improvement. An explanation is offered for the speed with which relevant research, legislation and practice burgeoned.

The literature concerning the 'what' of improvement is more extensive than that concerning the 'how'. The largest section of the thesis is devoted to the views of interviewees concerning how improvement is brought about, particularly how it is led. This analysis leads to the development of a new model for school improvement; most existing models are predicated on the turnaround of a failing school in a challenging setting, whereas this model emphasises the sustained improvement of schools in various contexts.

Analysing the impact of the 50-year history of school improvement is made difficult by the elusive nature of some data. The thesis highlights the neglect in the literature of underlying or fundamental improvement in schools (their improvement as inclusive, humane places for young people to learn in) and contends that Ofsted's potential to promote school improvement has largely been lost through changes in inspection methodology.

As well as contributing to practice through the development of a new model of improvement, the thesis clarifies how school improvement evolved and how leaders bring it about.

Key terms: school effectiveness; school improvement; improvement model; leadership for improvement.

Author's Declaration

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Gloucestershire and is original, except where indicated by specific reference in the text. No part of this thesis has been submitted as part of any other academic award. The thesis has not been presented to any other education institution in the United Kingdom or overseas.

Any views expressed in the thesis are those of the author and in no way represent those of the University.

Signed:

Date: January 2025

DOI: 10.46289/BMUN7859

Acknowledgements

This thesis is dedicated to my beloved late wife, Chrissie, whose love and support sustained me throughout the long time we were together and sustain me still.

Our children, Jonathan and Catherine, have offered constant encouragement and support: discussion with the insights of a skilled education professional from the one, and patient, expert support for my negligible IT skills from the other, and I am very grateful.

I could not have had better supervisors: Dr Paul Vare and Dr Colin Forster provided patient, skilled guidance, constant encouragement and the best kind of criticism: a shrewd blend of frankness, sensitivity and good humour. It is a cliché, but I would not have completed this work without them.

Over thirty colleagues gave generously of their time in interviews, sharing their experiences, joys, frustrations and love of the job of educating young people. I thank them, not just for their contribution to this work but for the huge amount they have done for young people.

Among the interviewees are two friends and colleagues, Jeffrey and Peter, with whom I have spent literally hundreds of hours over decades discussing education. Working with these two knowledgeable, skilled men, passionate about education and young people, has contributed greatly not just to this thesis and my work in education, but to my life. I offer special thanks to Jeff for his help in critically reading parts of the text and for sharing his considerable expertise in political and social history, and in academic writing.

It may be unusual to acknowledge a website, but the astonishing *Education in the UK: the history of our schools, colleges and universities* website compiled by Derek Gillard, has been invaluable.

When the five of us who formed the EdD Class of '19 first met, it was a source of wonder to me that five people could be so disparate in background and yet be within education; we have become good friends and mutual supporters, and I'm grateful to have met them through this programme.

TABLE OF CONTENTS	Page
List of tables and diagrams	vii
List of abbreviations	viii – ix
CHAPTER 1: INTRODUCTION	1 - 8
CHAPTER 2: LITERATURE REVIEW	9 - 54
2.1 Introduction	9
2.2 The process of conducting the review	12
2.3 Definitions	14
2.4 Testing the basic proposition	15
2.5 Review of the literature concerning the evolution of school effectiveness and improvement	18
2.6 Summary: key issues arising from the review, linked to Research Questions	52
CHAPTER 3: METHODOLOGY	55 - 76
3.1 Introduction	55
3.2 Research Design Decisions	56
3.3 Methodological Limitations and their Mitigation	72
3.4 Ethical considerations	72
3.5 Summary of decisions made	74
CHAPTER 4: FINDINGS	74 - 114
4.1 The basic proposition	76
4.2 Perceptions of ‘what is to be improved?’ have evolved and expanded.	84
4.3 Conditions for school improvement	92
4.4 Fundamental improvement has been neglected	109
Reasons for the rapid expansion of improvement are not fully explained	111
4.5 Summary of Findings	114
CHAPTER 5: DISCUSSION	115 - 165
5.1 Testing the basic proposition	116
5.2 The origins and early evolution of school improvement	117
5.3 Improvement in the context of unclear purposes for education	124
5.4 Conditions for Improvement: how school improvement happens	131
5.5 The impact of fifty years of effectiveness and improvement activity and prospects for its future	150

	Page
CHAPTER 6: CONCLUSION	166 - 178
6.1 The extent to which school improvement dominated educational thinking and action in the period	168
6.2 To ascertain why this has been the case	168
6.3 To analyse and explain the rapid evolution of the school improvement paradigm	169
6.4 To identify what key groups of educators have understood by 'improvement', and the actions they have taken as a result.	170
6.5 To identify the impact of fifty years of school improvement, and implications for the future	173
6.6 Unexpected Findings	174
6.7 Summary of Contribution	175
6.8 Limitations of the research	176
6.9 Future research	177
Epilogue	177
 References	 170-213
 Appendices:	 214-221
Semi-structured interviews	
I Information to potential interviewees; invitation to participate	
II Informed Consent Form	
III Interview Questions	
IV Interviewees: codes & one-line CVs	
V Sample post-interview letter/email of thanks to interviewees	
 VI Coding Tree	

LIST OF TABLES AND DIAGRAMS

Table 1	Research questions and relevant literature	p. 10
Table 2	Summary of research questions and methodology	p. 55
Table 3	Structure of findings: themes and sub-themes	p. 77
Table 4	Framework of skills for leading improvement	p. 142
Diagram 1	The Inverse Pyramid Model of School Improvement	p. 145

LIST OF ABBREVIATIONS

A-level Advanced Level (of the GCE, *qv*)

BAME Black, Asian and Minority Ethnic

BTEC Business & Technology Education Council

CEO Chief Education Officer (of a local education authority, *q.v.*) or Chief Executive Officer (of a MAT, *q.v.*)

CPVE Certificate of Pre-Vocational Education

CV Curriculum Vitae

ECM Every Child Matters

DCSF Department for Children, Schools & Families 2008 – 2010) different names

DES Department for Education & Science 1964 - 1992) given over time

DFE Department for Education 1992 - 1995) to the British

DFEE Department for Education and Employment 1995 - 2001) national ministry of

DFES Department for Education & Skills 2001 - 2007) education

EBacc English Baccalaureate

ERA Education Reform Act, 1988

FGM Female Genital Mutilation

GCE General Certificate of Education

GCSE General Certificate of Secondary Education

GNVQ General National Vocational Qualification

HMCI Her/His Majesty's Chief Inspector (of schools)

KCSIE Keeping Children Safe in Education

IDACI Income Deprivation Affecting Children Index

IT Information Technology

LA Local Authority

LEA Local Education Authority

LSES Louisiana School Effectiveness Study

MAT Multi-Academy Trust

NEET Not in Education, Employment or Training

NFER National Foundation for Educational Research

NHS	National Health Service
NPQ	National Professional Qualification
NPQH	National Professional Qualification for Headship
NVQ	National Vocational Qualification
OECD	Organisation for Economic Cooperation and Development
OED	Oxford English Dictionary
Ofsted	Office for Standards in Education
O level	Ordinary level (of the GCE <i>qv</i>)
ONS	Office for National Statistics
OPEC	Organisation of the Petroleum Exporting Countries
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
ROSLA	Raising of the School Leaving Age
SEF	Self-Evaluation Form
SEND	Special Education Needs and Disabilities
SES	Socio-economic Status
SIP	School Improvement Partner
SLT	Senior Leadership Team
SOP	Standard Operating Procedure
STEM	Science, Technology, Engineering and Mathematics
TMSS	Trends in International Maths & Science Study
TQM	Total Quality Management
TVEI	Technical & Vocational Education Initiative
UTC	University Technical College

CHAPTER 1: INTRODUCTION

“Schools try so hard for their pupils now. We feel responsible for them, enjoy their company and joyfully celebrate their successes ... None of that was true in my day Today’s adolescents are welcomed at the gate by staff who enjoy the relationship they have with them. We are at ease with our students ... and they with us. ‘Respect’ now forms a better and more lasting foundation for learning.” (Interviewee JJ ¹)²

Start-point: a personal reflection

I believe schooling in England to be considerably better than it was 50 years ago. I was beaten by teachers, and by prefects aged 17, for trivial offences such as running in a corridor. Radio 4’s Justin Webb describes his secondary school in the 1970s as “not a place of kindness and inclusivity ... for the vulnerable, a hellhole” (Webb, 2022, p.132), “torture both physical and mental” (*ibid*, p.130). Corporal punishment was abolished in maintained schools in England in 1986.

Current teachers strive to meet the needs of individuals: “students³ who sit in the same lessons do not make the same learning gains” (Allen, Evans & White, 2001 p.48). My secondary school had a preparatory department which I, in common with half of my class, had not attended; prep school ‘graduates’ had studied French for three years; I had done none. My first French lesson began with the teacher dictating French words which we were to copy. The first was *le crayon*, which I could both spell and understand, but next came a word for which I could do neither but nonetheless tried to copy. I now know it to have been *virgule*; since it separated each word in the list, I guessed it meant ‘comma’, ceased to write it down and crossed out my earlier attempts, thereby making a mess of the first page of my new exercise book. Half the class was bored stiff; my half was mystified. My first music lesson was also my last: the teacher called each pupil out and required us to sing the notes he played on the piano. I was labelled a ‘groaner’ and sent to the back with the quarter of the class who also groaned. For the remainder of that school year, he spoke not a word to us, engaging only with non-groaners; we did nothing. Because I was in the top stream, I was not taught art,

¹ A key to the anonymised initials of interviewees, with a one-line c.v., appears as Appendix IV.

² When a quotation is in *italics*, it is from an interviewee; when in plain text, it is from literature.

³ The term ‘pupil’ is used generically for all young people in school; ‘student’ is used specifically for those of secondary school age.

music or technology after age 12. Since the 1990s, the study of a broad curriculum until 16 has been mandatory.

Teamwork is an important skill for young people to acquire. In seven years of secondary schooling, I was never asked to carry out a classroom task with another pupil: tasks were conducted individually and in silence. Teaching almost invariably took the form of dictation, copying from the blackboard or answering questions from a textbook, including in four foreign languages.

When I went to university in 1963, I was one of 4% of the nation's school leavers to do so, the proportion rising to 14% by the end of the 1970s. In September 2019, the BBC website headline reported "The symbolic target of 50% at university reached" (BBC, 2019).

I taught the General Certificate of Education (GCE) O level, targeted at the top 20% of the ability range (roughly the grammar school population), and the Certificate of Secondary Education (CSE), introduced in 1965 to enable the next 40% of the ability range to leave school with some qualifications. Nothing was designed for the lowest 40%, with all the inaccuracies and potential biases that figure represents. In 1985, O levels and CSEs were merged into a single examination, GCSE; in 2022, around 94% of all 16-year-olds took at least five GCSEs (Ofqual, 2022).

The education system's disregard for less-able pupils had for many years been widespread. The D stream in my highly selective, four-form-entry grammar school was generally deemed ineducable. There must be a last carriage on the train. "It wasn't until 1970 that disabled children – previously classed as being unable to be educated – were given a right to participate in education" (Mencap, n.d.).

School leaders when I became one were trained only by apprenticeship. One of my interviewees said:

"My experience was probably that of most. I became a senior teacher, deputy and head without any systematic preparation for the roles. What I knew about leadership was based on the various deputies and heads for whom I had worked, and most were poor role models. I learned on the job and made many mistakes." (NN)

Only in the 1990s did the National Professional Qualification for Headship (NPQH) programme provide systematic preparation for headship, followed by programmes for senior leadership.

I could go on: working in partnership with parents in the best interests of pupil and family; removing gender bias in the timetable (woodwork for boys, home economic for girls), increased concern for pupils' safety and wellbeing.

Whether some of these changes constitute 'improvement' may be argued: grade inflation in examinations has been debated for many years, as has whether 50% of school leavers truly benefit from university. But there can be no doubt that the previous low rates of qualification and progression to university represented a huge waste of talent and a considerable bias in favour of more affluent, white students.

Has everything improved? Perhaps not: extra-curricular provision was arguably wider when I was a young teacher. Behaviour in schools may well have deteriorated; as I drafted this text, teachers at a school in Wales were on strike because of the apparent inability of the school's leaders to control the violent, abusive behaviour of some pupils. In March 2024 it was reported that 15% of teachers had recently been sexually harassed, and nearly 20% assaulted, by pupils.

I have come to this research very late after a lengthy career in education: teaching, leadership, administration, governance, inspection and consultancy. I began this study because that career had led me to the conviction of improvement described above, and yet arguments raged about whether schools had truly improved: "Much of what is claimed as school improvement is illusory" wrote Coe (2013, p.i) in a lecture titled "Improving Education: a triumph of hope over experience". When I looked behind the opinions, I found a lack of clarity about what school improvement meant and a narrow focus on one or two measures, especially examination results. The types of 'improvement' I describe above were rarely referred to, hence my title: 'Improvement in What?'

'School improvement' is a term rarely encountered before the 1980s but one which has dominated schools since. It can be said to have its origins around 1970, a time of growing disquiet about education in England: the replacement of grammar and secondary modern by

comprehensive schools, high-profile school failures such as William Tyndale School, and concern over so-called 'progressive' teaching methods all shook public confidence in education. Prime Minister Callaghan's Ruskin Speech (Callaghan, 1976) initiated the education 'Great Debate', opening schools, their curricula and classrooms, to public scrutiny for the first time.

It will be shown that school improvement was born out of the earlier concept of 'school effectiveness', which had a clear beginning when two key publications in 1979 (Rutter and Edmonds) demonstrated that schools have a significant effect on outcomes for their pupils. Subsequent studies, such as Mortimore *et al.* (1988), analysed the attributes of effective schools (those which achieve better outcomes for their students than others with similar intakes, Goldstein 1997), and ineffective schools (Stoll & Myers, 1988; Reynolds, Potter and Chapman, 2003). It is now argued that, in the words of one interviewee, *"school effectiveness has been done, completed"*, and the characteristics of such schools broadly agreed upon. It is also true that "We have learned much about how to improve individual schools" (Hopkins *et al.*, 2014, p.257), the wave of effectiveness research having evolved seamlessly into studies of how those attributes could be used to bring about improvement (Reynolds, 1993). Eventually, detailed practice manuals were produced, 'how to do it' guides to school improvement (Marzano *et al.*, 2014). And yet schools still struggle and are judged by inspectors from the Office for Standards in Education (Ofsted) to require improvement or special measures. *"Individual schools improve but system-wide improvement is lacking"*, said an interviewee.

My experience and a preliminary literature survey led me to identify gaps or contradictions in accounts of school improvement. First, my research is predicated on the widely held view that improvement dominated the school agenda in the period between, roughly, 1980 to 2020: that proposition had never been tested.

The burgeoning of improvement activity was remarkable for its pace and breadth: within 20 years of the Rutter and Edmonds papers, associated research had extended throughout Europe, North America, Australasia and East Asia, in schools of all types and departments within them. Legislation responding to effectiveness research had been enacted in several countries, notably England, where there were significant education acts on average every two years from the late 80s to the early 2000s. Just why this research resonated so widely and

catalysed pressures for change from within education and from wider society is a story only partially told: although many sources list the changes and identify some causes, none provides a comprehensive analysis and classification of the factors at play, nor an explanation of the exceptional speed at which the changes and the spread took place.

School effectiveness literature is clear: although articulated slightly differently by different researchers, the principles are broadly agreed. Were it simply a matter of ‘brute sanity’⁴, school improvement would also be ‘complete’, which it manifestly is not. This study therefore concerns itself with the gap between brute sanity and lived experience, between the known ‘what’ and the lesser-known ‘how’ of improvement. This includes identifying what researchers, administrators and leaders understood by ‘school improvement’. I have long sought to understand the meaning of ‘consistency’ in effectiveness and improvement literature, a term widely used, rarely defined and the impact of which is disputed. This is the thesis stage of a doctoral programme which began with a series of assignments: two of mine were devoted to aspects of consistency and form the basis of parts of the current work.

The pace of education legislation slowed from about 2012. The focus of improvement expanded, with issues concerning pupils’ health, wellbeing and safety gaining greater prominence.

The choice of the period 1970 – 2020 to attempt an account of the evolution and impact of school improvement was driven by:

- the frame created by the publication of the Black Papers (Cox & Dyson 1969-74), a key event in the disquiet leading to the period of rapid improvement, and the pandemic⁵, beginning in 2020, which changed so much in society
- the fit with my career in education
- the neatness of a fifty-year period.

I converted these apparent gaps or questions into five research questions:

⁴ Bernard Shaw’s phrase describing the tendency of those seeking to bring about reform to overlook complexity

⁵ The Covid-19 pandemic of the early 2020s

1. To test the extent to which improvement dominated educational thinking and action in the period.
2. To ascertain why this was the case: what combination of external (socio-econo-political) and internal (educational) factors led education to be dominated by school improvement during the period?
3. To analyse and explain the rapid evolution of the school improvement 'paradigm', identifying how it related to the effectiveness 'paradigm'.
4. To identify what key groups of educators (school leaders, policy makers at local and national level) have understood by 'improvement', and the actions taken as a result.
5. To identify the impact of fifty years of school improvement, and implications for the future.

I believe the study has the potential to add significantly to existing knowledge and practice in the improvement of schools, self-evidently an important topic, by bringing clarity to the evolution of improvement, and analysing its contribution and especially its processes through a new school improvement model based on the views of interviewees and the author's experience.

I am well positioned to undertake the study because of three dimensions of my professional experience. First, breadth: leading a large secondary school, being head of school improvement in a local authority (LA), leading inspections for Ofsted and the Department for Education (DfE), acting as school improvement adviser to governments and over 100 schools, and as governor in primary and secondary schools. This breadth has therefore spanned the roles of creator (research for the DfE leading to multi-academy trusts (MATs)); evaluator (inspector/reviewer for Ofsted and DfE); implementer (in schools and the LA); and interpreter/adviser (as school improvement consultant). And finally, the long period of my influential work coincides with that of intense improvement activity: improvement dominated school thinking from 1980 until the 2010s; I became a senior school leader in the mid-1970s, and a senior leader in the local government of education ten years later. Latterly,

I have spent over 20 years as education consultant/adviser/inspector to schools and governments, local and national, dwindling into semi-retirement from about 2015.

If not uniquely, I contend that this experience positions me distinctively well to undertake this study which, because it is for a professional degree, contributes to practice_as well as to knowledge. My aim is that the findings will point to how, were I back in school or LA leadership, the job could be done better.

At the outset, I recognise two limitations inherent to the study. First, its philosophy mirrors that of the research it analyses: it is inductive, in that conclusions are drawn from examples and case studies; and interpretivist, in that different researchers might interpret the data differently. These approaches have inherent strengths, mainly the ability to draw conclusions and explain patterns, enabling a measure of prediction; and weaknesses, chiefly the potential for inaccuracy through incompleteness. Although 32 is a significant sample and, I will argue, a rich one, a different sample might yield different results. These issues, with an account of attempts at mitigating the risks, are examined in the Methodology chapter. Second, although the literature reviewed is sourced globally, the study relates to England only.

A third limitation derives from the research design: many interviewees are people I know. It was important to achieve a rich sample, so I sought known expertise, while accepting the recommendations of trusted interviewees to introduce others. I have been scrupulous in seeking divergent opinions to offset this potential problem (Methodology chapter).

This preliminary analysis will be revisited in the Conclusion, both to review how the known limitations affected the study and to extend the list.

The thesis is set out conventionally: following this introduction are four major chapters:

2. Literature Review: a survey of current knowledge about effectiveness and improvement, including an analysis of gaps and contradictions.
3. Methodology: an explanation and justification of research design decisions, including the roles played by semi-structured interviews and autoethnography.
4. Findings: the themes (major recurring facets) emerging from analysis of the 32 interviews that constitute my research data, with data extracts to substantiate each theme.

5. Discussion: each section analyses the contribution of one of the themes in the Findings, first to answering the research questions and second to educational knowledge.

The final chapter (6) is the Conclusion: a high-level summary of what the study set out to achieve, its key findings, their implications and contributions to knowledge and practice. An evaluation of the limitations of the research and suggestions for further study are offered. The appendices mainly concern the interviews: interviewees, the information and consent form they were sent, and a list of the topics discussed.

When practitioner-led review blends personal narrative with scholastic analysis, the result can be compelling (for example Pring & Roberts, 2016; Brighouse & Waters, 2021); this thesis will aim to bring the same blend to a specific study of school improvement, its evolution, process and impact.

CHAPTER 2: LITERATURE REVIEW

The review is structured as follows:

2.1 Introduction: purpose of the review; relevance of the thesis theme.

2.2 The process of conducting the review; criteria for selection of sources.

2.3 Definitions of school effectiveness and improvement.

2.4 Testing the basic assertion (research question 1). Underpinning the thesis is the assertion that the period of 50 years from 1970 has been one of intense activity related to effectiveness and improvement, in schools, in government policy and legislation, and in research. This section sets out evidence to test this assertion.

2.5 Review of the literature concerning the evolution of school effectiveness and improvement (research questions 2 and 3, linked): to ascertain the causes of the surge in effectiveness and improvement activity, some from within education, others from a social, political and economic contexts; and to explain the rapidity of evolution of the effectiveness and improvement paradigms.

2.6 Conclusion: key points from the analysis in 2.3 – 2.5 above to inform the research.

The literature provides some evidence toward research question 4, but its nature (what key groups of educators understand by improvement and the actions they took as a result) means that most evidence will come from interviews. The final research question reflects on all evidence gathered. Links between the research questions and the literature reviewed are summarised in Table 1, p.10.

The review approach is therefore thematic and, since there is a focus on change over time, chronological within themes.

2.1 INTRODUCTION

A literature review is both a summary and a critical evaluation of current knowledge in the subject and how it evolved (Hart,1998; Rowley and Slack 2004; Galvan & Galvan 2017). By

Table 1: Research questions and Relevant Literature

RESEARCH QUESTIONS	RELEVANT LITERATURE
To test the extent to which school improvement dominated educational thinking and action in the period.	How frequently the themes of school quality and pupils' standards feature in research journals and government policy/legislation, pre- and post-1970. As well as 'school effectiveness'; and 'school improvement', relevant search terms include 'school quality' and 'pupils' standards'.
To ascertain why this has been the case: what combination of external (socio-econo-political) and internal (educational) factors led education to be dominated by school improvement during this 50-year period?	How major texts on the social history of late 20 th century England, including the history of education, link developments in education with wider social trends. <i>Mutatis mutandis</i> , do meta-reviews of school effectiveness and improvement research seek explanations beyond education?
To analyse and explain the rapid evolution of the school improvement paradigm, identifying how it related to the pre-existing school effectiveness paradigm.	Education research journals, especially those devoted to school effectiveness and improvement. To what do meta-reviews of school effectiveness and improvement research attribute the rapid rise and evolution of the paradigms?
To identify what key groups of educators (researchers, policy makers at local and national level, school leaders) have understood by 'improvement', and the actions they have taken as a result.	Most evidence will come from interviews. Compilations of the reflections of practitioners (Brighouse & Waters, 2021, Pring & Roberts, 2016) plus <i>passim</i> references in the texts referred to above. Grey literature, especially that from government and its agencies such as Ofsted.
To identify the impact of fifty years of school improvement, and implications for the future.	Meta-reviews. Hard data from inspection reports and test/exam results.

identifying gaps and inconsistencies in the findings of previous research, the review justifies subsequent research and ensures that it builds on, rather than repeating, earlier work. Reviews aim for breadth of reference while recognising that completeness is impossible; hence reviews distil the main ideas (Rowley & Slack, 2004; Galvan & Galvan, 2017). Distillation

implies selection of sources, which in turn requires criteria to ensure that those used are of the highest relevance, quality and reliability, including representing a broad spectrum of opinion. These criteria are detailed below [pp.12-13].

The field of this research is **school effectiveness and improvement**, paradigms which, according to the literature, developed from about 1980. Chapter 1 made clear that my involvement in the field began at much the same time and has continued to the present, spanning an era of near-continuous change and drive for improvement in schooling. I defined the research period from 1970, roughly the beginning of my career and the date of publication of the so-called ‘Black Papers’ (Cox & Dyson, 1969-74), one of the earliest manifestations of disquiet with British schooling; to 2020, when the pandemic marked a further set of major societal changes. The sense of history is an important theme in this research, as it seeks to analyse and explain this distinctive period, its causes, evolution and impact.

I contend that the research has significance because the topic of improving schooling has been the subject of lively, continuous debate for more than 50 years. Having been of sporadic political interest in England for the first three-quarters of the twentieth century, education, and specifically the quality of schools and schooling, was of continuous political interest from the mid-70s to the early 2010s. The debate has not been only at government level: in the press and around dinner tables as well as in academic circles, the state of our schools and the achievements of pupils have been of widespread interest and concern. The stakes are high for the pupil, several studies having established that those from ‘good’ schools out-earn those from less good (Betts, 1995; Card & Kureges, 1996; Sahadewo, 2023); and for the family, a study of the link between school quality and local house prices having found:

“House prices ... near the 10% best-performing primary schools are 8.0% higher than in the surrounding area. Near the 10% best-performing non-selective secondary schools, house prices are 6.8% higher.” (DfE, 2017, p.3)

[A headteacher interviewee whose school was judged by Ofsted to have slipped in quality, reports a parent saying to him “there goes my house price!”]⁶

⁶ I have included a small number of relevant, illustrative anecdotes from my experience within the review, but have highlighted them by bracketing and inseting them.

Education during the period has ‘enjoyed’ greatly-increased public involvement: the then prime minister’s Ruskin Speech (Callaghan, 1976), the first major speech on education by a serving British prime minister, inspired the opening of schools, classrooms and the curriculum to scrutiny; many of the policy developments of the 1980s and 90s sought to involve the public directly in the choice of school for their children, in the governance of schools and, latterly in sponsoring and creating free schools and academies (Montacute & Culliane, 2018; DfE, 2022).

The theme of improving schools remains not only one of interest and debate, but also of contention, within both the education community and society at large. As the review will show, there are authors (Coe, 2013; Allen, Evans & White, 2021) who cast doubt on the outcomes of improvement and the effectiveness of its processes [p.39].

The theme remains high on the socio-political agenda: the Education White Paper (DfE, 2022b) is shot through with references to ‘raising standards’, ‘world-class schools’ and ‘world class teaching’ (*ibid*, pp. 8-10).

2.2 THE PROCESS OF COMPILING THE LITERATURE REVIEW

The criteria used for the selection of sources are relevance, breadth and quality. To be ‘relevant’, the source must bear upon the research questions. ‘Breadth’ means diversity of viewpoint: attempting to hear all voices and present a balanced account. School effectiveness is a well-established concept but there is great variety of opinion concerning improvement, including whether it is ‘real’: the voice of the small but vocal group of ‘improvement deniers’ mentioned above must be included here.

Sources vary in ‘quality’: although stimulating ideas are found in sources like newspapers, greater weight is given to authoritative sources. These are identified by having key researchers as authors, by frequent citations and by appearing in peer-reviewed journals. However, the impact on schools and education practitioners of government grey literature (White Papers, manifestoes, legislation; Ofsted’s inspection frameworks, school and annual reports) provide important evidence of the steer government gives to the concept of improvement. Currency applies only partially as a criterion: although recent sources have

greater perspective by benefiting from earlier work, the differences in analysis at different points in time are illuminating.

Authoritative analyses of publications on the theme were located in two ways: a key word search using school effectiveness, school improvement, education reform; and the identification of a bibliographic analysis of the field. The work of Hallinger & Kovacevic (2019) proved valuable; their research scanned 22,000 peer-reviewed articles in the field across the period of my research and used science mapping to document the evolution of research, identifying the main journals and authors involved. Key journals specifically devoted to effectiveness and improvement were major sources; the starting point for much of the review is the journal *School Effectiveness & School Improvement* and the more recent *Improving Schools*. The focus was on articles dealing with the UK and on those with an international perspective. Key researchers, major figures in the field whose work has an international reputation, were identified and followed, such as Ball, Chapman, Fullan, Harris, Hopkins, Reynolds, Stringfield, Teddlie. Hallinger's & Kovacevic's work was also valuable here.

The circle was then widened to engage with research journals of similar standing but not improvement-specific, such as the BERA journal. Meta-reviews of school effectiveness and improvement, both in the journals above, such as those in the special edition of *School Effectiveness and School Improvement* (2004) devoted to state-of-the-art reviews of effectiveness and improvement; the more recent version by Mitchell et al., (2015) proved especially valuable.

Intersecting quality journals, key authors and meta-reviews enabled a high degree of relevance in start-points. The bibliographies of these works formed the next field for research.

The social, economic and political contexts of educational change during the period are only glancingly referenced in the sources listed to date. This field is less familiar to me, so I used *History Today* as a starting point to identify key articles; and the work of distinguished social and political historians such as Victor Bogdanor to identify authoritative texts charting the social history of late twentieth century England

I acknowledge the extraordinary website *Education in England* (Gillard, 2018) which provides the full text of every act of parliament relating to education, and of major reports such as Warnock and Bullock, and of significant speeches, such as Ruskin.

2.3 DEFINITIONS

The meanings of key terms in the field are ostensibly straightforward: effectiveness poses the question “what makes a good school?”, while improvement asks “how can a school be made better” (an improving school), or “how can more schools be made good?” (an improving system).

“Educational effectiveness can be defined as the degree in which an educational system achieve[s] specific desired goals and effects. Since in the context of educational systems goals and effects are represented in terms of achievement, an educational system that contributes more to better student achievement is considered more effective than others”. (Burušić, Babarović & Velić, 2016, p.10).

The terms ‘outcomes’, ‘aims’, ‘goals’ and ‘achievement’ are all open-ended: defining them is hampered by the lack of agreed purpose(s) for education. Trends in government thinking about purposes, and their implications for schools and leaders, are discussed below [p.119]. Though discussion is age-old, the impact of uncertain aims on school leaders is rarely considered [p.122].

Effectiveness research seeks to identify the attributes of those schools which achieve better outcomes for their pupils than others; this pursuit of ‘correlates with effectiveness’ is examined in detail below [p.31]. Distinguished authors in the field, among them Reynolds (1993) and Sammons (2007), use the term ‘paradigm’ to describe school effectiveness, but do not justify its use: this minor gap in the research is reviewed under Discussion [p.116 *et seq.*].

Effectiveness can be seen as a static concept, improvement as dynamic: improvement research evolved from effectiveness and investigates how schools achieve better outcomes for pupils over time. The correlates generated by effectiveness research are applied by school leaders to bring about improvement. In a major conference speech, Reynolds claimed that

“our knowledge of school effectiveness is extensive... (but) improvement is the weak link in the chain” (Reynolds, 2023), highlighting the problem explored throughout this thesis: how effectiveness knowledge can be used to bring about improvement.

Some definitions of improvement are broad, such as “efforts to make schools better places for pupils and students to learn in” (World Bank, 2005), which might reasonably include installing air conditioning; the theme of underlying improvements of this kind are largely absent from research literature, a point referred to in Chapter 1 and explored in Discussion [p. 149 *et seq.*].

Other definitions are tangential, such as “School improvement is how schools create an ever-better climate for the individual and groups of teachers to do their job” (Brighthouse. foreword to NAHT, 2020). Although flawed as a definition because it does not describe benefit to pupils, the approach is mirrored in the proposed school improvement model [p.136 *et seq.*]. Most definitions of improvement resemble that provided for its schools by the Welsh Government, highlighting experiences as well as outcomes for pupils:

“The overarching purpose of school improvement is to help schools give learners the best possible learning experiences and outcomes, whatever their background or circumstance, in order to achieve high standards and aspirations for all” (Welsh Government, 2022).

2.4 TESTING THE BASIC PROPOSITION

Central to this work is the assertion that effectiveness and improvement came to dominate the discourse of schooling from around 1980 until the 2010s. The evidence used to test this assertion comes from three sources: the volume and content of government education policies; the themes of education research; and later, in Chapter 4, from the views of interviewees. Anticipating those findings, in summary, all agreed with the proposition that their thinking during the period was dominated by quality and how to improve it; all referred to key events reinforcing the dominance: regular inspection by Ofsted, the annual sets of test and examination results, their publication in league tables, and the annual cycle of self-evaluation and re-drawing of improvement plans.

An authoritative account of the volume of research related to school effectiveness and improvement comes from Hallinger & Kovacevic (2019). Their systematic review of over 22,000 peer-reviewed articles from between 1960 and 2018 used science mapping to document the growth of different research themes, identify high-impact scholars and documents, and illuminate the “intellectual structure” of the field of educational administration research (within which they locate effectiveness and improvement).

Their study identifies a sharp increase across the period in the number of articles relating to the field, with four themes showing particular growth, one of which was effectiveness and improvement (*ibid*, Tables 2 & 3). In common with authors of meta-reviews noted above, they identify the rise of effectiveness and improvement research as a paradigm, and 1979 as the point of paradigm shift. This contention is examined below [pp.116-117].

They identify the 20 leading scholars (based on volume and citations) within the field of education administration and leadership in the period; six of them (Hopkins, Reynolds, Sammons, Fullan, Harris and Bryk) are listed as effectiveness and improvement scholars; their work is referenced extensively throughout this study.

Finally within the evidence from research, it is noteworthy that three of their 22 source journals (*ibid*, Table 1) are devoted specifically to effectiveness and improvement research: *School Effectiveness and School Improvement* (began 1990), *Improving Schools* (1998), *Journal of Educational Change* (2000). Other journals devoted specifically to the themes now exist, including the *Journal of Effective Schools* around 2000 and *Effective Schools* in 2019.

Although its universal significance and the amount of public money invested in it⁷ might suggest that education would be a constant political priority, the reality has been different. For much of the twentieth century, there was near silence from government in terms of major educational legislation. The first ever education White Paper was issued in 1943: entitled 'Educational Reconstruction', it presaged the 1944 Act's creation of a three-tier secondary education system. Between then and 1992 there were only three schools-related White Papers, the first themed around quality in 1985: 'Better Schools'. Since 1997, there has been an education White Paper on average every two and a half years, with quality-related titles such as 'Higher Standards, Better Schools' (2022) and 'The Importance of Teaching' (2010).

There were innumerable pieces of education legislation in the twentieth century, some of considerable social significance such as those controlling employment of children, but very few affecting the national architecture of education. In the period 1900 –1919, there were two major Acts; 1920 - 39, one; 1940 - 1959, one; 1960 -79, also one. Then, in the period of this research, acceleration: between 1980 and 1999, there were six major Acts, and a further six since 1999, although with a distinct slowing in the second half of that period (data from Gillard, 2018).

Even more significant to the key assertion is the nature of that legislation: until late in the twentieth century, the role of the national government in schools was limited to control of overall expenditure and to "ensuring that there were enough school places for all children and enough teachers to teach them" (Pring & Roberts, 2016, p.4). Legislation mainly concerned the structure of schooling, such as the school-leaving age and the powers of different types of school.

The Prime Minister's Ruskin College speech (Callaghan 1976) marked a genuine watershed. Only 20 years later, prime-minister-in-waiting Tony Blair, in his launch of the Labour party manifesto for the 1997 election, acknowledged that "nothing matters more [than education] ... to the future prosperity of this country" (Blair, 1996b); hence his policy mantra "education, education, education" (Blair, 1996a).

The origins of the Ruskin Speech and its place in the history of school effectiveness and improvement are considered in a subsequent section of this review [p.39]; it marked the first

⁷ 5.5% of UK government expenditure in 2020 (World Bank, 2021).

incursion of government into schools and classrooms, and was followed over the next two decades by a flurry of legislation aimed at circumscribing the leadership of schools, their performance and accountability for it, and their curricula. Key dates include open publication of the inspection reports of Her/His Majesty's Inspectors (HMI) (1983); the establishment of a National Curriculum (1988), a national system of tests and examinations at 7, 11, 14 and 16 (1988); the publication of those tests and examination results in public performance tables (1993); the establishment of a national school inspection body, Ofsted, with its system of regular inspection of all schools with the reports made public (1992). Roberts expresses this revolution in the role of the headteacher:

“whilst school leaders were given much greater control over financial and organisational matters, the educational objectives which they had to achieve were now largely dictated through the curriculum, public accountability measures and inspection regime.” (Pring Roberts, 2016, p 59).

[A headteacher in my LA put it more pithily: “yesterday I could teach what I like but couldn't spend more than a few pounds; today, I am told what to teach but have control of almost the whole of a multi-million-pound budget”.]

One final point securing the assertion that improvement dominated schools during the period was the requirement under the Education & Inspections Act 2006 that all schools should be allocated a trained and qualified adviser; the significance being that s/he should be called a '*school improvement partner*' (SIP).

2.5 REVIEW OF THE LITERATURE CONCERNING THE EVOLUTION OF SCHOOL EFFECTIVENESS AND IMPROVEMENT

The remainder of this review examines the state of knowledge concerning the development of the concepts of effectiveness and improvement, in particular their rapid evolution in the 1980s and succeeding decades. As 2.4 shows, there was a surge in debate, research and policy during this period, in the UK and more widely in the western world: many authors (Reynolds *et al.*, 2014; Hopkins *et al.*, 2014; Pring & Roberts, 2016; Ball, 2021) describe it as a period of wide-ranging reform: the urgent need to improve a system must arise from some

combination of disquiet with the current situation and changes in the social context or the market served. My hypothesis is that the period of rapid change was brought about by both sets of factors: dissatisfaction with aspects of education (push factors) interacting with changes in the wider social, economic and political environment (pull factors): causes are analysed in 2.5.1, while 2.5.2 analyses the response, the changes and improvements in schools.

2.5.1 Causes of the surge in school effectiveness and improvement activity.

It is ironic that the organised demand for school improvement began at a time when the prevailing research *zeitgeist* was that schools have no effect: attempting to improve the effectiveness of schools which allegedly have no effect is paradoxical and demonstrates that the 'no effect' paradigm had by that stage become stretched.

Albeit a generalisation, the twentieth century was, for most of the UK population, a period of improved living standards: across the century, national per capita income, index-linked, increased more than five-fold; life expectancy rose from the mid-40s to the late 70s as a result of improved nutrition, housing and health care; maternal death rates reduced by a factor of 79 (Horrell 2000, p.62). In the early years of the century, nearly two-thirds of Britain's population died before the age of 60; by its end, only 12% did so. Owner-occupation of houses rose from 10% of the population to nearly 70% (House of Commons, 1999).

Mechanisation in both the workplace and the home reduced grinding physical work. Overall, disposable income increased, accompanied by some increase in leisure time (Horrell, 2000, p.63) to enjoy it. As late as 1964, three million British households had no bath, inside toilet or running hot water; ten years later, half of homes had central heating, half had a telephone, 70% a washing machine (Benson, 1994, p.72). Although prime minister Harold Macmillan's statement in 1957 that the country had "never had it so good" became a music-hall joke, his expansion of it into "prosperity such as we have never had ... in the history of this country" (Macmillan, 1957) was pretty accurate. The expectation that 'things improve' had been established.

In education in England, the first half of the twentieth century was undoubtedly a time of *more*: compulsory primary education was extended to secondary, initially to age 14 under the Education Act 1918, to 15 (1947), 16 (1972), 17 (2013) and 18 (2015) (Gillard, 2018). When I

went to university in 1963, I was one of about 4% of the nation's school leavers to do so, the proportion rising to 19% by 1990 (*ibid*).

In the 1960s, questions began to arise about whether *more* necessarily meant *better*: the late 1960s and early 1970s were a time of growing public and political disquiet about education. This will be a repeated theme: internal pressure for change within education is supported or driven by external pressures from the economic, social and political environment. The literature tends to separate these factors; very few sources attempt a comprehensive account of factors, showing how they interact: Section 2.6 identifies this as a gap which the current research will address.

The disquiet is analysed by Jones (2003): "the problems of schooling [were seen] as a kind of condensation of all the worst effects of post-war history" (p.74). Both Jones and Ball (2021) classify those effects into two: problems in the economy, and the breakdown of order and moral values. Each of these is considered below, the second together with other societal changes. The same disquiet grew in other parts of the world, notably the US: that nation's Commission on Excellence in Education Report *A Nation at Risk* cites the same two sets of issues, asserting that US education had lost sight of the two key attributes of a good education "high expectations and disciplined effort needed to attain them" (NCEE, 1983, p.1).

The causes of rapid educational reform can be classified into three broad categories: economy, society and education itself; each is explored in the following sub-sections.

2.5.1.1 Education and Problems in the Economy

From at least the late nineteenth century, British education had come under almost continuous criticism for failing to prepare young people for current or future jobs. In his introduction to the Wolf Report (2011) into vocational education, Secretary of State for Education Michael Gove, wrote:

"Since Prince Albert established the Royal Commission in 1851 policy-makers have struggled with our failure to provide young people with a proper technical and practical education of a kind that other nations can boast. 160 years later the same problems remain. Our international competitors boast more robust manufacturing industries. Our

technical education remains weaker than most other developed nations.” (DfE, b Report, 2011, p.4)

In the mid-nineteenth century, engineer George Stephenson wrote that Britain’s neglect of technical education would lead to loss of business to continental competitors, especially Germany (quoted in Fores, 1971, p.39). A hundred years later, a government White Paper (Ministry of Education, 1956) describes the same situation as if it were a discovery:

“it is clear all these countries [US, Russia, Western Europe] are making an immense effort to train more scientific and technical manpower and that we are in danger of being left behind.” (para. 3).

Yet again, in 1981, the White Paper ‘New Training Initiative: An Agenda for Action’ re-discovers the same issue (DES,1981,para.6): “Our major competitors lay much greater emphasis on training young people than we do.”

From the time of the 1957 Russian launch of Sputnik 1, doubt was cast on the assumed superiority of western science and technology. Although utilitarian purposes for education had been voiced since at least the time of John Stuart Mill, they came to the fore as a response to this failure of competitiveness (Hodge, 2007; Hopkins, 2022), giving rise to an immediate surge in the demand for science and technology education in schools (Hopkins *et al.*, 2014). Harold Wilson’s speech to the 1963 Labour party conference spoke not only of “the white heat of technology”, but also of “the strength, solvency, the influence of Britain ... are going to depend, in the remainder of this century, ... on the speed with which we come to terms with the speed of change” (Wilson, 1963). Not only did he use the speech to point out the need for more scientists and technologists, but he also linked this to the waste of talent caused by “segregation” of children into grammar and modern schools at age 11, and said it should no longer be tolerated.

Sources are clear that the two specific criticisms of education in relation to the economy (failing to provide vocational programmes of a status equal to that of more traditional ‘academic’ courses, and failure to keep up with the scientific and technological demands of the job market) have been repeated consistently for at least 150 years, and that those criticisms are widespread, from both government and employers. Business leaders contributing to the 2014 Confederation of British Industry ‘Education and Skills Survey’

“wanted young people to be better prepared for life beyond school”, demanding improvements “in careers education and in school leavers’ technological and scientific skills, literacy and numeracy” (CBI/Pearson, 2014).

The extensive literature concerning the changing nature of employment in the second half of the twentieth century (Wyatt & Hecker, 2006, especially Chart 1, p.36; ONS, 2019 Section 5 and Figure 2b) shows the reduction in jobs in primary and secondary industries and the startling rise in “professional scientific and technical services including health and education” (*ibid*, Figure 2b): skilled jobs requiring qualifications. As early as 1956, Prime Minister Eden recognised this trend and its implications: speaking at Bradford, he referred to the world-wide scientific revolution (and again to the lack of technical and vocational education in Britain), adding: " The prizes will not go to the countries with the largest population. *Those with the best systems of education will win.*" (Eden 1956; my added emphasis). It was not until 1966 that the term ‘knowledge economy’ first came into common use, in the work of Peter Drucker (*The Effective Executive*, 1966), but the phrase captures Eden’s meaning.

The literature of social history is clear that the 1970s was a decade of considerable economic turmoil and social change in the UK (Jenkins, 1987; Sandbrook, 2012; Ball, 2021). These sources refer to events throughout the decade: the OPEC crisis of 1973 which tripled the price of oil; the bail-out of the British economy by the International Monetary Fund in 1976; endless trade union disputes leading to power cuts, the three-day week (1973-4) and the so-called ‘winter of discontent’ (1979); inflation surging to nearly 30% coupled with stagnating wages in the mid-decade. The consensus which characterised British politics in the post-war period began to dissolve (Toye, 2013); significant government intervention in the economy, a strong welfare state, the nationalisation of major industries and services, relatively high levels of taxation, strong trade unions and low unemployment began to be challenged (Dutton, 1997). Although the welfare state was initially conceived as provisions for the relief of poverty, in the 1940s the concept was enlarged to include a National Health Service and, via the 1944 Education Act, a national education service. Most authorities (Bogdanor, 2018; Ball, 2021; Brighouse & Waters, 2021) see education until the 1970s/80s as a component of the welfare state: a public good.

The economic turmoil and industrial strife of the 1970s undoubtedly contributed significantly to the weakening of the post-war consensus and its replacement by the new politics of

neoliberalism; Keynesian economics replaced by the market economics of Hayek and Friedman (Friedman, 1955; Ball, 2021, p.2; p.17 *et seq.*): smaller government, lower taxation, more choice and individualism. According to Jones (2016, p.2) neoliberalism comprises

“a free-market ideology based on individual liberty and limited government that connected human freedom to the actions of the rational self-interested actor in the competitive market-place”.

Keith Joseph is seen as the architect of what became known as ‘Thatcherism’ (Ball, 2021; Bogdanor, 2018; Brighouse & Waters, 2021). He believed in a smaller role for the state and a greater role for the market:

“because excessive state control and ownership limits the liberties of all citizens as well as leading to impoverishment ... the blind, unplanned, uncoordinated wisdom of the market is overwhelmingly superior to the well-researched, rational, systematic, well-meaning, cooperative, science-based, forward-looking statistically-respectable plans of government.” (Joseph, 1974a)

Joseph was also one of the major critics of the breakdown in traditional order and values; the same speech contains the following trumpet call:

“Should we not deal with matters which concern the nation; respect for other people and for law, the welfare of young people, the state of family life, the moral welfare of all the people, cultural values, public-spiritedness or its lack, national defence, the tone of national life?” (*ibid.*).

Thus the man who was to become the longest-serving Secretary of State for Education of the modern era was one of the leading critics of education, for its roles both as part of the welfare state and in the breakdown of discipline and order.

The social context for education was therefore no longer the welfare state but the market. It was into this context of instability, and the social turmoil that accompanied it, that the prime minister’s pivotal education (Ruskin) speech, was delivered (Callaghan, 1976).

2.5.1.2 Education in relation to changes in society.

Educational change in the late twentieth century was also driven by societal changes, summarised here under three headings:

- aspiration and choice: affluence and the increased proportion of the population who owned their houses (15% in 1914, 67% in 1996) and who identified as middle class
- breakdown in traditional order and moral values. Like the economic issues, the breakdown in order in society was laid in part at education's door: schools were accused of lax discipline and a failure to inculcate strong moral values in the young
- society's attempts to understand, accept and promote greater diversity and inclusion of groups and individuals from different social, ability, ethnic, gender and sexual orientation backgrounds, and to close gaps in the attainment of different disadvantaged groups.

Aspiration and choice: the 1950s and 60s saw a sharp increase in consumerism and personal choice. The rise in ownership of houses and consumer goods has already been described. In 1950, only about 15% of households owned a car; by 1970, the proportion had reached 50% and by 2020, about 77% (National Transport Survey 2021, Chart 7; earlier data from BBC News website, 2017).

It is understandable that pressure began to be exerted for citizens to be able to exercise the same choices in aspects of public service such as health and education. The proportion of pupils attending independent schools, although small, has edged up since the mid-1970s, from just below 6% (Bolton, 2012) to the current figure of just under 8%, and 16% of sixth-form students (ISC Census, 2022, Fig.5 p.8). Although some parents chose to send their children to schools other than the local, the concept of catchment areas persisted until the Education Reform Act (ERA) (1988), the terms of which made open enrolment a legal requirement and firmly established an education marketplace [p.40 *et seq.*].

Breakdown in order and moral values: Under this heading can be gathered aspects of the so-called 'permissive society'. Best (2009, p1) describes its positive side as "liberation from the oppression of laws that stood between individuals and their self-determination and fulfilment": the decriminalisation of abortion and homosexuality, the lifting of censorship, the reform of harsh divorce laws. The permissive society also encompassed 'the pill', abortion on demand and increases in drug-taking and pornography.

The characteristics of the 'swinging sixties' do not require detailed analysis here; suffice it to use the words of Florence Sutcliffe-Braithwaite in her book *The Decline of Deference, 1968 – 2000*, that this era was one which saw the "rejection of acceptance of and respect for social hierarchies, authority and tradition" (Sutcliffe-Braithwaite, 2018: Abstract).

The lifting of National Service in 1960, after which young men no longer experienced military discipline, was accompanied by the rise of sub-cultures such as teddy boys, mods and rockers in the 1950s and 60s. The student riots and protests of 1968 across Europe and the US added to a pervasive sense of unrest.

Keith Joseph's powerful attacks on the damage inflicted on the British economy by the post-war consensus have been noted above; his Edgbaston speech (1974a) described society in England as characterised by:

"delinquency, truancy, vandalism, hooliganism, illiteracy, decline in educational standards. Some secondary schools in our cities are dominated by gangs operating extortion rackets against small children. Teenage pregnancies are rising; so are drunkenness, sexual offences, and crimes of sadism." (Joseph, 1974a).

In a later speech in the same sequence (Joseph, 1974b), he asserted that, in addition to impoverishing the British economy, the post-war consensus had "destroyed the moral foundations of society", and that the optimism of the previous decades had "crumbled under the weight of rising crime, social decay ... the decline of traditional values ... have we really become a nation of hooligans and vandals, bullies" The collapse of community he attributed to "vast factories, *huge schools*, sprawling estates" (Joseph, 1974b: my added emphasis).

Thus the Secretary of State for Education, was deeply critical of the social role of schools and keen to bring about change.

Diversity and inclusion. The final group of external factors pressing for educational change came from awareness of the unequal rights and differential performance of sub-groups within the school population.

The British Social Attitudes report on gender roles (Scott & Clery, 2013) shows that whereas around half of those surveyed in the early 1980s believed in 'traditional' gender roles (men go out to earn, women look after home and children), by 2012 only 13% did so (p.115).

Whereas the number of men in full-time employment stayed roughly static during the period 1984 - 2012, the number of employed women increased significantly (*ibid*, p.118, Fig. 5.1). Traditional assumptions about the interests, skills and career paths followed by girls and young women were forced to change.

The ethnic diversity of the school population increased sharply from the mid-twentieth century: “The ethnic minority component of the age 5-16 maintained school population in England has increased from 16.8% in 2003 to about a third in 2020” (Migration Watch, 2021, para. 24 *et seq.*).

Throughout the twentieth century, the proportion of children with a learning disability who had access to education increased. The 1918 Education Act provided for handicapped (*sic*) children to be educated but only in special schools or classes; the 1944 Act created a variety of special school provision for different types of disability; the 1970 Education (Handicapped Children) Act transferred responsibility for the education of severely disabled children from health to education authorities (Gillard, 2018: Parliamentary Business, 2006). The Warnock Report (1978) changed the language of learning disability, introducing the phrase ‘special educational needs’, emphasising the concept of inclusion and the presumption of access to mainstream education for all but a very few children (Parliamentary Report, 2006). The 1981 Education Act, receiving all-party support, enacted many of the recommendations of Warnock, changing the language and stressing inclusion.

The emphasis on adapting schools to the demands of increasing diversity, whether by gender, ethnicity or ability, initially concerned attitudes, training of staff and resources. Later, especially when performance league tables began to be published in the early 90s, the emphasis shifted to differences in performance in tests by different groups. The phrase ‘achievement gap’, first used in 1963 in the US to describe differences in achievement between ethnic groups (Walker, 1963), began to be applied more widely, to gender and ability, as well as ethnic, groups. In the 1990s, it was realised that difference in attainment between children of different levels of social advantage dwarfed those of, for example, gender groups; the Sutton Trust, which “aims to improve social mobility and address educational disadvantage” was set up in 1997 (Sutton Trust website). The most convenient statistical proxy for the average affluence of a school was found to be the proportion of its pupils eligible for free school meals. From 2011, this proportion was recalculated based on

eligibility for the new pupil premium grant, a more sophisticated proxy measure for disadvantage (EEF, 2017).

It is noteworthy that, when the world's national education systems are rated and compared, the index is no longer simply pupils' educational performance; this forms the x axis of a graph whose y axis is a measure of the gap in attainment between advantaged and disadvantaged pupils. Good examples are provided by Knowles & Evans (2012) in the graphs and charts in the appendices, which analyse the PISA data for 2009 by attainment and social advantage gaps; and especially by Hopkins's (2022, p.6 Figure 1) use of OECD PISA data to graph the quality of nations' school performance against equity.

In summary therefore, the literature shows that, since the 1970s, education has been under pressure to change from at least five external sets of forces:

- a shift in political ideology which changed the context of education from a part of the welfare state to a marketplace
- a persistent critique that it is not preparing students adequately for the jobs of the future
- increased scrutiny because education lies at the heart of the knowledge economy
- part responsibility for the ills of society, whether those be childhood obesity, knife crime, teenage pregnancy, poor behaviour in public ... Even if responsibility does not equal blame, major responsibility for tackling those ills is conferred on schools
- a more diverse school population and the demand not just for equality of opportunity but also for equality of outcome for different groups.

These pressures argue not only for reform but for improvement: a more relevant curriculum, higher standards not just of attainment but also of personal development and behaviour, and for closing attainment gaps.

2.5.1.3 Changes from within education.

This section deals with internal pressures for change, although it will become clear that these overlap with the external set. An illustration of this overlap concerns the closure of grammar schools: although several comprehensive schools had already opened, the main impetus for their spread came from Circular 10/65, a message to LAs from Secretary of State for Education

Anthony Crosland to “request[s] local education authorities, if they have not already done so, to prepare and submit to him plans for reorganising secondary education in their areas on comprehensive lines” (DES, 1965, Introduction). The rationale for this request, in the Introduction to the Circular, quotes a motion passed in the House of Commons:

'That this House, conscious of the need to raise educational standards at all levels, and regretting that the realisation of this objective is impeded by the separation of children into different types of secondary schools, notes with approval the efforts of local authorities to reorganise secondary education on comprehensive lines which will preserve all that is valuable in grammar school education for those children who now receive it and make it available to more children” (DES, 1965, Introduction).

Over the following 15 years there was some yo-yoing in policy toward comprehensivisation between Labour and Conservative administrations, but comprehensive schools rapidly became the great majority: by 1976 they were educating a third of secondary-age pupils (Sandbrook, 2012, p.196) and by 2020, of the roughly 3400 secondary schools in England, only 163 were grammar schools, educating roughly 5% of the pupil population (House of Commons, 2020).

Pressure to get rid of selection at age 11 arose from several sources. There was popular disquiet about the finality of separating ‘sheep’ from ‘goats’ at age 11 (Jenkins, 1987, p.267-8; Sandbrook, 2012, pp.196-7) based on a test whose validity was discredited in the early 70s (Joynson, 1989). Popular literature such as Barry Hines’s *A Kestrel for a Knave* (Hines, 1968) painted a picture of uncaring life in a secondary modern school. The policy to create comprehensive schools met with widespread initial support (Sandbrook, *ibid*). Research had revealed that “success at 11+ had more to do with class than merit” (*ibid* p.197).

However, early enthusiasm for comprehensives waned and public opinion became ambivalent. The creation of the comprehensive school may have been popular, the closure of grammar schools was not. Not only did grammar school closure dissatisfy many in the middle class who saw in it the provision best matched to their perceived social status, but also many in the working class, because closure cut off access to the upward social mobility the grammar schools allegedly provided, and which matched the aspirational, property- and share-owning society promoted by the Conservative governments of the 80s.

[In 1987, I was the junior member of a two-person team from the LA at a consultation meeting with parents and the community concerning the planned closure of a long-established grammar school through merger with the nearby secondary modern school, on the latter's site. A hall packed with local residents implacably hostile to the idea of closing the grammar school, subjected us to two hours of verbal battering. Driven to exasperation by the unreasoning, as we saw it, 'debate', my senior asked for a show of hands of those in favour of the retention of the secondary modern school. Not a single hand went up. "Then we have a problem of logic" he said.]

Secondary modern schools were generally unpopular because they symbolised failure: few actively wanted to go to them, most 'ending up there'; many were of poor quality and poorly resourced, grammar schools receiving three times the amount per pupil (Sandbrook, *ibid*, p.197), but they represented the inevitable corollary of selection (Partridge, 1965).

In the public imagination, the rise of the comprehensive school became conflated with the growth of so-called 'progressive', child-centred methods such as mixed-ability teaching, and what was seen as the 'dumbing down' of the curriculum, less-demanding courses replacing study of, for example, classics. The report of America's National Commission on Education described the US high school curriculum as a "cafeteria" in which "the appetizers and desserts can easily be mistaken for the main courses", with most students opting for less-demanding courses in physical and personal development (NCEE, 1983). As referenced above, these changes in education were blamed, at least in part, for the breakdown in order and traditional values in society.

Dissatisfaction with these trends from within the profession led to the writing of the Black Papers (1969-77), a series of articles critical of the changes; the papers' titles, such as 'Crisis in Education' and 'The Fight for Education', contributed to the national sense of disquiet. The authors of the first two papers asked, "How far are we witnessing the progressive collapse of education?". Notable authors such as Kingsley Amis and Iris Murdoch contributed to the series, which spread into a broad critique of the closure of grammar schools, the loss of setting by ability and the growth of 'discovery' methods, especially in primary schools. The papers were published in the *Critical Quarterly*, the website of which described them as "... an attack on the excesses of progressive education and the introduction by the Labour Party of a system of 11-18 comprehensives to replace the grammar school" (Critical Quarterly website).

Hunter Davies carried out a detailed study of a large comprehensive school (Davies. 1976) and produced a report that told a balanced story; when the BBC Panorama programme followed up the report, to demonstrate disorder, young people were shown smoking in the school, but it was later revealed that the shots were taken in local flats and the children had been encouraged to smoke by the programme's producers. Parts of these accounts Davies called "sneering, smearing". In 1978, BBC began screening *Grange Hill*, stories from a fictional North London comprehensive; in its 30-year run, the series tackled issues including drugs, teenage pregnancy, knife crime, sexual assault, child abuse and alcoholism. Realistic it may have been, but the association with comprehensives (no-one made a similar series about grammar or independent schools) was unfortunate. In 1979, Pink Floyd's *The Wall*, in which the children of Islington Green School sang "we don't need no education", was issued.

The criticisms have persisted, Alastair Campbell describing "bog-standard comprehensives" (Campbell, 2001) and Andrew Adonis "secondary modern comprehensives" (Adonis, 2012, p.12). In his introduction to the same book, Adonis says "I saw failing comprehensive schools, many hundreds of them, as a cancer at the heart of English society" (*ibid* xi).

The popular view of permissiveness and disorder in schools was fuelled by a small number of high-profile school disasters. William Tyndale Junior School and Risinghill Secondary School, both in Islington, embraced extreme versions of child-centred education, such that children were given the choice of what, and indeed whether, to learn. Risinghill opened in 1960 and closed less than five years later, ostensibly because it was unpopular with parents; contradictory views of it are recounted in detail in Berg (1974). Whatever the rights and wrongs, the closure of a large, early comprehensive school shook public confidence. The William Tyndale affair achieved much greater notoriety and took place in the era of the Black Papers; so high-profile were the school's problems that a parliamentary inquiry was carried out (Auld Report, 1976). This concluded that the approach to discipline and the curriculum gave rise to anarchy in the school. Eventually, five teachers were sacked and the school re-organised. The impact was far greater however: the LAs involved seemed incapable of controlling Tyndale and schools like it, leading to a great weakening of confidence in them, and to the Inner London Education Authority being first reconstituted (1986) and then abolished (1990). Davis (2002) provides a full account. Other so-called 'free' schools in Liverpool and Islington faced similar problems and were closed; these incidents contributed

to the weakening of the powers of LAs in the 1980s and 90s (Roberts & Pring, 2012, chapters 1, 5 and 14).

The final pressure for change and improvement from within education itself came with the launch of international comparisons of student performance. The first such tests were TIMSS (Trends in International Maths and Science Study), launched in 1995. In that year, the superiority of countries such as Singapore and Korea was striking, and England's mid-table position came as a surprise to many (TIMSS Highlights, 1995). The best-known test, PISA, describes itself (OECD PISA website, 2023) as "the OECD's Programme for International Student Assessment. PISA measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges." It began in 2000, is carried out every three years and has steadily increased both its geographical range (now 90 countries) and the variety of the disciplines tested. Although in 2000 England scored well in science, it was mid-table for language and mathematics, bringing further sharp pressure for improvement.

2.5.2 Review of the literature concerning the evolution of the concepts of school effectiveness and improvement

Section 2.5.1 analysed the causes of the surge in effectiveness activity; 2.5.2 analyses what the literature says about responses to the demand for change.

Logic would suggest that research into school effectiveness preceded that into improvement; most sources therefore begin their accounts of improvement with the beginnings of effectiveness research in the mid-to-late-1970s (Fullan, 2008; Reynolds, 2012; Brighouse & Waters, 2021). In summary, effectiveness research established that schools do have an effect, sought to quantify it and provided the framework of factors or correlates (what effective schools have/do/are) to steer the work of improvement. This straightforward view of 'what happened next' in research is provided by both Barber (2009) and Hopkins (2014):

“the school effectiveness research in the 1980s ... gave us increasingly well-defined portraits of the effective school, that led in the 1990s to increasing knowledge of school improvement (i.e. how to achieve effectiveness).” (Hopkins, 2014, p.88)

The same author confirms the point ten years later:

“factors associated with the effective school soon led to my involvement in the nascent school improvement movement, where we used this research to begin to develop strategies and policies for making the school, and later school systems, more effective.” (Hopkins, 2022, p.2)

Accounts of the history of school effectiveness reference the work of Coleman, Bernstein and Jencks (Fullan, 2008; Reynolds *et al.*, 2014; Brighouse & Waters, 2021, chapter 7; Hopkins, 2022), who concluded that schools “bring little influence to bear on a child’s achievement that is independent of his background and general social context” (Coleman, 1966, p.325). This was no passing whim: sources agree that Coleman’s was a major piece of research, and when Jencks (1972) re-evaluated Coleman’s data and methods, he came to the same conclusion. The same quotation is frequently used to summarise the thinking of this period: “education cannot compensate for society” (Bernstein, 1970). Even the Plowden Report (Plowden, 1967), which achieved the status of holy writ among primary educators, emphasised the influence of home and community on children’s prospects, and the marginal effect of the school in comparison (*ibid* para. 76 *et seq.*). David Reynolds describes being introduced at a conference in 1976 as “[one] who persists in believing that schools make a difference’, after which there was prolonged, thunderous laughter” (Reynolds, 2023).

Thus the late 1960s/early 70s was a period when the prevailing paradigm was ‘no school effect’; thus attempts to increase their effectiveness would be a logical inconsistency. Nonetheless, both Hopkins and Reynolds separately (Hopkins *et al.*, 2014) and together (Hopkins & Reynolds, 2001) acknowledge earlier origins of improvement. They point to a major early study in the US (Aikin, 1942) and claim that the intellectual underpinning of the improvement movement lies in the organisational change management work of Lewin in the 1930s and 40s, and of Miles (1967) work on ‘organisation health’ (p.459). Miles investigated schools’ ability to cope with change and found them unprepared for change because of

professional isolation, likening them to a set of individual classrooms held together by a shared heating system.

[Classrooms in the first school in which I taught were autonomous, sometimes fiercely so: when a new headteacher arrived, he announced in his inaugural staff meeting his intention to get to know all staff by visiting our classrooms. The very long-serving head of mathematics said “headmaster, the moment you enter my classroom, I shall stop teaching.” What took place in a classroom was a matter solely for that teacher.]

Such an organisation is not prepared for change, and a wave of change was about to hit schools. Miles’s research recommended school- and team-based in-service training; such collective training did not arrive in England until ERA introduced the five statutory training days for all schools called, after the secretary of state responsible for the Act, ‘Baker Days’ (Education Reform Act (ERA), 1988).

Further early examples of school improvement effort are provided by the work of Stenhouse to promote curriculum and teacher development through the ‘teacher as action researcher’ movement (Stenhouse, 1980) and by the numerous curriculum development projects developed by the Schools Council (1964 - 84).

Despite the growing pressures for change, these developments had limited impact because they took place at a time when the prevailing *Zeitgeist* was that “academic achievement was related to race, class, gender or genetics” (Hopkins, 2022, Intro p.2). Effectiveness research had yet to create a theoretical framework upon which to build improvement efforts:

“many of the initiatives associated with this first phase of school improvement were ‘free floating’, rather than representing a systematic, programmatic and coherent approach to school change” (Hopkins & Reynolds, 2001, p.459).

The same authors point out that the emphasis of the early work was on organisational change, not overtly linked to practice in classrooms or to outcomes for students (*ibid*).

[The crudity of some of these projects was caught by Prof Ted Wragg who, in a meeting of the National Curriculum Committee, of which both he and I were members, described them to me as “‘give ’em a haircut and a homework diary’ models of school improvement”.]

A later account by Stringfield & Teddlie (2011) describes how Coleman's conclusions had, by the mid-70s, come to be seen as counterintuitive. In his history of school effectiveness research, Reynolds (Reynolds *et al.*, 2014) showed that small-scale studies challenging the Coleman conclusion began to appear, such as those by Weber (1971) and by Reynolds himself (1976), each of which pointed to a school effect on student outcomes.

However, all accounts of the evolution of improvement and effectiveness research agree that the work of Rutter (1979) was pivotal: this is captured by Hopkins (2022, p.2):

“This was one of the first major studies to demonstrate unequivocally that schools **do** make a difference and that the school a child goes to, indeed the classroom they join, makes a significant difference to their academic progress and subsequent life chances. Rutter and his colleagues compared the ‘effectiveness’ of ten secondary schools in South London, on a range of student outcome measures, and found that despite similarities in intake and socio-economic context, some schools performed better and were more effective than others”.

As well as demonstrating the ‘school effect’, Rutter and his colleagues measured the effect size and began the process of identifying those attributes characteristic of the effective schools, such as “the degree of academic emphasis, teacher actions in lessons, the availability of incentives and rewards, good conditions for pupils, and the extent to which children are able to take responsibility” (Rutter *et al.*, 1979, p.178).

They further identified a multiplier effect: the effect of the factors operating together greatly exceeded that of any of them operating alone. Thus they were able to postulate that “individual actions or measures may combine to create a particular ethos, or set of values, attitudes and behaviours which will become characteristic of the school as a whole” (Rutter, 1979, p.179).

Tim Brighouse, then chief education officer of Oxfordshire, recalls (Brighouse & Waters, 2021, p.288) the “emotional excitement” brought by Rutter's revelations to a group of his headteachers.

[I felt the same: I discussed Rutter's findings, and how they empowered schools, at an interview for deputy headship in 1980. The phrase ‘school improvement’ was not in

common currency, but we could dimly see that the attributes Rutter was describing made sense, were accessible, in the sense that we could take practical actions to make them happen; and were convincing: they would make a difference.]

A parallel piece of work was published in the US also in 1979: Edmonds's study of schools for the poor also challenged Coleman's conclusions and, like Rutter, identified common characteristics of schools that achieved more for their pupils than others:

"[Edmonds] identified five correlates of school effectiveness: purposeful and focused leadership from the principal, a focus on instruction, a safe and orderly climate, high expectations on the part of teachers, and frequent monitoring and assessment of student achievement." (Fullan, 2008, p.1)

What Fullan calls "Edmonds' ... famous challenge" (*ibid*) does not merely suggest that knowledge of effectiveness can lead to improvement, it rather demands that:

"We can, whenever and wherever we choose, successfully teach all children whose schooling is of interest to us. We already know more than we need to do that. Whether or not we do it must finally depend on how we feel about the fact that we haven't so far" (Edmonds, 1979, p. 1; quoted in Fullan *ibid*).

As I contend elsewhere in this text [p.117], these works marked a genuine paradigm shift from 'no effect' to 'significant effect', a shift that can be dated precisely to 1979.

The literature describes but does not explain the flowering of research, policy and legislation related to school improvement that followed the paradigm shift.

I now analyse the expansion of activity under three headings: the contributions made by research, government policy and Ofsted.

The contribution of research

All major sources describe the expansion of effectiveness studies following 1979 as seeking to ratify, refine and extend the list of effectiveness correlates. Fullan (2008) observes that two of the most respected such studies were led by members of the 1979 teams: in the UK, Mortimore, from the Rutter team, led the junior years study titled, significantly, *School*

Matters (Mortimore *et al.*, 1988), which confirmed the Rutter findings and refined the list of effectiveness factors, identifying eleven. In the US, Lezotte, from the Edmonds team, researched his *Correlates of Effective Schools* (1991), in which he modified Edmonds' five-factor theory to seven. There is considerable overlap between the various lists (Fullan, 2008); the Lezotte 7-factor theory provides the substance: instructional leadership, clear and focused mission, safe and orderly environment, climate of high expectations, frequent monitoring of student progress, positive home-school relations, and opportunity to learn/student time on task (Lezotte, *ibid*; listed in Fullan, 2008, p.2). Although criticised for their naivety (Reynolds, 2023), these seven factors still feature in lists of effectiveness correlates.

In the same piece, Fullan offers the insight that the Mortimore and Lezotte studies lead in different directions: Mortimore's study leads to further effectiveness studies with increasingly scientific and nuanced findings, whereas Lezotte is more action-focused, his work leading on to improvement. Fullan therefore suggests a divergence, the European tradition pursuing increasing sophistication in effectiveness, the American pursuing improvement.

Seductive as that argument is, the facts only partially bear it out. It is certainly true that, in the US in the 1980s and 90s, several school reform (Improvement) programmes were set up including Sizer's *Coalition of Essential Schools* (1989), Glickman's *Renewing America's Schools* (1993) and Slavin's *Success for all* (1996) (Reynolds, Potter & Chapman, 2003). Groups of schools working together, with the assistance of a support team, seem more effective in bringing about improvement than schools working alone. Nunnery (1998) posits this view, and my own experience bears it out: when working in support of a number of group improvement programmes in the 1990s and 2000s, and when carrying out the field and literature research for *A Review of Research & Practice* (Reynolds, Potter & Chapman 2003), I became convinced that the content of the programme, assuming it was not perverse, mattered less than the communality of the effort. Bringing energy and focus to bear on improvement; providing some resource; sharing experience within and between schools; adding expertise, facilitation and impetus from the support team appeared more impactful than the detail of the programme's content.

It is also true that, in the UK and the Netherlands in particular, further effectiveness studies were carried out: by the 1990s and early 2000s, adding nuance to the identified correlates:

- the use of more sophisticated statistical analysis techniques such as multi-level modelling (Aitken & Longford, 1986) to isolate school-, department- and classroom-level effects
- whether a school's effectiveness varied by subject department (Harris, Jamieson & Ross, 1995), or by the gender, ethnicity or social class of its students (Kyriakides, 2003)
- a parallel group of studies emphasised teacher, rather than school, effectiveness, and made a strong case for the significantly greater impact of the former (Muijs *et al.*, 2005). Analyses of effective teaching techniques in the Muijs study, the work of Rosenshine (2012) and Coe (2019) have generated clear guidance to teachers
- others studied the durability of effectiveness: especially in areas of low socio-economic status (SES): whether changes to the school demographic or leadership may lead the changes in quality (Bellei *et al.*, 2020)
- ascertaining whether effectiveness principles are international (International School Effectiveness Research Project [ISERP]: Hopkins, 1987)
- research into differential effectiveness among teachers and departments led to a branch of effectiveness research termed 'within-school variation', based on the insight that such variation dwarfs that between schools (Reynolds, 2007).

It has become fashionable to rail against what authors caricature as the 'one-size-fits-all' nature of school effectiveness principles:

"it has been found that rigid and "one-size-fits-all" policies can have a negative effect on schools when they leave little space for teachers and school communities to learn, innovate, and shape their own improvement strategies" (Giles & Hargreaves, 2006).

School Effectiveness and School Improvement devoted an entire issue (2019 Vol. 30 Issue 3) to the theme. In it, Scherer & Nilson state:

"The days of the "one-size-fits-all" perspective on school effectiveness and improvement are numbered, as more and more evidence is surfacing that shows how certain groups of students, teachers, schools, educational systems, or even countries may not benefit from certain educational approaches, programs, or conditions to the same extent".

I doubt that Rutter, Mortimore and Lezotte ever believed the principles they developed were universal and for all time; it is clearly unproductive to start from individual contexts, since individuality can only be studied in relation to norms and principles.

To qualify Fullan's view [p.36], major studies continued in the 'opposite' area: the longest-running effectiveness study, the Louisiana School Effectiveness Study, is American, and dates to 1980 – 92 (Stringfield & Teddlie, 1988); while the High Reliability Schools Project (1995 – 2000), a highly-structured attempt to apply effectiveness principles to bring about improvement in three LA areas, took place in the UK (Reynolds, Stringfield & Schaffer, 2006).

By the 1990s strong arguments were being made to unite the two disciplines:

“in their [effectiveness and improvement] orientation to outcomes, input, processes, and context in education, they also have much in common. In the theoretical part, different orientations have been analysed and combined in a model for effective school improvement. Based on this analysis, an evaluation framework was developed for the analysis of the case studies of school improvement projects ... The theoretical model and the results of the analyses of the case studies were combined in a framework of effective school improvement (Creemers, Peters & Reynolds, 1989, Introduction).

The two disciplines have thus been aligned since the early 90s, the one creating the theoretical basis, and therefore the evaluation framework, for the case studies and projects of the other. In interview, a leading researcher expressed the view that school effectiveness “*has been done*” (QQ). It may be that studies continue to elaborate context-specificity, but the general principles have been established: meta-analyses of the field by Creemers & Kyriakides (2008) and by Reynolds *et al.* (2014) tell their story with finality.

To a degree, the same is true of improvement: in their state-of-the-art review, Hopkins *et al.* (2014, p.257) assert that “we have learned much about how to improve individual schools, but successful efforts at systemic improvement have been less common”. System improvement lies outside the scope of the current study, but it is important to state that governments will clearly be interested in system-wide improvement studies (Fullan, 2008) because they help identify the policy levers. Such became the interest in system improvement that McKinsey, the management consultancy, worked with two major figures

in school improvement, Michael Barber and Mona Mourshed, to produce a report on high-performing national education systems, an international system effectiveness analysis (Barber & Mourshed, McKinsey, 2007).

Reference is made below [p.44] to UK government-led improvement programmes such as the literacy and numeracy programmes of the 2000s. Research established that these 'next generation' school improvement projects have common characteristics:

- "a clear and comprehensive model of reform
- strong leadership at the regional level
- substantive training related to the goals of the programme
- implementation support at the school level
- an increasingly differentiated approach to school improvement" (Hopkins, Stringfield, Harris & Mackay, 2014 p.265).

The authors list (*ibid*, pp.263-4) some examples of successful projects with these characteristics: in England, Improving the Quality of Education for All and High Reliability Schools; in Scotland, the Improving School Effectiveness Project; in Canada, the Manitoba School Improvement Program; and the Dutch National School Improvement Project.

This account of research from the 1980s to the 2000s may give the impression of linearity and consensus. It is true that the movement to support improvement in schools through a thorough understanding of how quality is built has moved with momentum, but there have been dissenting voices. These come in two groups: those who question the methodology of effectiveness research, and those who question the impact of improvement.

The first group is typified by studies such as Cuttance (1982), who mounted a critique of the relatively small sample of schools used by Rutter, and the invalidity of extrapolating from a small sample in one area to schools of all types nationwide. Such cut and thrust (proposition, critique, re-evaluation and reaffirmation or adaptation) is the stuff of science and advances knowledge. This study has already mentioned adaptations to the critique of 'one-size-fits-all' approaches. Most recently, the drive to close attainment gaps has led to enhanced investigations into how the various effectiveness correlates operate in different contexts

(Chapman *et al.*, 2011). Anyone who doubts the venom of argument in effectiveness research might study the exchange of papers between Gorard (2010, 2011) and the rebuttals by, *inter alia*, Reynolds *et al.* (2012), concerning the validity of statistical data analysis techniques. Gorard uses the phrase “Numbers are like people; torture them enough and they will tell you anything” in the Abstract to his 2010 article.

A leading sceptic concerning the efficacy of improvement activity is Robert Coe. The opening paragraph of his inaugural lecture as director of the Centre for Evaluation and Monitoring at the University of Durham states:

“standards have not risen; teaching has not improved; research that has tried to support improvement has generally not succeeded; even identifying which schools and teachers are good is more difficult than we thought” (Coe, 2013, p.1).

Coe argues that evidence of rising standards is equivocal or elusive and provides a valuable check on the problems of grade inflation, especially in GCSE examinations.

The contribution of government policy

Just as there is a pivotal date in school effectiveness [1979, p.4], so there is in government action to reform and improve schools: 1976: Prime Minister Callaghan’s 1976 Ruskin speech. Earlier text described the slow pace of major education policy statements and legislation in the first three-quarters of the twentieth century and its remarkable acceleration in the final quarter; and the shift of policy focus from the organisation of education to reforming the processes of schooling and improving the quality of its outcomes.

Although, as earlier text also explained, there was a significant gap between the Ruskin speech and major government action, the serious economic and electoral problems facing governments in the late 1970s deflecting political energy away from education, nonetheless Ruskin represented a hinge-point. The early years of the Conservative government, elected in 1979 with the most radical neoliberal Secretary of State for Education in Keith Joseph (from 1981), saw a flurry of education legislation, but little of it directed at school reform and improvement: the 1979 Act dealt with the repeal of the requirement for LAs to prepare plans for their secondary schools to become comprehensive; the 1980 Act was more radical,

including plans for assisted places at independent schools and the growth of parent power on governing bodies; and the 1981 Act was mainly concerned with the implementation of the Warnock report's proposals for pupils with special educational needs. The 1984 and 1986 Acts introduced the theme of central control over education (such as requiring LAs to earmark funds for nationally determined programmes), a theme that was to expand rapidly over the next few years. The second Education Act of 1986 required schools to write and publish their curriculum policies and for their governors to report annually to parents on the progress of their school: openness and parental influence, prompted by Ruskin ten years earlier.

The direction of travel is clear: the government response to the criticisms and pressures for change, and to the growing understanding of school effectiveness, can be seen in the twin themes of greater control at the periphery (schools and parents) and at the centre (government), with a corresponding reduction in control by LAs. On the one hand, increased local management of schools and greater parental involvement, in line with Joseph's small government, market forces neoliberalism; but on the other, increased government control of funding and the curriculum. The steps at this stage are small: ten years after Callaghan's tentative questioning whether the nation should mandate a curriculum ("It is not my intention to become enmeshed in such problems as whether there should be a basic curriculum with universal standards - although I am inclined to think there should be" (Callaghan, 1976)), very little change had taken place. Thereafter, change was rapid and, following ERA (1988), England had a fully-fledged National Curriculum covering ten subjects (plus religious education), applicable to all pupils between the ages of 5 and 16.

Although there was a flurry of significant school-focused legislation over the following ten years, two acts of parliament stand out as creating fundamental change in the educational model and the placing of improvement centre stage: those of 1988 and 1992. Their effects were to bring market forces to bear on education; to weaken LAs mortally; to put headteachers and governors firmly in charge of schools and their budgets; and to apply a tight set of accountability measures to schools (a required curriculum, national tests and league tables, ubiquitous inspection).

Stephen Ball (2021) describes the outcome as: “The state funds and manages the system, using performance measures and requirements to ‘steer at a distance’ ... with a primary concern for global economic competitiveness” (p.21).

The terms of ERA were skilfully crafted to create a marketplace within which parents could choose. The logic chain ran as follows, with the main clauses of the Act underlined. To be in a marketplace:

- schools had to be funded such that good schools would attract more pupils and expand their budgets: per pupil funding. Little was said about the corollary that less-good schools would dwindle
- to enable good schools to grow, they had to be able to attract pupils from anywhere, not just a notional catchment area: open enrolment
- almost all the budget, most of which had previously been held by LAs in the form of staffing and premises costs, was transferred to headteachers and governors: local management of schools. Schools that wished could apply to leave the LA altogether, receive the entire budget and become, effectively, independent schools within the maintained sector, so-called grant-maintained schools, the ancestors of academies
- being in control of their budgets meant that schools would make different choices and would vary in style and quality, enabling real choice for parents, the obverse of open enrolment for schools being parental choice of school
- parents must be able to compare schools on a fair basis. A National Curriculum would ensure like-for-like comparison and provide entitlement for pupils
- schools therefore began to market themselves, including through a mandatory prospectus, latterly a website, containing statutory information such as policies and examination results, to help parents choose
- to assist their choice, it was argued that parents also needed hard data on school performance, and thus there would be national tests for all pupils at ages 7, 11, 14 and 16 (GCSE), the results to be drawn up and published by school in its prospectus and in government-produced league tables (from 1993).

The final piece in this elegantly crafted legislation to subject schools to market forces and increase their accountability was added by the 1992 Education Act, which created Ofsted,

which would employ independent inspection teams to inspect all schools regularly to a standard framework, and publish each school's report, including a summary issued to all parents. Key features of the new inspections in relation to school improvement are considered below, [p.154 *et seq.*]. To fund the new inspections, the government withdrew money from local authorities. With another education act (1993) making it easier for schools to become grant-maintained, it is unsurprising that LAs were described as “unloved and unwanted” by Howard Davies of the Audit Commission (Davies, 1991). As Ball (2021, p.85) points out, the government then gave Ofsted the power to inspect local authorities.

Underlying these legislative changes are a mistrust of teachers, based on 1970s stereotypes of ‘anarchic’ schools, a belief that standards were not high enough, based on data from international comparison tests, and that those standards would be raised by a combination of accountability measures and competition (Ball, 2021). Just before it lost power, the Conservative government published a White Paper *Self-government for Schools* in which it summarised the principles that had guided its work in schools:

“Over the past decade the Government has introduced a wide range of reforms to raise standards ... based on two key principles - that parents should have more choice in deciding the education they want for their children, and that schools should have more freedom to run their own affairs.” (White Paper, 1996, p.iii para.1)

School improvement featured strongly in the agenda of the Labour governments from 1997 – 2010, following Blair's “Education, Education, Education” mantra (Blair, 1996). Ball (2021 p.91), in a briefing paper to the Labour party conference of 1998, listed no fewer than “47 education-related policies, initiatives and funding decisions”.

It has been remarked (Crisp *et al.*, 2009; Ball, 2021) that there was more continuity than change between the Conservative government of Major and the Labour government of Blair; “a remarkable degree of continuity ... reflecting the shift in British politics away from traditional left-right divisions and towards a neo-liberal orthodoxy.” (Crisp *et al.*, 2009, p.58). Only one element of the changes introduced by the Conservatives was reversed by Labour: the grant-maintained programme. Indeed, in response to Blair's “education, education,

education”, John Major remarked that he had the same priorities but in a different order (Major, 2011).

Conservative governments created the framework of the drive to raise standards; the Labour government intervened more directly in improvement. Leading improvement researchers, Michael Barber and David Hopkins, were hired as senior education advisers to the Department for Education, and David Reynolds led the National Numeracy Strategy.

In a White Paper *Excellence in Schools* (DfEE, 1997), published within two months of taking office, the new government committed itself to place education at the heart of government, raise standards for the many, not the few, and to put standards before structures. A Standards and Effectiveness Unit was established under Michael Barber.

Policy developments to deliver these objectives can be summarised under four main headings:

- strategies to raise standards across the piece, such as appointing a SIP to all schools (DfES, 2004a, p.9), and programmes to improve leadership (the National Professional Qualification for Headship, NPQH, from 1998) and governance in schools (Ambition Institute, 2023)
- strategies to tackle weak or under-performing schools, such as the academies programme (Education Act, 2000), designed to raise standards in areas of long-term, serious deprivation and educational under-achievement by creating new, beacon schools; and group of initiatives to tackle low-achieving schools. Schools Facing Challenging Circumstances, the Improving Schools Programme and the National Challenge were all large-scale programmes employing the concept of a floor standard: all schools failing to reach an arbitrary standard in either the primary school tests or GCSE examinations were entered into a programme of support and challenge and provided with additional resources
- large-scale curriculum programmes, especially the strategies, totalling 14 *National Strategies Programmes* (DfE, 2011a) spanning the age-range from early years to post-16, many subjects and aspects such as behaviour and attendance

- in part a response to the tragic deaths of Victoria Climbié and others, the Green Paper *Every Child Matters* (ECM) (DfES 2003) formally widened the goals of education to include children's safety, their health and wellbeing and their enjoyment of life and education. Chapters 4 and 5 show the great importance attached by my interviewees to this document, and the developments like Sure Start Centres that flowed from it.

Labour also committed itself from the outset to close gaps in attainment, the 1997 White Paper vowing "to raise the attainment of ethnic minority pupils" (DfEE, 1997).

.

The Coalition government which came to power in 2010 signalled its intentions for education, and specifically for school improvement, as rapidly and firmly as its predecessor had in 1997; within months, policy statements and rapid legislation underlined two beliefs about school improvement:

1. That it is the job of schools rather than government, local or national, to improve schools:

"the time is right for central government to step back from much of the central initiatives that have been developed over recent years and to consolidate resources and decision-making at school level, allowing schools to determine their own needs" (DfE, 2011. p.3)

In their later review of school-based improvement carried out for the government, Greatbach & Tate (2019) use the phrase "a self-improving school-led system" (p.3) to encapsulate the intention. The change in approach prompted another surge in research: Greatbach & Tate based their review on 154 studies, 87 of which took place between 2015 and 2018 (*ibid*). The government based its approach on the early work of Potter (2004), Chapman & Allen (2006) on the effectiveness of what were then called 'federations' and later re-named multi-academy trusts (MATs) (DfE 2016a), with their potential to share skilled leadership, expertise and resources.

2. That standards in English schools were not high enough and measures should be taken to drive them up. Secretary of State Gove claimed that the 'improvements' of the Labour years had in significant measure been the result of 'dumbing down' (less challenging curriculum and testing) and of 'gaming the system' to meet targets such as five or more higher-grade GCSEs. He made it clear that "the fundamental problem with our schools is that not enough of them are good enough" (Gove, 2012).

In pursuit of the first belief or objective:

- Several of the then-current requirements were dubbed “regulation and bureaucracy” and removed, such as those for schools to have a SIP and to produce a Self-Evaluation Form; the *National Strategies* programmes were abruptly ended. The Education White Paper, published in November 2010 (DfE, 2010a), strongly rejects the interventionist approaches of the previous government:

“Over recent years, centralised approaches to improving schools have become the norm. Government has tended to lead, organise and systematise improvement activity seeking to ensure compliance with its priorities. We think that this is the wrong approach. ... [it] reduces the capacity of the school system to improve itself. Instead, our aim should be to support the school system to become more effectively self-improving. The primary responsibility for improvement rests with schools, and the wider system should be designed so that our best schools and leaders can take on greater responsibility, leading improvement work across the system” (*ibid* paras. 21 & 22).

- enhanced collaboration between schools would be facilitated by freeing them from the control of LAs: interested groups were enabled to set up free schools, and it was made easy for schools to become academies, the right of governors, local authorities or parents to oppose academisation having been removed (Academies Act, DfE 2010b)
- the Labour government’s concept of ‘academy’, in the words of Secretary of State Blunkett when launching the programme, was to “break the cycle of low expectations” (Blunkett, 2000) in deprived communities. By the time Labour left office in 2010, there were about 200 academies open (Gillard, 2018). The Coalition/Conservative concept of academy was akin to that of the grant-maintained school: a strong school freed from LA control, and the intention was that academy status should become the norm. At various times in the 2010s, announcements were made of the intention that all schools must become academies. The strongest were given the power to set up MATs and take on weaker schools, becoming responsible for their improvement. The number of academies expanded rapidly, to around 2000 by 2012 and over 10,000 by the end of 2022. Around 80% of the secondary school population, and 40% of primary, attend an academy (DfE, 2022a)

- local education authorities were renamed ‘local authorities’ in 2010.

To pursue the second belief or objective:

- the performance of UK schools was compared, especially through PISA, with the highest-achieving school systems such as Finland and Singapore. The opening paragraph of *The Case for Change* (DfE, 2010a) refers to international comparative studies and to the influential McKinsey study (Barber & Mourshead, 2007) to prove that the previous administration’s strategies had not worked, and that change was needed: Secretary of State Gove reinforced the point in his speech (2012) to the Spectator Conference.

“Over the past 15 years, a number of major studies have examined systematically how well students perform in literacy, mathematics and science in different countries of the world at different ages. These studies have begun to expose how well different education systems are doing – and have cast the education debate in this country in a wholly new light. In the latest round of tests of 15-year olds (PISA), England was 17th in reading, 24th in mathematics and 14th in science – ahead of countries like Spain, the USA and Italy, but still well behind, for example, Finland, Hong Kong and Canada” (DfE, 2010a, para.1)

- the phrase ‘world-class schools’ came into frequent use, such as from the Prime Minister’s introduction to the White Paper *The Importance of Teaching* (DfE, 2010c, p.3), accompanied by ‘policy-borrowing’ from nations deemed to have world-class systems, such as from Singapore and Shanghai for mathematics
- making the argument for school-based improvement: *The Case for Change* focuses significantly (para. 61 *et seq.*) on the way countries successful in PISA tests have shifted their national education policies away from central control towards improvement programmes controlled by schools themselves:

“the top-down approach was seen to be limiting to the further development of the system. There is also a significant international trend towards greater curriculum autonomy in schools” (para. 64)

“There is strong evidence that the most effective systems in the world seek to combine significant operational independence for schools with effective accountability.” (para. 71)

The paper also (para. 79) uses evidence from the English academies programme to show that autonomous schools perform better academically

- focusing relentlessly on under-performing schools, increasing the power of government to direct such schools into MATs which would bring about their improvement or closure
- closing attainment gaps, especially that between disadvantaged pupils and their peers, particularly through the introduction of the Pupil Premium Grant in 2011
- emphasising a knowledge-based curriculum through a thorough revision of the National Curriculum in 2015. The purpose, as described in the consultation document (DfE, 2013 para. 1.1), would be to: “ensure that the new National Curriculum embodies rigour and high standards and to ensure that all children are taught the essential knowledge in the key subject disciplines”. Tests and examinations at both primary and secondary levels were made more demanding, with nine levels replacing the previous eight grades at GCSE
- emphasising the importance of subject content and knowledge through the introduction of the English Baccalaureate (EBacc), a virtual qualification achieved by gaining good passes in each of English, mathematics, sciences, history or geography and a modern foreign language
- for the first time, determining acceptable pedagogy, such as the requirement to teach synthetic phonics and promoting ‘mastery’ programmes, especially in mathematics. Through innovations such as the Education Endowment Federation and Oak Academy, the government promotes targeted research and development to improve classroom practice, especially to raise the standards achieved by disadvantaged pupils
- building the means of system improvement. The Labour government had introduced the National Leader of Education (NLE) in 2006, high-quality leaders of high-quality schools who wished to help increase the leadership capacity, and thus to help raise standards, in other schools. The Coalition built on this programme by enhancing leaders’ roles and adding further components to system leadership:
 - National Leader of Governance (2012) to play a role for governance analogous to that played by NLEs in school leadership
 - professional leadership qualifications for senior and middle leaders
 - Teaching Schools (Academies Act, 2010), a network of outstanding schools which would provide high-quality training and development to school staff and support to schools and teams in difficulty

- Multi-Academy Trusts (MATs): from early in the history of academies, some had worked together, taken over weaker neighbours or cloned themselves. Research within the DfE Academies Division (Potter 2004) pointed to the success of a number of such school federations in turning round weak schools. Federations became academy chains and then MATs and have become the government's preferred delivery model for the improvement of weak schools. Schools which are deemed inadequate are compulsorily converted into academies and placed with a MAT chosen by government (DfE, 2016a).

The contribution of Ofsted

Before the creation of Ofsted in 1992, school inspection was haphazard. In my 18 years in teaching, none of the five schools in which I taught was inspected. The annual inspection of all schools, established in the mid-nineteenth century, was abandoned in 1902, and for the next 90 years schools had been inspected, if at all, by a small national force of HMI, or by LA inspectors, to individual and mostly unpublished standards (DES, 1970; 1982).

Although under the terms of the Education Act 1944 the Secretary of State was charged with the duty to "cause inspections to be made of every educational establishment ... at appropriate intervals" (Ministry of Education, 1944, p.56, Sec.7(2)), the force of HMI who carried out these inspections was too small to inspect every school (DES, 1970). When schools did receive a full HMI inspection, the criteria for judgement were unpublished and inconsistent between schools, and the reports were confidential and unpublished until 1983.

The Education Act 1992 mandated frequent, systematic inspection of all maintained schools. Ofsted brought standard, published criteria for judgement; standard formats for inspections and reports; a regular cycle of inspection; published reports, including a summary sent to all parents; and a requirement for action plans to be written by the school to respond to the issues raised by the inspection, also made available to parents. Schools were categorised on a 5 (later 7) point scale. Inspectors were required to report when they judged schools to be 'at risk': either as having 'serious weaknesses' (weaknesses so serious that the school is unlikely to escape from them unaided) or 'requiring special measures': failing to provide its pupils with an acceptable standard of education and not possessing the capacity to make the

necessary improvements (Ofsted, 1995). Between 1992 and the first inspections in 1993, Ofsted produced its *Handbook for Inspection*, the manual for inspectors and schools laying out the purposes and conduct of inspections; this contained the *Framework for Inspection*, a statement of the content of the inspection and the report, the judgements to be made and how they were to be arrived at, and a summary of such law as related to each area. The *Framework* has been modified frequently in the subsequent 30 years, often in detail, sometimes, such as in 2005 and again in 2019, radically.

Ofsted's engagement with school improvement has varied: the 2022 version of the *Handbook* does not refer to the aims of inspection, while earlier versions did: in the 2013 version (Ofsted, 2013, p.4), one of the three purposes of inspection is listed as to "promote the improvement of individual schools and the education system as a whole" (the other two purposes are to provide information to parents and to parliament) and goes on to explain how it "drives and supports" (*ibid*) school improvement:

- "raises expectations by setting the standards of performance and effectiveness expected of schools
- provides a sharp challenge and the impetus to act where improvement is needed
- clearly identifies strengths and weaknesses
- recommends specific priorities for improvement for the school and, when appropriate, checks on and promotes subsequent progress
- promotes rigour in the way that schools evaluate their own performance, thereby enhancing their capacity to improve." (Ofsted, 2013a, p.4)

As Rosenthal (2004, p.143) notes, "The general philosophy of [Ofsted's] approach is well described by its motto: "Improvement through Inspection". This strapline was only in use for a small part of Ofsted's life, being replaced by "raising standards, improving lives".

As Reynolds *et al.* (2012) note, from quite early in its life, Ofsted based its criteria around educational effectiveness research, in particular commissioning a review of research leading to a composite set of 'school effectiveness criteria' from Pam Sammons (Sammons *et al.*, 1995).

Research evidence of the impact of inspection on school improvement is varied, cautious and not strongly positive: the National Audit Office (NAO, 2018) began its report by saying that “The ultimate measure of the value for money of Ofsted’s inspection of schools is the impact it has on the quality of education, relative to the cost”, and concluded that:

“Ofsted does not know whether its school inspections are having the intended impact: to raise the standards of education and improve the quality of children’s and young people’s lives” (para. 20).

However, “There is some evidence that inspections are helping schools to improve” (*ibid* para. 22), a judgement based only on the perceptions of headteachers. Rosenthal (2004, p.143) claims the opposite but is cautious: “there exists a small but well-determined *adverse*, negative effect associated with the Ofsted inspection event for the year of the inspection.” The report into inspection commissioned by Ofsted from NFER (McCrone *et al.*, 2009) was more positive: “Inspection was generally perceived to have achieved a direct positive impact on school improvement in terms of assessment and, to some extent, quality of teaching, and to have contributed to improved attainment” (p.iv). Case, Case & Catling (2000) deliver a bitter attack on Ofsted’s impact on primary teachers, concluding that:

“Despite the evident intensity of the Ofsted experience, teachers in our study uniformly indicate that, one year after inspection, it has had no lasting impact on what they do in the classroom” (Abstract, p.605) and that “Ofsted is stage-managed public ‘accountability’” (*ibid*).

Most evaluations of the impact of Ofsted, like those cited above and that by Scanlon (1999), focus on the impact of inspection itself on schools and teachers; it is unsurprising that teacher perceptions tend to the negative. A much greater impact comes from schools’ close attention to Ofsted’s criteria in editions of the *Framework* and in inspection reports, and adapting to changes therein. Following race riots in the north of England in 2001, the government introduced a legal duty upon schools to promote community cohesion; Ofsted was required to monitor and report on how effectively they did so (DCSF, 2007). Ofsted removed this criterion from its schedule in 2010, since when I have seen not a single mention of community cohesion in any of the several dozen school improvement plans I have read.

A much larger-scale example of how Ofsted criteria drive school priorities relates to the curriculum. Earlier versions of the *Framework* placed the curriculum as a ‘contributory factor’ (Ofsted, 1992) and later as one aspect of leadership and management (Ofsted, 2015, p.12). In 2019, Ofsted announced a major revision of the *Framework*, arguing that:

“the main thrust of the typical inspection conversation has come to be about recent outcomes, assessment of current ‘pupil progress’ and expectations of future progress. Schools have responded to this with workload-intensive management models that focus on data and prediction..... We accept that developments in our inspection practice over time have contributed to this imbalance in the accountability system.” (Ofsted, 2019b, p.3)

This acceptance that the Ofsted process had contributed to a distortion in school practice, to schools being obsessed with the collection and analysis of pupil performance data, led Ofsted consciously to re-focus schools on what is taught and learned: Ofsted believed that “a focus on substance, on the knowledge and skills that we want pupils to acquire, is often lost” (*ibid* p.4). The subsequent focus in schools on ensuring that subject leaders can articulate the intent, implementation and impact of their curriculum is ubiquitous.

2.6 SUMMARY: KEY ISSUES ARISING FROM THE REVIEW; THE STATE OF KNOWLEDGE IN THE RESEARCH FIELD; LINKS WITH THE RESEARCH QUESTIONS

The key ideas to emerge from this review together summarise the state of effectiveness and improvement research.

1. It is argued, including by Reynolds (2023), that school effectiveness research is “complete”: barring minor refinement, we have a secure understanding of what characterises an effective school, one that produces better outcomes for its students than other schools with similar intakes.
2. Reynolds (*ibid*) further suggests that school improvement, the implementation of the known principles, is a “weak link”. The ‘what’ is well known, the ‘how’ less so.
3. Other authors, notably Hopkins (2014), claim that we have considerable understanding of how to improve individual schools, but that our ideas of system-wide improvement are less developed.

4. However, significant research such as that by Barber & Mourshead (2007) clarifies the characteristics of effective school systems. What appears to be missing is therefore principles of implementation.
5. A number of the key ideas within school improvement are not defined with clarity: this is partly attributable to the lack of nationally determined aims for education. As a result, improvement is frequently analysed narrowly, generally in terms of test/examination performance, and thus it is unclear whether aspects of rising importance, such as young people's welfare and mental health, are included within 'school improvement'.
6. Likewise, 'consistency' is a term used frequently in the field of effectiveness and improvement, but one which is rarely defined or circumscribed: which aspects of schooling benefit from consistency is an issue to be addressed.
7. Because of the narrow focus, school improvement literature neglects what I have termed 'underlying improvement': aspects of schooling that make schools better places in which to learn.
8. Analyses in the literature of the causes of the surge in school improvement activity in the last quarter of the twentieth century are partial and piecemeal.
9. The literature does not provide an explanation for the rapidity with which school effectiveness and improvement activity spread and diversified.
10. There is an ongoing debate about the reality, not of school effectiveness research, which is not seriously disputed, but of school improvement: there are relatively few but vocal and articulate school improvement deniers.

The following five paragraphs are reflections on the influence of these summary points on my research questions.

The first 'question' concerns the validity of the untested underlying proposition that improvement dominated educational thinking and action in the period. Evidence from the work of Hallinger & Kovacevic (2019) strongly supports the proposition; evidence from interviews is necessary to confirm or deny the proposition.

The second research question asks why this was the case: what combination of external (socio-econo-political) and internal (educational) factors led education to be dominated by

school improvement during the period? Evidence from the literature review (point 8 above) is patchy and not joined up: this research will combine the literature evidence with that from interviews, especially with those with a background in social, economic and political history, to create a more comprehensive analysis, classifying the causes and showing how they combined to drive improvement.

Point 9 above states that the literature lacks an explanation for the rapid evolution of the school improvement paradigm and its relationship to the effectiveness 'paradigm'. the research will seek an explanation for this, the third research question.

The largest single component of the research will concern the fourth question: to identify what key groups of educators (school leaders, policy makers at local and national level) have understood by 'improvement', and the actions taken as a result. Points 1 – 6 above form a chain of reasoning suggesting that clarifying how school leaders and others interpret 'improvement', and how they bring it about will lie at the heart of the thesis. Since it would be idle to suggest that genuine school improvement (as opposed to apparent improvement through outside influences such as a major change in a school's intake) happens spontaneously, the *leadership of improvement* will be a major theme.

The final research question, identifying the impact of fifty years of school improvement, and implications for the future, responds in particular to point 10, but also engages with complexities of definition raised by point 5.

CHAPTER 3: METHODOLOGY

3.1 INTRODUCTION

The chapter begins with a restatement of the research questions and an account of the structure of the chapter, the bulk of which details and justifies the main research design choices (3.2), with an analysis of their attendant problems and limitations (3.3) and of ethical consideration (3.4). Decisions are summarised (3.5).

Table 2: Summary of research questions and methodology

Research questions	Methodology
1. To test the extent to which school improvement has dominated educational thinking and action in the period.	Scrutiny of research in both scholarly and especially grey literature Interviews, especially with school leaders
2. To ascertain why this has been the case: what combination of external (socio-economic-political) and internal (educational) factors led education to be dominated by school improvement during this 50-year period.	Literature review Interviews, especially with policy makers
3. To analyse and explain the rapid evolution of the school improvement paradigm. Identifying how it related to the pre-existing school effectiveness paradigm.	Literature review Interviews, especially with researchers and policy makers Autoethnography *
4. To identify what key groups of educators (researchers, policy makers at local and national level, school leaders) have understood by 'improvement', and the actions they have taken as a result.	Interviews, especially with headteachers Autoethnography *
5. To identify the impact of fifty years of school improvement, and implications for the future.	Literature review and interview, but particularly reflection on all evidence Autoethnography *

* Elements of autoethnography validate and triangulate researched data (explained p.63 *et seq.*)

3.2 RESEARCH DESIGN DECISIONS

Decisions concerning the research design (3.2) analysed and justified here are:

Research philosophy (3.2.1): “what the researcher perceives to be truth, reality and knowledge” (Ryan, 2018, p.2). When the research questions are addressed to the phenomena being researched, which approaches (quantitative/qualitative, positivist/interpretivist; inductive/deductive) are appropriate and why?

Research strategy (3.2.2): the plan of methods, describing and justifying how the research was conducted and the data collected (including the sampling strategy).

Time Horizon (3.2.3): decisions about the period over which the data were collected.

Data analysis (3.2.4): the choice of methods explained and justified.

3.2.1 Research philosophy

The phenomena studied here are school effectiveness and school improvement; the philosophy of this study mirrors that of the research which it investigates. From data generated by sampling the literature of the subject and by interviewing a sample of practitioners, I developed theories or models of how schools improve: an *inductive* approach. Research studies into effectiveness and improvement have, over 50 years, also been predominantly inductive: samples of schools were studied to ascertain which attributes correlate with high performance and theory was developed from the sample. The methodology of the LSES is typical: identify a number of schools whose performance makes them, over time, positive or negative outliers in their socio-economic group, and research the characteristics they have in common.

I carried out the work in the knowledge that a different sample might generate some different data, and that a different researcher might draw some different conclusions: thus this study is also *interpretivist*. Effectiveness and improvement research is similar, accepting that there are multiple versions of reality, and that different researchers and practitioners see the concept of school improvement differently, based on their context and experience; thus they are also interpretivist.

School effectiveness studies may be correlative by nature, seeking to identify which attributes of schools are associated with high performance, and so involve a measure of quantification,

but in essence they are *qualitative*. They gather and analyse data about school attributes such as values, teaching styles and quality, and leadership, often in the form of opinion. The current study is similarly qualitative.

In detail, however, the journey from the first research question to the last represents one from a more deductive/positivist approach to one which is inductive/interpretivist, although it remains qualitative throughout.

The first question takes the hypothesis that the theme of improvement dominated schools during the period from around 1970 and “designs a research strategy to test [it]” (Wilson, 2010, p7), falsifying or corroborating it: a definition of *deductive* method. Although not quantitative, as such studies most commonly are (Soiterman, 2010), the process drew evidence from research literature and the experience of practitioners to test the hypothesis. Incidentally, by identifying the period of rapid change and focus on improvement in education in England, the response to the first research question also explained and justified the choice of research period: 1970 – 1920.

By association, the same part of the study is *positivist* in that it pursues a verifiable answer to a question. The questions in question 1 are more susceptible to such answers than other parts of this study but, by asking “to what extent”, the question is not liable to a yes/no answer but rather one in the form “to a considerable extent”. The great majority of energy and activity in a school is devoted to the maintenance function: the organisation and delivery of teaching and care to students. The margin of time, energy and money left over may, at any one moment, in any one school, be devoted to issues such as finance, staffing or safeguarding. Answering this first research question positively, as I do, means that, over time and across schools, this precious margin of ‘discretionary time’ was spent predominantly on issues of effectiveness and improvement.

The questions associated with question 2, which seeks reasons for the surge in importance of school improvement, retain elements of the deductive approach: they relate to the development of effectiveness and improvement over time, and parts of the answer are historical ‘truths’. However, the gap in the research literature lies in the identification of contextual factors: the range of social, economic, political and educational factors that drove

the burgeoning of improvement. This is not the testing of a theory but the building of one from a diverse evidence base: a range of documentary sources with divergent theories of economy, society and politics; and evidence from interviews, the views and lived experiences of a diverse group of educationists. The task of creating a coherent and comprehensive account of the context is thus inductive, developing a theory; by drawing on a range of views and imposing a framework on them, it is also interpretive.

The questions related to the second question attempt to explain origins ('why' questions) those around the third seek to explain the rapid evolution of the school improvement paradigm: 'how' questions. The journey through this third question to the final two shows the increasing predominance of an inductive, interpretivist approach. Questions related to question 4 (what educationists have understood by 'improvement') are answered mainly through interview, drawing on policymakers', researchers' and school leaders' understandings of effectiveness and improvement, how they came to those understandings and translated them into action. The process of eliciting these opinions and crafting them into a coherent narrative and theory is fundamentally inductive and interpretivist.

The final question uses all the evidence from questions 1-4, combined with the author's diverse experience of effectiveness and improvement policy and practice, to evaluate the impact of school improvement: again interpretivist.

As noted, the study researched choices and opinions; I chose interviewees and the questions to ask them; I classified data in particular ways to generate hypotheses. All of these involve the researcher in interpreting the field and are thus by definition inductive and interpretivist in approach. This is a study which, by using qualitative data from semi-structured interviews and autoethnography as its methodologies, actively promotes the notion that reality is shaped by the observer and is therefore liable to various interpretations. The Saunderson *et al.* (2019, p144-145) analysis of research philosophies describes the axiology of interpretivist studies as typically involving the researchers in the study; its autoethnographic component clearly demonstrates that this is true of the current study, which is necessarily selective and, to that extent, subjective. Both Hudson and Ozanne (1988) and Saunderson *et al.* (*ibid*) describe the ontology of interpretivism as complex, multiple and relative, since such an approach treats the phenomena being studied as undergoing revision through interaction with researchers' different interpretations. Finally, Interpretivist epistemology, how it defines

acceptable knowledge, generalises to construct theory from observation. It is subjective in its analyses, accepting that researchers will, through their interactions, affect the subjects of the research and vice versa. As a result, reality in this study is necessarily co-constructed (Berger & Luckman, 1969).

Saunders *et al.* (2019, pp.154-5) describe the strength of induction as its ability to enable prediction: generalising from many observations makes it possible to state what is likely to happen. Observing a thousand white cats doesn't prevent the thousand and first being black, but the next cat encountered is *likely* to be white. The approach also promotes flexibility, the variety of interpretations provided by different researchers strengthening the resulting theory. Effectiveness and improvement studies work this way and have been cumulative; by drawing on them, and by working with a sample of interviewees, so is the current study.

The inherent weakness of an inductive approach is, *mutatis mutandis*, the potential for inaccuracy through incompleteness. Hume's challenge to induction was that it lacks a chain of reasoning, merely aligning to the Uniformity Principle: what happens is assumed to be the same as what has happened before (Hume, 1739, Book 1, part iii, section 6). Had researchers studied other schools or interviewed other leaders, they might have found conflicting data; had they correlated different variables, they might have generated different hypotheses. As the sub-section about interviews shows [p.61 *et seq.*], selecting interviewees and shaping conclusions from the analysis of interview transcripts by a process of coding, necessarily encounters the Hume challenge inherent in all inductive studies; [p.68] shows how I sought to mitigate the problem.

3.2.2 Research Strategy

The study uses three methods: scrutiny of documents, both grey and scholarly; semi-structured interviews and an element of autoethnography. Each question was investigated using all three methods [table p.55], but in different proportions.

In the case of the first question, the views of school leaders gathered in interviews provided an important perspective on what had dominated their thinking and actions, while the review of literature provided historical evidence. Research into the origins of the school improvement 'movement' (second question) also leaned heavily on literature, while most of

the evidence for the third and fourth questions was drawn from interviews. Findings related to the final question are based on reflection on the evidence from all sources.

The autoethnographic component, my own experience in the field, informs all questions and acts to verify or triangulate the researched material. An account of this experience and a justification for its use in the study, appears on [p.67 *et seq.*].

3.2.2.1 Documentary review

Here I categorise the literature studied, and the filtering techniques and search terms used to reduce the potential volume to manageable proportions.

Four types of grey literature are significant to the study:

- Political papers: major party manifestoes for general elections; education White Papers and legislation. These denote the changing political prominence of education, the changing profile of school improvement, and current political thinking about how best to bring about and support improvement (especially questions 1, 2, and 4).
- Educational outcome data (test and examination results, attendance data) track improvement (especially research question 5). The choice of what to measure and publish conveys agencies' concept of significance, what they see quality and improvement to be about (research question 4).
- Evaluation reports: the school inspection reports of HMI were infrequent and had restricted circulation. The fact of universal, published inspection reports since 1993 is a major indicator of the rising social and political importance of school effectiveness; that such inspections were to be repeated every few years demonstrates that changes in school effectiveness over time (improvement) was seen as equally important.
- Two Ofsted publications were analysed to track changes in the indicators of quality: the annual reports of Her/His Majesty's Chief Inspector (HMCI) and the numerous editions of the Framework for Inspection, the inspector's rule book. Trends within these two sources show how the concepts of quality and improvement evolved (especially research questions 3 and 4).

The scholarly literature reviewed falls under two main headings:

- Effectiveness and improvement literature, the analysis of which enabled the evolution of the paradigms to be tracked (research questions 1 and 3). The sample includes both the

flow of new literature, especially in key journals, and meta-analyses summarising the evolution of research knowledge. Major journal websites have a key word search function; using the same search terms (effectiveness, improvement, quality, standards) to scan consecutive periods enabled analysis of changes in the interpretation of quality.

- Analyses of social, economic and political change to identify the social trends which, together with elements of changing educational thought, provide the context promoting the rapid expansion and evolution of effectiveness and improvement. Although relevant ideas from these fields appear in the education literature, it was necessary to go to specialist literature and to interviewees to fill in much important detail.

3.2.2.2 Semi-structured interviews

Interviews with leading practitioners lay at the heart of the research. Semi-structured interviews were chosen to secure comparable data from the standardised interview, and rich detail from freer discussion (Adams, 2015). Hopkins *et al.* (2016, p.8) define such interviews as “using key questions and associated probe questions designed to elicit [the interviewee’s] thoughts and feelings linked to the focus areas”. A core of questions (Appendix III) was asked of all interviewees, with probing supplementaries which varied with the individual’s experience and responses to core questions. The analysis of semi-structured interview technique by Cohen & Crabtree (2006) emphasises conversational, two-way communication. They list richness (the ability to go into detail) and the balance between structure and freedom, preparation and spontaneity as representing the best of both worlds: reliable, comparable data, not just answers but the reasons for the answers, new information and ‘leads’ for the researcher. They stress the need to engage sufficient, diverse people to make comparisons, emphasising “probing, open-ended questions”; they suggest that semi-structured interviews are especially appropriate when interviewers “need maximum latitude to spot useful leads and pursue them” (pp. 493-494). As a result of these strictures, interviews were designed as follows:

Questioning

The seven core questions (Appendix III) were broad and open-ended, allowing each interviewee to draw on experience.

Question 5 (Any key moments/critical events in your own involvement with school improvement?) differed in style, designed to stimulate vivid memory and contribute to the authenticity of the study. It was placed centrally in the list to disrupt what might otherwise have become a linear, convergent sequence of ideas. The intention was to ‘flip’ the interviewee’s thinking, to provide graphic illustration and to re-channel thinking during subsequent questions. The questions were usually posed in order, but this varied because some interviewees’ response suggested a different order.

I considered the use of focus groups since these can lead to stimulating debate and to the ideas of one member triggering ideas in others (Accella, 2011), but decided against, chiefly because I judged this potential to be outweighed by that for opening up issues confidential to the individual. With one exception: during the data collection phase, a school I had supported during its move from dysfunction to soundness was re-inspected and judged outstanding. I took the opportunity to run semi-structured interviews with members of the leadership team, individually or in small groups, touching on the other questions but focusing on their analysis of the critical factors underlying continuous improvement sustained over more than a decade.

Interviewee sample

Whereas “probabilistic or random sampling is used to ensure the generalizability of findings by minimizing the potential for bias in selection” (Palinkas *et al.*, 2015, p.535), such an approach is not possible in the context of this study: the concept of a representative interview sample of school leaders, who number well over 50,000 (approximately 25,000 schools with a minimum of two leaders each) makes no sense; nor does such a sample of researchers or government officials. Neither was a random sample desirable: the purpose of the study was not to come to conclusions about the distribution of school improvement, nor its frequency, but rather to gain understanding of its processes and effectiveness.

My interviewees therefore constituted a non-probability sample, selected in a non-randomised and non-representative way: the sampling strategy could best be described as ‘purposive’, a technique that involves judgement on the part of the researcher “to select respondents that are most likely to yield appropriate and useful information” (Kelly, 2010, p.317), and are thus likely to use limited research time effectively (Palinkas *et al.*, 2015).

In purposive sampling, interviewees are selected based on specific characteristics relevant to the research questions. I therefore approached:

- school leaders who had been successful in bringing about improvement in one or more schools
- government officials, local and national, who had been successful school leader/improvers and whose knowledge had been increased by working to promote and observe improvement in a variety of schools across the nation or region
- researchers whose work made a significant contribution to the study of school improvement.

Accordingly, I used my experience and range of contacts to select interviewees who met these criteria and who, in addition, I knew to be thoughtful, analytical leaders in their fields. Because of the richness of their experience, I accepted recommendations for additional interviewees from these known and trusted sources.

The sample could be characterised as ‘rich’: as Campbell and his collaborators (Campbell *et al.*, 2020) express it “the reason for purposive sampling is the better matching of the sample to the aims and objectives of the research, thus improving the rigour of the study and trustworthiness of the data and results” (p.652).

My approach cannot be accused of ‘convenience sampling’, the choice of interviewee based on ease of access, on geography. Such a methodology is advantageous only in terms of speed and cost (Give, 2008), criteria not relevant to me. It is true that the interviewees are mainly from the southern half of England, but this reflects the criteria of *purposive* rather than of *convenience* sampling.

The potential flaw in such sampling is self-evidently its propensity to bias: the choice of respondents by a single agent, albeit from a rich and extensive background, makes non-randomised samples liable to the charge of various forms of bias (Palinkas *et al.*, 2015). I tested my sample and sampling strategy against the criteria of the fourteen types of bias listed in the Palinkas article. Two stood out as significant risks which I sought to mitigate.

In the choice of interviewees, *selection bias*: reasons for the non-random selection of participants have been explained above. The threat is choosing participants known to have convergent views. Mitigation involved:

- increasing sample size from the initial group of around 20 to 32, mainly by recommendation. especially of individuals identified as holding contrary views
- the process of asking interviewees to identify others who would make a valuable contribution can constitute ‘snowballing’, gathering similar material and travelling in one direction: this is emphatically not the process pursued, as I have consistently sought divergent views and especially from school improvement deniers, the relatively small but vocal group of academics who argue that school improvement is illusory, or even a contrivance. Since I was not able to interview members of this group, I sought their views in print especially carefully
- seeking respondents from a variety of backgrounds, types of school, roles in government, to achieve breadth and balance of views
- likewise, seeking the views of different research ‘schools’: older and newer, those who have driven improvement and those who are ‘deniers’⁸
- “Making explicit the approach used for participant sampling provid[ing] improved methodological rigour” (Campbell *et al.*, *ibid*), thereby maintaining transparency concerning the selection process.

During interviews, *confirmation bias*, “the tendency to search for, interpret, favor [*sic*], and recall information in a way that confirms one’s preexisting beliefs or hypotheses” (Nickerson, 1998): mitigation here involved:

- avoiding leading questions: all questions are open, leaving each interviewee to respond as s/he wishes
- actively seeking divergent views, because of the richness, depth and nuance a range of views brings to the analysis
- rigour in trawling interview transcripts to identify different views and represent them.

- maintaining what Finlay (2002) terms ‘reflexivity’: frequent checking on my assumptions throughout the research process, questioning them and considering alternatives
- maintaining the same transparency over the collection and analysis of data as was used in the selection of interviewees [p.63]: “Clearly documenting and explaining the rationale behind the choices made during the research process enhances the credibility and methodological rigor of the study” (Campbell et al., 2020b).

The process of asking interviewees to identify others who would make a valuable contribution can constitute ‘snowballing’: gathering similar material and travelling in one direction: this is emphatically not the process I have pursued because of the rigour in trawling interview transcripts to identify different views and represent them. I have consistently sought divergent views, especially from school improvement deniers, the relatively small but vocal group of academics who argue that school improvement is illusory, or even a contrivance. Since I was not able to interview members of this group, I sought their views in print especially carefully.

Interviews were conducted with 32 individuals, amounting to over 50 hours of discussion. The interviewee sample was designed to have a high proportion with school leadership experience (around 80%). The perspective of school leaders, accountable as they are through Ofsted, governors and parents, is uniquely powerful in evaluating school improvement.

The sample layers other perspectives over school leadership, forming a web of varied and profound school improvement experience. Many, including the author, have experience that spans most of the categories that follow: the one-line *curriculum vitae* (CV) provided for each interviewee in Appendix IV shows the diversity and richness of their backgrounds:

- nearly 80% have experience as school leaders, most in more than one school
- 40% have experience of system leadership, working to a significant degree beyond their own setting to support the improvement of others
- a third has experience as CEO or chair of a MAT
- 20% have undertaken research in a field related to effectiveness and improvement, a few of them among the world’s most respected researchers in the field

- over 20% have experience as the most senior officers of local (education) authorities while 25% have worked as national government officers and advisers, including two senior advisers within the DfE
- 20% have been Ofsted inspectors, half of them lead inspectors.

While transcribing and analysing the interviews, I found that I had subliminally chosen many with experience of leading a school in severe difficulties: all but two of the school leaders interviewed (therefore over half of the total sample) have led a school in serious weaknesses or special measures. I challenged myself as to whether this unbalanced the sample and rejected the challenge because:

- all those interviewees had also led an Ofsted ‘good’ school, so the skew is illusory
- the experience of leading a school living under the Ofsted monitoring regime provides valuable insight into what is meant by improvement and how to bring it about in a high-stakes, urgent and public setting.

Even more significant to the depth of experience in the sample than whether they have lived under a ‘special measures’ judgement is the degree of challenge provided by the school. There is an extensive literature (Ofsted, 2018; IDACI, 2020; Ehren *et al.*, 2023) concerning the correlation between the level of deprivation and the Ofsted ‘overall effectiveness’ judgement: only one of the sample headteachers’ schools placed in special measures served a community of average SES and an intake of average prior attainment. All the others led schools with very high levels of deprivation and social challenge; one talked of the straightforward-ness of improvement in what he called ‘unconstrained’ schools, where pursuing effectiveness principles will generate improvement; and of how much more complex improvement is when faced by the multiplicity of factors in ‘constrained’ settings. Perhaps it is fairer to describe ‘added constraint’, over and above those experienced by all schools, such as budget, recruitment and the pandemic.

My conclusion was that the sample was rich rather than skewed: the breadth of interviewees’ experience, the contrasts among the schools they led and the high-stakes publicity they received provided vivid as well as authoritative evidence of the realities of school improvement.

The list of interviewees (Appendix IV) is anonymised, and I have not noted individuals' genders: the Conclusion [p.169] comments on the skewed gender, ethnic and geographic balance of the sample. The protocol papers sent to interviewees appear as Appendices I & II.

During the interview period, I made two adaptations to the basic pattern: I ran three pilot interviews, asked for feedback, and tweaked the process accordingly, such as by adding an invitation to reflect on the interview and communicate additions, withdrawals or amendments. In later-stage interviews, in addition to the core questions, I took the opportunity to float emerging patterns, to gain interviewees' perspectives.

3.2.2.3 Autoethnography

Autoethnography is "an approach to research ... that seeks to describe and systematically analyse personal experience" (Ellis, Adams & Bochner, 2011, Abstract). Laslett (1999) asserts that it "offers a new vantage point from which to make a unique contribution to social science". Because of the breadth of my experience, set out in Chapter 1, I found the opportunity to make my distinctive contribution appealing, and this study was initially conceived to be predominantly autoethnographic but, because of the concerns listed below, I reduced that reliance considerably.

Through original research, inspection, devising policy nationally and locally, observing and coaching teachers and leaders, since 1984 it has been my job to help schools improve, specifically through:

- the observation and evaluation of over 5000 lessons
- critical reads (and writing) of over 200 school self-evaluation documents and improvement plans
- the inspection of nearly 40 schools and Ofsted-standard reviews of around 50 more
- coaching over 100 headteachers to 'put their best coat on' for inspection
- supporting governing bodies in appointing 30 heads and in moving a similar number on
- training and coaching over 50 inspectors for the LA and Ofsted, monitoring reports and judgements
- acting as a governor in four schools.

While autoethnography provides the ideal vehicle for such a practitioner to reflect analytically on a lengthy involvement with effectiveness and improvement, I have four concerns:

- Plausibility: the key question, posed by Poulos, is “Can I really write about myself? Does that count?” (Poulos, 2021, p.3). There is a natural reticence about using one’s own life experience as either typical or illuminating.
- Memory: I am concerned that the experiences to which I refer should represent accurate rather than convenient memory, and there is copious research on the unreliability and degradation of memory.
- Oppositionality: autoethnography is arguably a manifestation of post-modern rejection of dominant theory and conventional method (Richardson, 2000). “By writing themselves into their own work as major characters, autoethnographers have challenged accepted views about silent authorship, where the researcher’s voice is not included” (Holt, 2003, p.2). The use of the verb “challenged” in this text typifies this opposition to traditional methods, described in autoethnographic literature as ‘canonical’ (Spry, 2001; Ellis, Adams & Bochner, 2011; Denshire, 2013), a term which I find to have pejorative overtones.
- Struggle to be heard: the issues addressed by autoethnographers are frequently impassioned attempts to give a voice to people and issues that have previously lacked one. Such writing forefronts the life experiences of individuals with a strong personal relationship to a significant societal issue, providing “pathways for stories of the socially marginalized to enter the discourse of academics” (Marx, Pennington & Chang, 2017). Thus the experience of those who have undergone abuse, gender ambivalence or adverse early-life experiences feature strongly and show evidence of a struggle to be heard in academic circles. While respecting these authors and their views, this thesis does not set out to challenge existing approaches to accounts of school improvement; it rather seeks to cast new light from the point of view of a practitioner in many roles over a long period. The more dispassionate focus of the current work avoids the complex relational ethics implicit in much autoethnography (Ellis, 2007), and lacks the focus on “emotional, sensitive topics” (Ellis, Kiesinger & Tillmann-Healy, 1997, p.121).
- Fictive: autoethnographic writing frequently constructs imaginary narratives (Denshire, 2013; Clough, 1999); ‘fictive’ carries overtones of ‘fabricated’ (Cambridge English

Dictionary). I view a preoccupation with detail and accuracy as essential characteristics of sound research.

Given my professional background, it is perhaps inevitable that I am cautious about the ability of autoethnographic writing to sustain academic rigour in the study of school improvement. By triangulating personal experience with established literature and with the experiences of influential education leaders, this study attempts to benefit from the insights of both epistemologies: autoethnography's authentic personal experience and more traditional research methods.

I have therefore sought to optimise the power of autoethnography to bring authenticity and vividness in two ways: first, using my own experience to triangulate and ratify ideas from interviews and literature, and second, seeing the value of my own recollections, I asked my interviewees to do likewise, inviting them to identify 'flashbulb memories', a term coined by Brown & Kulik (1977) to describe moments of considerable consequence that endure in the memory with particular vividness. There is substantial evidence (Finkenauer *et al.*, 1998) that personal significance (consequentiality) and proximity to the event are factors influential in fixing the memory and in sustaining its accuracy for longer periods. The individual nature of the memories creates unique insights that are a strength of autoethnographic writing, and interviewees' flashbulb memories feature occasionally in the text.

3.2.3 RESEARCH TIME HORIZON

For reasons explained in Chapter 1, the research covers the period from 1970 to 2020. Occasional references appear to events before 1970 and post-2020, but the publication of the *Black Papers* (Cox & Dyson, 1969-74), a watershed in the rising disquiet about education, and the pandemic, which has changed so much in society, form a frame to the study.

Although data were collected within a short period and therefore might be seen as cross-sectional, the data from interviews are based on reflection and provide evidence of the key period of rapid change described above; the data collected are thus essentially *retrospective*, describing a dynamic situation.

3.2.4 DATA ANALYSIS

Transcripts of over 50 hours of semi-structured interviews, combined with notes from many written sources, generated many tens of thousands of words which, triangulated against my personal experience, created a considerable reservoir of qualitative data. Authors define the analysis of qualitative data as “bring[ing] meaning to a data set including a wide range of materials observations, and ... interviews” (Lester, Cho & Lochmiller, 2020, p.96). The notion of finding meaning in, or bringing meaning to, such datasets is to be found in all definitions of data analysis, meaning construed in this context as contributing to answering the research questions. Patton (1980, p.306) noted that

“inductive analysis means that the patterns, themes, and categories of analysis ... emerge out of the data rather than being imposed on them prior to data collection and analysis”.

Induction also, I believe, implies that the researcher, by re-reading the data, may identify “patterns, themes and categories”, but that these are in part characteristics of the data, in part a construct of the researcher. Savin-Baden & Major (2013, p.452) balance the notions of patterns existing within datasets with patterns developed from them by researchers; they assert that “analysis involves uncovering patterns in data and interpretation involves uncovering meaning”.

The risk remains, even in this perhaps overly neat distinction between analysis and interpretation, that data can be forced to fit an emerging theory; I have attempted to mitigate this risk by remaining constantly alert to the threat, remaining open to the possibility of different interpretations: my initial concept of how schools in deep difficulties were turned round by new leaders was a simple one, based on the see-saw model [p.143], but discussion with a number of leaders led to a much more refined, varied version [p. 136 *et seq*].

The analysis process used in this study was based on a distillation of authors’ accounts of the various types of textual analysis. Lester, Cho & Lochmiller (2020, pp. 98 – 102) make fine distinctions between discourse analysis (analysis of the language used in interviews); narrative analysis (what people say when they tell stories); content analysis (what is being said); and thematic analysis, the identification of patterns and themes. The analysis process used here is hybrid, acknowledging the significance of the words interviewees use (discourse analysis) and the illustrative examples they choose (narrative analysis). The study draws some

conclusions from these, such as interviewees' consistent use of adjectives like 'remorseless' and 'relentless' to describe the pace of change (discourse) [p.81], and the impassioned references of many to their poor experience of teaching when they were at school themselves compared with current practice in their own schools (narrative), or to many narratives around the impact of deprivation ["twin cities", p.103]. However, the study concerns itself more with their ideas (content analysis) and the identification of commonality and difference in their interpretations of events and policies (thematic analysis), including attempting to identify what is not being said.

The analytic process used was drawn mainly from Lester, Cho and Lochmiller (2020 *ibid*) and Miles and Huberman (1994, p.9), and can be summarised:

- transcription of interviews; their careful electronic assembly and accurate filing and indexing
- editing the transcriptions by removing extraneous chat, digressions, emphasising key points, retaining the original phraseology
- coding the data: affixing "a short, descriptive word or phrase that assigns meaning to the data related to the researcher's analytic interests" (Miles and Huberman, 1994, p.100). This was carried out by thematic analysis ("approaches aimed at identifying patterns across qualitative datasets", Braun *et al*, 2019, p.844), in stages:
 - light analysis: a complete read-through making marginal notes and underlining noteworthy words, phrases, ideas; noteworthy either because they recur in the transcripts, and/or because they seem significant. Since the interviews, and thus the transcripts, were organised into my framework of seven questions, this first analysis could be described as 'etic', a term used by Cresswell (1998, p.96) to describe an analysis based on the framework developed by the researcher
 - incipient key ideas were noted: some were very specific, such as 'watershed 1980' (the point of paradigm shift) or 'un/constrained schools'; others were broader headings, such as 'school culture'. These are listed in the second column of the coding tree (Appendix VI)
 - the analysis was then re-cast, testing the validity of these incipient key, recurrent ideas *across* the seven questions. By Cresswell's (1998) definition, this was an emic analysis, one deriving from the views of the participants, identifying ideas raised by them which

cross the artificial boundaries of the seven questions. A process of aggregation joined most individual codes into categories; others, such as 'watershed 1980' remain freestanding, as the third column of the coding tree (Appendix VI) shows. Some codes, such as 'dominance of test/exam results' occur in more than one category

- the codes and categories were gradually, over several iterations of amalgamation, reduced into a small set of generalisations: themes which form the final column of the coding tree (Appendix VI). Some of the categories were of greater significance and became discrete sub-themes, such as 'headteacher leadership'.

In accordance with the evaluation criteria for rigorous qualitative research propounded by Anderson (2017, p.127 *et seq.*), I have sought to ensure that the study has clear audit trails, enabling readers to understand the pathway by which conclusions were arrived at.

3.3 METHODOLOGICAL LIMITATIONS AND THEIR MITIGATION

These were listed in Chapter 1 and throughout this chapter, and are summarised here.

- Problems inherent in an inductive study: incompleteness and selectivity, exacerbated by problems of convergence [p.63] in a modest sample.

I have mitigated these risks by constant awareness, very wide reading and by consciously seeking divergent views, in text and in interviews.

- The skew in the sample.
- My concerns over the intellectual rigour of autoethnography, and how I have circumscribed its use to validation and triangulation.

3.4 ETHICAL CONSIDERATIONS

As well as abiding by the University of Gloucestershire's regulations concerning ethical research (University of Gloucestershire, 2020), I also adhered to the ethical research guidelines of the British Educational Research Association (BERA, 2018). *Inter alia*, this included securing written, informed consent from interviewees; ensuring the confidentiality of information given; and anonymity throughout the process. The University of

Gloucestershire's Research Ethics Approval Application Form received approval from the University Ethics Board on 22nd September 2022.

It is worth mentioning that, to qualify to be a member of the sample, considerable experience as a very senior leader was necessary. By definition, therefore, no interviewee could be inexperienced or vulnerable (in the sense of being very old, young, frail, sick or mentally impaired). Participation in interview was carefully circumscribed: all participants received the information about the interviews in advance, including the areas of questioning, the guarantees of confidentiality and anonymity and the right to withdraw, as set out in Appendices I and II, and all signed their agreement to take part under the given conditions. Participation, it was made clear, was voluntary with a right to withdraw up to one week after the interview. I was the sole researcher conducting the interviews. These conditions comply with the BERA ethical research guidelines alluded to above.

No potential interviewee declined (though four did not reply to the invitation), none circumscribed the use of their views or words, and only one withdrew before interview, through illness. All were asked whether they wished to qualify or withdraw anything said; none did. Because of the potential for interviewees to become 'carried away' during the discussion, I considered it important to emphasise their right to withdraw, or to decline permission for me to use, any statement. The right to withdraw, in whole or in part, was therefore not only included in the letter of invitation (Appendix I), but was also repeated at least twice, once at the end of the interview and once in a subsequent email. I wrote 'thank you' emails to all interviewees a short while after their interview. The latter varied in detail because of my relationship with the individual and the content of the interview: a sample is included as Appendix V. It will be noticed that the right to withdraw is repeated a third time in this message.

The only significant ethical concern identified was the potential for reputational or psychological damage to interviewees. They are people of high reputation, several internationally, from prestigious institutions; it was conceivable that, in the context of a semi-structured interview, an interviewee could say something which, if quoted, could lead to reputational damage. The interviews also had the potential to visit, inadvertently, painful or uncomfortable memories for the interviewee. However, these are people of experience and wisdom, so the likelihood is low. The author is also a person of long experience in highly

sensitive settings, well able to spot potential risk and either avoid using that source or checking with the interviewee whether s/he is happy for words or ideas to be used. At no point in any interview did the interviewee express concern about either personal or reputational damage or discomfort.

Interviewees were never asked to comment on other institutions or individuals; comparative comments useful to my findings occurred twice, and I have identified one other occasion when a very valuable comment was made which could cause damage to people or institutions. On each occasion, I have protected the interviewee to a still greater degree by removing even the coded initials. Those data extracts are not attributed, and I have explained why in the text.

Extracts from interviews are attributed to fictitious initials (AA, BB etc); Appendix IV lists these initials with a one-line CV (“headteacher of schools both strong and in serious difficulties, subsequently CEO of a MAT”). These are carefully written to render the individual unrecognisable.

To help ensure the quality of interviews, their arrangements, paperwork, content and conduct, and to identify potential risk areas, I arranged three pilot interviews which provided feedback on all aspects of the process, as well as the opportunity for me to reflect and improve. One change made as a result was to strengthen the ‘right to withdraw’ statements as described above.

3.5 SUMMARY OF METHODOLOGY

- The study is *qualitative*, concerned with how improvement happens, its origins, impact and future, based heavily on evidence from the views and opinions of practitioners.
- By seeking to develop theory from the sampled evidence, the study is predominantly *inductive*.
- Because the researcher selected the samples of literature to study and practitioners to interview; and because the researcher sought patterns within the evidence collected, the study is *interpretivist*.

- My first research question sought to test a proposition or hypothesis and was, by contrast, inclined to a *deductive* and *positivist* approach, while remaining qualitative.
- The sources of evidence to answer the research questions were the existing literature, scholarly and grey, on the theme of school effectiveness and improvement; the scholarly literature concerning the social, economic, political and educational context of the research period; interviews with leading practitioners: researchers, government officials involved in the local and national governance of education, and school leaders.
- The author's long experience as creator, implementor, interpreter and evaluator of school improvement means that an *autoethnographic* component, in the form of validation and triangulation of other evidence, is a component of methodology.
- The accumulated data ran to many tens of thousands of words. Although the predominant method of analysis was *thematic*, because the identification of patterns of themes was the purpose of the analysis, the analysis methods were hybrid, because the precise words used by interviewees (discourse analysis), and the illustrative stories they tell (narrative analysis), lent not only detail but also authenticity and vividness to the analysis.

CHAPTER 4: FINDINGS

This chapter presents the five themes that emerge from the research data, with extracts from interviews substantiating each. ‘Themes’ are recurring facets in the data; ‘sub-themes’ are notable elements within the organising concept of the theme. Themes are presented as separate entities; links with the literature review and the research questions are stated but not elaborated. The Discussion chapter that follows interprets the themes in the light of the literature, comments on their possible significance, and forges links between them.

The structure of the chapter is displayed in Table 3.

4.1 THE BASIC PROPOSITION

The first question of this research was to test its underlying proposition: that effectiveness and improvement dominated thinking in English schools between about 1980 and 2010. The literature review supported the proposition by confirming that, during that period, three relevant developments took place:

- considerable and rapid expansion of research into school effectiveness and improvement
- the emergence within educational research of a specialised field of effectiveness and improvement, and a group of key authors
- a considerable expansion of government interest in educational improvement, leading to a large volume of fundamental legislation.

The first two developments provide the context and academic underpinning for change in schools; headteacher interviewees, when asked to identify the source of their understanding of improvement, rarely referred to academic research; their discourse was dominated by references to legislation and especially to Ofsted:

“although headteachers do not all have a clear knowledge of the principles of school effectiveness research, they all know the Ofsted framework” (CC).

This data extract stresses the importance of the third development: legislation, including the creation of Ofsted, added *force majeure* to the intellectual arguments. It is a defensible

Table 3: Structure of Findings: themes and sub-themes

	THEME	DESCRIPTION	SUB-THEMES
4.1	Testing the basic proposition	This research asserts that school effectiveness and improvement dominated thinking in English schools between about 1980 and 2010. This theme tests that proposition from evidence provided by interviewees.	4.1.1 Since around 2010, issues of individual pupil welfare have risen in importance to, arguably, become the new dominant.
4.2	Perceptions of 'what is to be improved?' have evolved and expanded	Analysis of interviewees' perceptions of what school 'improvement' involves show that the concept has both evolved and expanded.	4.2.1 Improvement goals include school actions as well as pupil outcomes. 4.2.2 Despite the expansion of improvement aims, test and exam results remain the 'default setting' for school improvement. 4.2.3 There are divergent views of the goals of improvement. 4.2.4 Some improvement is not real but the effect of strategies to 'game the system'.
4.3	Conditions for school improvement	In seeking to articulate how schools improve, interviewees refer to 'conditions for improvement': actions or attributes that promote improvement. This theme assembles these identifiable conditions that promote improvement, which are inchoate in the evidence.	The role in improvement played by: 4.3.1 Ofsted 4.3.2 self-evaluation and the balance between it and external evaluation 4.3.3 consistency of practice 4.3.4 differences in school improvement in 'constrained' and 'unconstrained' schools 4.3.5 the processes needed to get a school 'off the bottom' may not be the same as those needed to make a sound school, good
4.4	Fundamental school improvement has been neglected	While pursuing the 'technical' measures of school improvement, evidence of underlying improvement in schools has been neglected.	
4.5	Reasons for the rapid expansion of school improvement are not fully explained	Interviewees had patchy, limited ideas of the sources of school improvement.	

proposition that the existence of Ofsted, the inevitability of inspections with published reports, and their frequency in weaker schools, is itself a guarantee that quality and improvement remained high on headteachers' agendas. AA describes the 1990s as:

“a period when we were heavily focused on results, league tables and Ofsted as public measures of our success and accountability.” (AA).

To explore the interview evidence, I need to clarify the phrase “dominated thinking”. A very high proportion of schools' resources of time, energy and money is devoted to 'maintenance' functions: teaching pupils, supporting them and parents, managing resources, especially people, and to administration and maintaining the plant. An analogy is domestic non-discretionary expenditure, that which must be spent; although efficiencies may be made and decisions challenged, this set of activities will absorb a very high proportion of the total resource. There remains (hopefully) a small margin which constitutes discretionary spending, a matter of choice for those managing it. It is this margin of time, energy and money that I assert became heavily devoted to the effectiveness and improvement function.

Because the question “was your time as a headteacher dominated by effectiveness and improvement?” is clearly leading, it was not asked directly but explored through discussion of the role and its priorities. The result is that this section draws evidence from answers to other questions [Appendix III].

Five members of the sample (plus the author) became senior school leaders in the 1980s; direct experience of the early days of effectiveness is therefore limited in the sample. All have clear recollections of the excitement brought by the confirmation that schools have an effect, and that it could be quantified and analysed. The literature (for example Brighouse, 2021, p.288) gives support to the same idea. One of these interviewees referred to the stimulus provided by the work of Rutter (1979) and Mortimore (1988); s/he described:

“a clear list of what effective schools do and hence, implicitly, a steer for school leaders as to how to make their schools more effective” (NN).

JJ referred to “schools' confidence in their own agency” because of the new research. NN set about the improvement of a school in difficulties informed by the newly acquired principles

of school effectiveness, and I referred to the same ideas in interview for deputy headship in 1980.

Discussion with the 26 school and MAT leaders was dominated by themes of quality and improvement: that this was conditioned in part by the context of an interview themed around effectiveness and improvement is unavoidable, but opportunity was provided for interviewees to pursue other themes. They discussed resource and recruitment problems, especially in recent years; several contrasted this with the late 1990s and early 2000s: *“a golden age of priority and funding”* (CC); *“propitious times for education, a time of vision, priority and investment for education”* (LL).

But all those discussions of resource were set in the context of seeking improvements to teaching, the curriculum and leadership: *“the building blocks of school improvement”* (PP).

In discussing the changing targets of school improvement, an interviewee described the relentlessness of improvement goals:

“from simple academic outcomes, through the goals of Every Child Matters to concepts like value-added and gap closing” (GG).

This evolution of ideas is considered below [p.86 *et seq.*]; the relevance here is that the relentless challenge of changing goals ensured that issues of quality and improvement remained uppermost leaders’ minds. Another interviewee refers to:

“constant shifts in emphasis: from teaching to learning, management to leadership, raw results to progress and gap closing, a focus on outcomes to one on curriculum” (FF).

The publication of the work of Rutter and Edmonds drove an immediate expansion of research and publication on the themes from the early 80s, but the post-Ruskin-speech so-called Great Debate did not immediately result in great changes in schools. An interviewee described the 1970s as

“an era of choice for headteachers”, a time when the school could *“design a curriculum to suit [its] aims and needs”* (GG) contrasted with the late 80s and early 90s, *“the beginning of an age of prescription, inspection and accountability [when] legislative change related to school improvement came thick and fast”* (GG).

Another former headteacher claimed that: *“until the late 1970s, education was the sole property of educationalists”* (NN) and contrasted this with the accelerating pressure for improvement from beyond the school in the next decades. The same respondent contrasted the 1970s Schools Council approach to improvement, which s/he characterised as *“here’s an interesting idea, why not give it a try?”*, with the pace and compulsion of the late 1980s and 90s, characterised by *“a focused drive for school improvement”* (NN).

The consensus was that the late 1980s and early 90s were a ‘watershed’, a term used by half a dozen interviewees and implied by others. LL, who worked as a school leader, adviser for local and then national government, said the term *“‘watershed’ applied across all those educational environments”*, visible in terms of *“the dominance of the agenda by the accelerating drive for improvement”* (LL). AA described the 90s in particular as being:

“characterised by a flurry of change and increasing accountability through Ofsted, the National Curriculum, and league tables” (AA).

Asked to analyse the nature of the watershed, all interviewees pointed to the growing panoply of school accountabilities at the time, brought into being particularly by the major Education Acts of 1988 and 1992: everyone referred to the advent of universal inspection by Ofsted and the publication of inspection reports, most also to the annual sets of test and examination results, their publication in league tables, and the annual cycle of self-evaluation and re-drawing of improvement plans. As a primary headteacher interviewee noted:

“the annual publication of the school’s results was familiar to secondary schools but a transformation in primary” (BB)

S/he explained that primary schools had no equivalent of the school-leaving examination (GCSE) until the SATs tests introduced by the 1988 Education Act, nor any tradition of the accumulated results achieved by individual children becoming a measure of school performance. S/he argued that the attention of primary school headteachers became, for the first time, dominated by performance.

It is interesting to note that KK, analysing from a position at the heart of government policy making, was clear that:

“what attracted politicians to school improvement was the clarity and authority of school effectiveness research”

and that effectiveness and improvement came to dominate ministry thinking in education from the late-80s onwards.

Half a dozen interviewees noted that, In the 1990s, the term ‘accountability’ entered the educational lexicon, representing most graphically the pressure for improvement. An analysis of interview transcripts reveals that the words ‘relentless’ or ‘remorseless’ (to describe the pace of change) were used in two-thirds of interviews, and in three-quarters when synonymous phrases are included. GG, for example, describes *“the relentless pace of change”*. The language used reinforced this sense of breathless change, of urgency: several, such as NN, LL and PP, talked of *“driving improvement”*. LL, FF and GG also spoke of *“a sense of beating past performance, based on continuous self-evaluation”* (LL).

The adjective ‘continuous’ to describe change is also used in nearly half of interviews.

Little wonder then that improvement-related issues dominated school leaders’ thinking. Although many interviewees suggested that the pace of change had been excessive, two articulated the damage done:

“too many interventions with the assumption of rapid improvement: interventions that assumed too steep an improvement gradient” (JJ).

In the same vein, HH declared that *“there is a maximum speed that an engine can go”*.

My conclusion from analysis of interview data was that issues of quality and improvement were at the forefront of school leaders’ minds through the research period, and thus the research project’s basic proposition was justified.

Sub-theme 4.1.1: The rise of pupil welfare.

The evidence also made clear that the focus of leaders’ attention shifted, and this became a clear sub-theme: interviewees were asked, as a supplementary question, whether they still saw improvement as the predominant theme in schools; almost all said that it had been joined (some said supplanted) by issues of pupil (and, to a degree, community) welfare. Interviewee

KK described this as *“widening the boundaries of school leadership”* in the last decade, and especially post-pandemic. The composite list of new and developing foci, drawn from all respondents, comprised the following, with the three most-frequently-mentioned placed first:

- safeguarding and child protection which, at the prompting of government and Ofsted, are uppermost in schools’ minds: at every staff meeting in one school in which I was a governor, the headteacher began with the message that *“safeguarding is job #1”*
- issues of pupils’ (and families’) mental health
- expansion in the numbers of pupils with special educational needs and disabilities (SEND), and the severity of their needs
- gang culture
- drugs and county lines
- the implications of social media
- gender uncertainty and transition
- the expectation of frequent, detailed communication with parents, including feedback on their children’s progress, behaviour and welfare. As a young teacher, I wrote a one-line report on my pupils twice a year and met their parents once. Frequent text messages to parents are now the norm.

HH described *“the monster safeguarding”* and suggested that *“a compliance culture now dominates schools, making them risk-averse”*.

It is not surprising, given when the interviews took place, that the impact of the pandemic featured strongly in the list of issues headteachers considered dominated their time and attention, especially in the evidence provided by serving headteachers. They described not only setbacks to pupils’ progress, especially those from disadvantaged families, but also how Covid broke links between school and family, built painstakingly over years, again especially with disadvantaged families. KK described how, in his MAT, *“there has been a large increase in spending on various forms of pastoral support”*. VV referred to:

“the growing focus on family welfare during and after the pandemic, signposting families to food banks, how to access benefits, the school becoming a centre for mass vaccination and the giving of health advice” (VV).

The issues in the list above are complex, time-consuming and sensitive: the handling of some is potentially damaging for schools as recent events, for example concerning gender assignation of school lavatories and searching pupils for drugs, have shown. Many of the issues are 'new territory' for headteachers.

Some were emphatic that this widening of the aims of leadership for improvement is *"a temporary, post-pandemic phenomenon"* (BB). EE stated:

"Schools must be secured as places of learning, not therapeutic communities; the focus on mental health as a school or headteacher priority must be a temporary, post-pandemic priority, not a permanent accretion to the role" (EE).

Others believed *"these changes are here to stay"* (PP, VV). What is certain is that the present is a time of tension in what is meant by improvement.

Two of the most telling data extracts came from the head of an inner-city school: *"I can either work to raise standards in my school or try to deal with the welfare and health issues of the community, not both"* (BB). But to emphasise the significance of the pursuit of school quality, s/he went on to say: *"Ofsted frameworks and local authority briefings come and go, but this community lives or dies by what goes on in this school"* (BB).

In summary, the basic proposition of the thesis, that effectiveness and improvement dominated thinking in English schools between about 1980 and 2010, is upheld by evidence from the literature:

- considerable and rapid expansion of research into school effectiveness and improvement
- the emergence within educational research of a specialised field of effectiveness and improvement, and a group of key authors
- a considerable expansion of government interest in educational improvement, leading to a large volume of fundamental legislation

The view from literature is strongly supported by interviews, especially those with school leaders, who stated that issues of quality and improvement were at the forefront of their minds through the research period.

From around the year 2000 onwards, the emphasis on issues of pupil welfare and safety increased considerably.

4.2 PERCEPTIONS OF 'WHAT IS TO BE IMPROVED?' HAVE EVOLVED AND EXPANDED.

Interviewees were asked to reflect on their understanding of what constitutes improvement. The key facet of the evidence, recurring in almost every interview, is that the goals of school improvement have evolved and expanded significantly over 40 years. The following text develops this theme, identifying four noteworthy elements as sub-themes:

- 4.2.1 Improvement goals mainly concern pupil outcomes but also involve actions by the school.
- 4.2.2 Improvement embraces more than students' academic outcomes, but the latter form the default setting.
- 4.2.3 The evidence contains divergent views of improvement goals.
- 4.2.4 Some 'improvement' is false rather than meaningful and sustainable: such improvement is illusory, having been achieved through various forms of 'gaming the system'.

This section concludes with the paradox that improvement is simultaneously complex and straightforward.

There is a 'hidden' sub-theme which turns the practical dilemmas of the theme into a philosophical discussion: the difficulty associated with the term 'improvement' in the context of a lack of consensus regarding purpose(s) for education. *"The nation is unclear what it wants from its schools"* (SS); *"what is school for?"* (LL); *"what are we educating for?"* (UU). Although raised as an issue in around half of the interviews, summarised as: *"how are we to interpret improvement in the absence of agreed purposes for education?"* (FF), the evidence says little more. The literature review suggests that a consensus may be emerging; the sub-theme is explored in the Discussion chapter [p.119 et seq.].

Rutter's original 1979 work took an effective school to be one that promotes both the academic *and* personal development of its students, and almost every interviewee's first response to the question 'what is to be improved?' reinforced this. CC is typical:

"Improvement is not solely academic, although that is how it is most often measured not only are personal development, good behaviour, habits of work and manners valid educational objectives in their own right but also, unless personal characteristics improve, academic outcomes won't either; or will be limited." (CC)

TT was insistent that

“improvement is not either/or, it is both/and: it encompasses both exam results and the wealth of skills and values that enable students to take the next step in their lives, with confidence” (TT).

Several interviewees (especially GG, FF and AA) claimed that these wider objectives took a back seat in the 1990s and early 2000s, when test and examination results, and their publication in league tables, took priority. They further argued that the importance of personal development goals was not truly established until 2003. Many (GG, AA, HH, TT, FF) referred specifically to *ECM* (DCSF Green Paper, 2003) as a sea change or watershed, after which: *“outcomes for young people other than the academic were validated and ratified”* (GG); **the improvement agenda was permanently widened”* (LL). HH called it:

“a key marker in the development of school improvement ... the addition of a set of measures validating areas of school activity which schools themselves have always placed centrally”.

This Green Paper, prompted by the deaths of children at the hands of those who should have cared for them, became the Children Act 2004, codifying what were previously looser ideas of personal development into ‘the five ECM outcomes’: at least four interviewees could quote the five, twenty years on, that young people should:

“be healthy, stay safe, enjoy and achieve, make a positive contribution and achieve economic well-being” (FF, GG, UU, AA; see also DfES, 2004b, pp.17-18).

It will be noted that the ECM outcomes overlap personal development with academic success, as did several interviewees, who saw:

“the two aims as inextricably linked”; personal development “is a route to academic outcomes, through improving students’ motivation and commitment” (AA).

Most interviewees went on to discuss the shifts in academic goals, pointing to two problems inherent in raw results, first:

“Although I accept that attainment lies at the heart of what schools are about, the inherent unfairness, the far from level playing field, of school intakes ... and hence the shift from attainment to value added, to progress, to gap closing and equity” (PP).

I have already quoted GG on the same issue [p.79] to illustrate the breathless rate of change in these outcome measures; s/he echoes PP's analysis of the changing goals of improvement: *"from simple academic outcomes, through the goals of ECM to concepts like value-added and gap closing"* (GG).

Both emphasise the need to provide fairness for those who may have had low starting points, not reached benchmarks like five or more higher-grade GCSEs, but nonetheless made considerable progress. The measures were also designed to show a school effect, eliminating the influence of higher- or lower-achieving intakes, showing instead how much progress students made, individually and on average.

"While the headline measure of school attainment has remained the percentage of pupils achieving five or more good GCSEs, the headline measure of school progress has changed from 'value-added' (2002-2005) to 'contextual value-added' (2006-2010) to 'expected progress' (2011-2015) to 'progress 8' (2016-)" (Leckie and Goldstein, 2016, p.3).

These increasingly sophisticated measures were adopted by Ofsted as key parts of their suite of judgements of school quality.

The second problem with raw attainment as an improvement measure was the 'gaming' practised by some schools: this was mentioned so frequently as to become a sub-theme [p.90].

Almost all interviewees spoke about the theme of equity or 'gap closing' as an increasingly important attainment improvement goal, with a focus on the gap in achievement between the children of socially advantaged and disadvantaged families. None referred to gaps in attainment by gender or ethnicity, which became of government interest in the 1990s. The advantage gap first became a major focus in the middle of the first decade of the twenty-first century (Andrews, Robinson & Hutchinson, 2017) and gathered momentum after the establishment of the Pupil Premium Grant in 2011, which provided a definition of disadvantage. GG linked the focus on the underachievement of poorer families with the ECM agenda of 2003-4:

“ECM marks a profound change in the focus of school improvement: not just goals for personal development, but social justice for disadvantaged families The change from simple academic outcomes to a focus on progress and gaps.” (GG)

In an analysis of *“shifts within the pursuit of school improvement”*, FF includes *“latterly a shift from academic outcomes to progress and thence to gap closing.”*

Thus, the meaning of improvement began with pupils’ attainment and their personal development. The latter was crystallised by ECM legislation in 2004: the former ramified, first into various means of measuring progress, then into closing gaps in the attainment of different pupil groups. In the annual surveys of student opinion across the large MAT he chaired, KK reported that issues of equality and equity consistently came out on top: the issues of the age. More than one interviewee foresaw:

“a sharpening focus on the attainment of particular groups of students: large groups, such as those with English as an additional language or having special needs; and much smaller groups, such as those attending alternative provision or who are young carers.”
(KK)

Sub-theme 4.2.1: Improvement goals involve actions by the school, not only outcomes

Although interview evidence shows the core improvement goals were seen to concern pupil outcomes, all interviewees referred in some way to improvement involving the school’s actions. Some recalled that, when Ofsted came into being in 1992, their first Handbook (Ofsted 1992)

“distinguished between outcomes for pupils and ‘contributory factors’, such as leadership, teaching and the curriculum” (NN).

Although the term ‘contributory factors’ has been discarded by Ofsted, interviewees considered that the concept remains relevant: hence all, without exception, discussed improvements to aspects of the school’s work, notably leadership, teaching and curriculum, not just as key *means*, but also *goals*, of school improvement. Thus JJ (and NN) saw Ofsted’s

“ratification of the difference between outcomes and contributory factors as one of their key contributions to school improvement. Schools saw improvement in both as critical” (JJ).

Sub-theme 4.2.2: Academic attainment is the ‘default’ setting for school improvement

Given the variety of improvement goals in the foregoing, it is unsurprising that several interviewees described the idea of improvement as complex. However, a clear sub-theme is that the system reverts to academic attainment as a default. This reality enabled one interviewee to describe improvement as being:

“at the one time both elusive and obvious, straightforward and complex” (PP).

Set against elusiveness is one interviewee’s emphatic view that:

“although the ‘Every Child Matters’ aims underpin the school’s work, as a school, your job is to maximise examination results to optimise young people’s chances and opportunities. And this is especially true for disadvantaged students” (CC).

The idea that, despite all the complexity and expansion of goals, academic achievement is the default setting for school improvement is very widespread in the data. UU stated that:

“the government’s view remains that improvement is about exam results ... it remains very difficult to do well in an Ofsted inspection if the school’s academic results are not strong” (UU).

On the same theme, AA reflected on the school s/he led:

“an over-emphasis on personal development and students being happy, and a consequent under-emphasis on examination results, was the main reason why my school went into special measures” (AA).

Sub-theme 4.2.3: Divergent views of improvement goals

The preceding text summarises the interviewees’ majority views; there were in addition a small number of divergent views of the goals of improvement. The first, quite large group

(CC, EE, HH, FF, BB, PP and TT), significantly headteachers of schools in very deprived communities, considered that their improvement aims included families and communities. CC described how s/he had:

“set about helping families to take part, and interest, in their children’s education, by attending plays, events, parents’ evenings....” (CC).

thereby hoping to increase the students’ motivation and commitment. Others in this group echoed these sentiments and added in the effect of the pandemic which

“not only lost learning, especially for the most disadvantaged, but also widened gaps that had started to close, and fractured home-school relationships” (FF).

The next variant came from the headteachers of two University Technical Colleges (UTCs), a small group of schools with a focus on the STEM (science, technology, engineering and mathematics) subjects. The strapline on their website (UTC Website, 2023) reads “Today’s Education for Tomorrow’s Careers”, encapsulating their chief goal: employability. Their aim is to achieve *“the best possible career routes for our students”* (UU) and *“to have all our students leave to jobs with training, or to further or higher education, with no-one Not in Education, Employment or Training (NEET)”* (PP). Although these aims are familiar to other types of school, the focus in UTCs is stronger, and

“links with employers lead us to focus on improving skills for employability as seen by those employers, such as the ability to work independently and in teams, self-confidence and creativity” (UU).

Although many hinted at it, three interviewees (AA, FF, TT) explicitly stated that a rich diet of extra-curricular activity is a valid improvement objective. No doubt others would, if asked, have agreed with this aim, but only three saw it explicitly as one they pursued under the heading of school improvement.

Five interviewees consciously rejected school improvement in favour of *“doing the right things”* (FF, YY) or the *“pursuit of the good school”* (JJ). Another interviewee’s version of this idea is:

“I interest myself rather in excellence and a great education, arguably a more productive strand than school improvement, with which it often gets jumbled up” (RR).

These four interviewees are joined by at least six others in believing that the pursuit of improvement can lead to narrowing the curriculum and to the types of ‘gaming the system’ described below, including ‘win-lose’ approaches. The Discussion chapter [p.124] pursues the important distinction between excellence and improvement.

Sub-theme 4.2.4: ‘Gaming the system’

As noted above, a major problem with raw attainment as a school improvement measure was the ‘gaming’ practised by some schools: as RR notes:

“the superficial improvement in five or more higher-grade GCSEs achieved by multiple IT courses⁹ or by off-rolling¹⁰ may appear to represent school improvement but conflicts with the idea of a great education for all” (RR).

‘Gaming the system’ is a clear sub-theme in the research data: over half of the headteachers interviewed referred to various methods of gaming. UU believes that *“the focus on exam results was so total that it is hardly surprising that various forms of gaming took place”*.

Some methods were extreme: one interviewee (identity protected) had been:

“instructed by the headteacher to remove 12 students from the Year 11 roll before the official date¹¹. Parents of students who were likely to achieve badly and who were disruptive were asked to remove their children to home education or they would be permanently excluded.”

Another headteacher spoke with passion about the need for improvement to be system- or area-wide; neighbouring schools, knowing it had vacancies, excluded weak and badly-behaved students in the knowledge that they would be *“‘dumped’ in my school”*. S/he (identity protected) talked about *“win-lose improvement: local heads believe that they will win if my school loses”*. Some interviewees described the use of what they termed ‘bogus’ qualifications which at one point had GCSE equivalence.

⁹ GCSE Information Technology courses that provided four or five high-grade GCSEs through a portfolio approach.

¹⁰ Off-rolling’ is a means of ‘improving’ a school’s results by moving pupils likely to achieve poorly to another establishment.

¹¹ The date on which numbers of pupils become fixed, for the purpose of calculating examination percentages

Describing some approaches as ‘gaming’ is contestable: the multiple-GCSE, portfolio courses described above, and narrowing the curriculum for marginal students to help them achieve five higher-grade GCSEs including English and mathematics, described as *“trivialising approaches to improvement”* by FF, had the merit that they gained more students more qualifications, enabled them to go on to college courses and, arguably, raised their self-esteem and commitment to education.

Gaming shows that over-valuing examination success can distort the meaning of improvement.

Conclusion: The paradox of simultaneous complexity and straightforwardness.

PP argued this paradox:

“attainment is the heart of what school improvement is about, and schools pick up so many other things under the umbrella of school improvement” (PP).

S/he went on to elaborate, saying that its breadth (examination results, ECM, contributory factors, safeguarding, gap closing) and its shifting nature as successive administrations and Ofsted *Frameworks* moved goalposts, meant that s/he could describe improvement as *“a sham or falsehood ... every measure of school improvement is a proxy for real school improvement”* (PP). S/he concluded by claiming that *“The cynical definition of school improvement is whatever DfE and Ofsted say needs improving and that changes over time.”* (PP).

The same sentiment was expressed by UU: *“my understanding of school improvement comes directly from what we have to do for Ofsted”*.

In summary, although school leaders have always seen academic and personal development as equal and linked goals, test and examination results have generally been the ‘default setting’ for judging school quality and improvement. Measures of academic achievement have been refined to achieve greater fairness. Excessive concentration on narrow measures of academic success have led to various methods of ‘gaming the system’. The boundaries of school leadership have expanded to encompass the mental health of pupils and families.

THEME 4.3: CONDITIONS FOR SCHOOL IMPROVEMENT

Of the seven main interview questions, by far the fullest answers were to question 4: 'From your experience, how do you believe that improvement happens in schools?'. The volume of transcribed answers to this question exceeded that of the combined total of the other six. Whereas the literature abounds with references to the 'what' of school improvement, it is less rich concerning the 'how'. The literature often suggests 'brute sanity': here are the school effectiveness criteria; do more of those things. All interviewees were profoundly interested in how improvement happens.

*"School improvement works by good **leaders** applying the **principles** of school effectiveness (KK)." "School improvement is about having **systems** built around the right **people**" (DD) (my added emphasis).*

These two significant extracts structure this chapter. Analysis of responses, especially to question 4, shows constant allusion to there being **conditions for improvement**: features of schools conducive to improvement. Four interviewees (BB, LL, NN, SS) used the specific phrase, all of them in the sense that the conditions are created, not spontaneous. All interviewees espoused the idea that some features of the school are propitious for improvement and others unfavourable. Rather than listing the researched effectiveness criteria, the 'what' of improvement, interviewees sought to get behind the generalisations of those lists, delving into the 'how' of unspecific phrases like "purposeful leadership" and "consistency among teachers" (Mortimore *et al.*, 1988, pp.250-251).

The data extracts above suggest four groups of conditions: leadership, principles, systems and people; these form this section's sub-headings.

Sub-theme 4.3.1: Leadership

Interviewees argued that, since the conditions for improvement are created, the key to how improvement happens is the leadership that shapes them, so interviewees' analyses of the attributes of effective school leaders dominated the discourse. They drew on their experience of leadership, personal and through their careers, to outline what they had learned about the management of improvement: what each role and each leader they had worked under had added, positively and negatively, to their understanding, frequently naming individuals; and exactly which piece of learning each had contributed. LL and EE, for example, both named

the same headteacher who had *“brought about striking levels of improvement in the school by rigorous use of Total Quality Management”* (EE). Although there were many references to what those leaders did, much more of the discourse analysed how they went about their work. Tim Brighouse summed this up in conversation: I was briefing him about the headteacher conference to which I had engaged him to speak; he asked: “how many headteachers do you have in your LA who know the words but don’t know the xxxxxxxx tune?” (Brighouse, 1993): ‘tune’ equals ‘how’.

Almost all the headteachers in the sample had, at some point in their career, taken on a school in serious difficulties and turned it round. A trope in interviews was ‘what I found when I took the school over’: *“I didn’t realise the depth of the difficulties it was in”* (CC). With passion, they described their struggles and reflected on the keys to success.

AA reported having worked for two weak headteachers, in whose schools *“there was no sense of corporate effort towards something called improvement”*. Then s/he joined a well-led school, with an improvement plan

“which I had never heard of before suddenly there was a sense of purpose, collective effort, something around which we could collaborate I was made responsible for delivering part of the plan and held accountable for it” (AA).

This sense of individual and collective purpose is echoed throughout the evidence.

TT was newly appointed to a first senior leadership post:

“on my first GCSE results day, the leadership team turned up to get the results. They were poor; we didn’t know why; we didn’t know what results the students should have achieved; what was most alarming, we didn’t know what to do about it.”

From the role of leader of a MAT, PP was clearest about the centrality of leadership:

“all the schools in this Trust were in difficulties, some of them severe, until sound leadership arrived in each despite all the words about it being a team etc, it is impossible to over-estimate the impact of a single, high-quality leader at school or department level” (PP).

From the data, the four attributes central to effective leadership are:

- facing up to the need for change
- setting (often transforming) the culture and values of the school
- focusing on the right things
- building teams (overlaps with section 4.3.4, People).

Facing up to the need for change

“The school will make no progress until it faces up to the realities of its position” (EF).

“There would be no change until we faced up to our problems” (FF).

“The critical decision is to know and accept where the school is: to face up to the facts, sometimes brutal, before it can move on” (RR).

TT's experience is common to those taking over schools in serious difficulties: his/her willingness to tackle poor performance encountered challenge from a group of staff who reacted to visits to classrooms by taking out a collective grievance. All radical change programmes described to me involved overcoming resistance and challenge; hence determination is a pre-requisite.

KK's view that *“To change something, you have to understand it”* introduces both a note of caution and the concept of evaluation, pursued below.

Setting (often transforming) the culture and values of the school

When XX took over a school in long-term difficulties, s/he found a staff whose attitude could be summed up *“Tell me what to do and I'll do it”* (XX), disempowered and demoralised by the contrary directions set by different headteachers and advisers, *“if told to stand on their heads, they would have done so”*. S/he described a group of people who *“turned up to work with little sense of engagement or agency”*. The key to turning the school round was the building of trust, arguing that

“Unless there is trust in the leadership, staff will not take the risk to innovate ... creativity and innovation require a culture that encourages and enables risk-taking. That in turn requires trust and confidence” (XX).

S/he set about building trust and confidence through a culture of openness and collegiality: budget issues were explained and staff reassured; expectations of teachers were made clear, so that lesson observations lost their threat.

Two of the commonest words used to describe the rebuilding process were “culture” and “expectations.” Almost all spoke of the need to rebuild the school culture. BB was particularly emphatic that this is the key role in bringing about improvement: s/he described taking over a “*busy school*” but one without accountability to a clear vision of what it believed in, and without a clear set of expectations. Reflecting on both schools s/he has led, “*there is a temptation to see school improvement as a list of jobs, when what really matters is the creation of a culture*” (BB).

TT was equally clear:

“there is no change without cultural change, winning staff over to it or removing them; ... creating critical mass buy-in ... Those who share the vision are drawn into the centre; those who can’t are trained; those who won’t are removed” (TT).

CC’s model of improvement echoes this sentiment:

**[leaders] set standards and expectations in discussion with staff; hold everyone to account to those standards; improve through professional development those who aren’t but can; shed those who won’t*” (CC).

Pressed on what they meant by ‘culture’, interviewees referred frequently to consistency, a sub-theme (Systems) considered below. YY said it is “*vital that everyone signs up to the school’s core values*”. BB described culture as “*what it means to work here ... a culture is something people buy into, creating professional loyalty*”. S/he talked a lot about “*re-booting*” and “*this school is a place where xxx happens*” and goes on to propose that “*a well-articulated culture provides a yardstick by which we hold each other to account*” (BB). At this point, my marginal note on the transcript reads: “NB each other, not just top down”. “*Whether or not they call it vision or culture, everyone, including parents, needs to be able to articulate what the school is about*” (BB).

YY leads a school with remarkable levels of consistency and continuity; the large senior leadership team (SLT) is entirely home-grown and has stayed together for nearly a decade. S/he speaks of *“home-grown staff: we grow their good habits”* (YY).

Different individuals in the sample used different words to describe ‘culture’: thus *“schools only prosper when they have an agreed ideology, a philosophy, a belief system”* (FF). S/he inherited a school in a poor state: *“in crumbling buildings, in deficit and in the lowest bracket of achievement ...I went for a cultural shift: let’s establish who we are, what we’re for”*. This cultural shift *“established clear values and priorities, written in kid-speak, displayed and constantly referred to, to show what we are about”* (FF).

Pressed still further on how the renewed culture would become tangible, many including FF, LL, EE, NN, BB mentioned *“everyone having high expectations of each other, staff and students alike”* (NN). A large number referred to language: **the courtesy with which staff and students were expected to speak to and of each other”* (FF), and several to inclusivity, which could be seen as a corollary of consistency: *“including everybody, ensuring that everyone matters and is valued”* (LL). Others referred to **ambition, for the whole school community”* and to *“the focus is on the pupils, not the job”* (YY).

EE fixed on *“a culture of openness and the absolute need to protect it”*, which BB expressed as *“improvement won’t happen until we are challenging each other, openly and without blame”* (BB).

LL’s version is to *“sell the story to all staff, get them emotionally involved and so be able to change their behaviour”*, emphasising the centrality of *“continuous professional development alongside change, to effect improvement”* (LL). HH’s notion of culture centred on teachers’ *“lack of agency”*: in schools s/he took over:

“we gave them a belief in the possibility of improvement: things don’t have to be this way. ... Improvement is intimately linked with institutional confidence” (HH).

CC describes *“a sense of shared purpose”*.

Focusing on the right things

“My leadership style focused on doing the right things rather than obsessing about outcomes: do the right things and outcomes will follow” (FF).

AA names *“the ability to set and pursue clear priorities within the core business of education”* as a key attribute of headship, while LL describes the process of leading school improvement as *“vision first, build the team to it and then focus on teaching”*. TT agreed with this priority order, with *“leading the improvement of teaching”* at the centre. LL focused on *“heads’ strategic ability to plan”*. With a dazzling mix of metaphors I can only admire, an interviewee summarised focusing on the right things as: *“the ability to filter the noise, identify the gold in every innovation but not jump on each bandwagon; and retain the baby while ditching the bathwater”*.

FF and HH link this ability to focus on key priorities with skill in deploying resources, especially money. DD has been a head in three MATs and said that one, with a national reputation for consistency that some interpret as rigidity:

“spends money and time on the right things; focuses the attention of leaders on teaching, curriculum and staff development. Yes, it is competitive but the organisation learns and grows by studying performance, comparing schools within it to see what works ... the organisation learns” (DD).

Identifying ‘the right things’ implies the skills of self-evaluation; the term occurred rarely in interviews, with only few mentions such as *“the skills to self-evaluate at the heart of leadership skills”* (AA). I expected interviewees to extol the virtues of self-evaluation over scrutiny by Ofsted and other external bodies, but they rarely did so. However, when I looked behind and between their words, I realised that this was because they treated self-evaluation as integral to leadership for improvement and the self-managing, self-improving school. Much of what they said was predicated on routine self-evaluation: *“knowing where you are is a necessary condition of improvement”* (AA); *“what school is going to wait for Ofsted to judge it without judging itself?”* (NN). Instances of allusion to self-evaluation without naming it include:

- the need to monitor against the standards set: LL described *“learning from what has and has not worked in the school, and reviewing progress against the standards set”*; CC referred specifically to *“observing lessons and scrutinising data to challenge consistency”*, EE to *“systematic lesson observation”* as one way to *“ensure fidelity of implementation”*

- FF's *"unwavering focus on lesson quality"* implies scrutiny. BC and CD described the 1000 lesson observations which took place in a school year to *"ensure that expectations were met, every time"*
- KK: *"to change something, you have to understand it"*, describing how the MAT s/he chaired *"used to employ consultants to tell heads about the quality of their schools"*, and how they devised a process which *"internalised that knowledge by seconding members of its own staff, not only to understand the strengths and weaknesses of individuals and teams, but also to support and work alongside them to bring about improvement"*
- TT described shock at a new school's failure to understand a poor set of results, to know what those students should have achieved, and what to do about it
- like LL, PP described the crucial importance of headteachers' ability to understand not just what worked, but why, a skill I have termed 'the ability to deconstruct success' and expand in the Discussion chapter [p.143].

Interviewees thereby asserted the centrality of self-evaluation to improvement. Ofsted reinforced this in its inspections, for example by co-observing lessons with senior staff to check the accuracy of the school's self-evaluation. From 2006 to 2010, it was a statutory requirement for schools to produce a Self-Evaluation Form; although that is no longer the case, Ofsted still includes "a summary of any school self-evaluation and/or improvement plan or equivalent" (Ofsted Handbook, 2024, para.112) among the papers it asks to see during an inspection.

Building teams

One of the strongest messages from the evidence was that **school improvement is a team game**. EE described the building of strong teams of senior and middle leaders as a key role of headship, and crucial to the pursuit of improvement *"since teamwork is the basis of consistency"*. PP described how *"headteachers build teams, coach them and monitor performance"*, matching DD's description of how s/he went about turning a profoundly weak school round: *"recruit good leaders, give them responsibility for parts of the improvement plan, and then coach and monitor how they execute the roles"* (DD). YY involves staff in decision-making by distributing leadership across large senior and extended leadership teams.

In addition to poor performance and unwillingness to be held accountable, TT's criteria for shedding staff included *"those unwilling to work in a team"*.

UU and RR, in very similar words, believed that *"improvement was impossible during the era when the teacher's classroom was her own domain"* (RR).

{I described earlier the unwillingness of the head of mathematics in my first school to allow the headteacher to enter his classroom: he believed in the supremacy of the teacher in his domain and that for another adult to enter was an impertinence. I never saw my first headteacher about the school corridors; he certainly never entered my classroom.}

School improvement is impossible in an individual culture: it must be a 'team game'.

Sub-theme 4.3.2: Principles

These comprise chiefly the school effectiveness criteria, rehearsed in the literature review and treated only briefly here. These principles of improvement lie at the interface between leadership and systems; PP believed that:

"somewhere along the spectrum from 'one size fits all' to total context specificity (at which point everything is individual to the school) is a set of common-sense principles" (PP).

Most of the principles of improvement referred to during interviews were references to the effectiveness criteria in the research literature. Hence KK *"School improvement works by good leaders applying the principles of school effectiveness"*, with specific reference to the work of Mortimore and Reynolds. Reflecting those criteria, the centrality of good teaching to improvement was frequently mentioned, such as TT's determination to observe classrooms, FF's commitment *"never to waver in the focus on classrooms"* and EE's conviction that *"school improvement is about pedagogy and practice"*.

Other principles have been referred to under Leadership, above, such as a belief in teamwork and in inclusivity, valuing everyone: the earlier reference was not just to valuing all staff, but also all students. Many interviewees (including JJ, NN, LL, FF, RR) referred to the far greater attention paid to the needs of students of more limited ability, including those with special needs, and how in the past such students had often been under-valued.

Sub-theme 4.3.3: Systems

Consistency

A system implies some degree of consistency; since consistency doesn't occur spontaneously, it must be led or, to use AA's term, "*driven*"; having clear systems and insisting on consistent implementation are widely described in the evidence as major components of leaders' techniques to drive improvement. If improvement is a team game, games have rules and team members should play by them.

Every interviewee referred to consistency, though not necessarily with the same meaning. I have written elsewhere (unpublished doctoral assignments) that the limits of consistency (what it is worth being consistent about) are rarely defined, leaving open the possibility of caricature and allegations of straitjacketing.

Frequent reference will be made in this research to Collins's (2001) flywheel image: in a book about why some companies become 'great' while others do not, he envisages an organisation as a heavy flywheel; it takes massive effort to begin turning it but, when spun in the same direction (consistency) over a sustained period, it spins rapidly with little input of energy, highlighting the importance of multiple small actions rather than a few large ones. I have always found this image compelling, and RR referred to it to illustrate the importance of momentum and consistency.

Interviewees such as LL and EE described how they had learned from distinguished headteachers who designed consistency into their schools: one who determined the design and contents of "the standard classroom atSchool"; another, who developed "the good lesson at School "; and another who "*re-engineered an entire school based on the principles of "Total Quality Management"*" (EE). S/he acknowledged that the crucial element had not been the specific elements of TQM, but "*fidelity of implementation*" (EE), a phrase much used in American improvement programmes and in interviews, including with DD, who stated that "*what matters is fidelity of implementation [of systems], rather than their exact content*". This echoes other interviewees' views, summarisable as "it ain't what you do but the way that you do it".

Earlier research mainly concerned consistency of actions or behaviours, an interpretation used by several interviewees: classroom routines such as taking a register and managing behaviour. CC referred to "*routines*", EE to "*non-negotiables or SOPs*" (standard operating

procedures). However, most discussion of consistency concerned school culture, discussed under 'Leadership'. This argument featured strongly among the many interviewees who had taken over weak schools, some failing dramatically, others *"bumbling along"* (HH). FF discussed how the school developed *"a common language, written in kid-speak, to describe who we are and what we will do ... and the need to develop consistency of values, not just actions."* HH: *"the first thing to do [when taking on a weak school] is change the culture"*. NN spoke of *"the joy of building a culture around which we built staff loyalty to do the right things, consistently"*.

There was a lot of discussion of where consistency becomes uniformity, with the common assumption that consistency is desirable, uniformity is not. When pressed as to the hinge-point between them, the comments almost always had a pejorative slant: AA described one MAT's rigidity, as he saw it, as *"militaristic"*; EE asked, *"How like Starbucks should we be?"* and fears *"a robotic approach"*. GG referred to some MATs having *"improvement in a box"* or an *"engine out, engine in"* approach: when a component of one school in the MAT fails, the same component is lifted from another school to replace it.

GG went on to say: *"why not clone practice if it works?"* and this was the commonly held, pragmatic view. HH: *"it is easy to caricature MAT consistency as 'Big Brother' but, well run, such systems work"*. An insider's perspective from DD described a MAT, renowned for its consistent practice, as having *"built strong systems to save time and bring about improved outcomes, not consistency for consistency's sake"*. YY spoke at length about designing and refining systems for managing behaviour and for teaching a good lesson *"to save teachers' time and effort, enabling them to teach at their very best, every lesson"* (YY).

Although several appeared to hold their noses when saying so, many more interviewees espoused a substantial measure of consistency, even uniformity, than expressed doubts. SS asked, *"Is there a formula for school improvement and, if so, what does it look like?"*. BB provided an answer, saying *"I do not believe there is a 'formula' beyond the building and sustaining of a distinctive culture"*, thereby providing a formula!

Two leaders in particular (RR, BB), perhaps coincidentally the youngest interviewees, were *"in favour of more prescription, not less"* (BB). They lead very different schools, but both argue the need for greater uniformity, reasoning that the demand for personalisation of learning,

with a rising proportion of pupils with additional needs, means that teachers' attention should focus on tailoring the material to the needs of individuals, not to developing that material.

The Discussion chapter [p.146] proposes a detailed definition of consistency.

Systems in particular settings: improvement in constrained and unconstrained schools

The terms 'constrained' and 'unconstrained' were used in these interviews by SS, and I can find no reference in the literature to their use. S/he was presenting a view that *"schools are victims or beneficiaries of the houses that surround them"*.

Sensu lato, all schools are constrained by budgets and recruitment difficulties, but SS's meaning was narrower. Almost all senior leaders interviewed had led schools in difficulties almost all of them SS's 'constrained schools'. Assembling their views, there emerges a set of schools that serve communities of considerable, diverse disadvantage, with high levels of poverty, social problems and/or homes where English is not the first language. Because academic achievement tends to be low and behaviour problems substantial, such schools are often unpopular and therefore have spare places and become homes for migrant and refugee families as well as for pupils excluded from other schools. In need of help and support those children and families undoubtedly are, but they rarely enhance the school's academic performance, so the spiral continues, and the difficulty of improvement grows. Several drew attention to the considerable difficulty of recruiting staff to such schools, and of retaining them: promising teachers are often 'poached' by more successful schools with larger budgets (CC, DD, HH, SS, RR, LL, EE FF, PP).

The significance is the greater difficulty of school improvement in constrained schools. SS argues that, in unconstrained schools, *"sound leadership that applies the principles of school effectiveness will lead to improvement"* but that this is not the case in constrained schools. DD described the school as *"set in conditions inimical to school improvement"*. CC spoke passionately about *"the depth of the challenge to improvement the huge scale of deprivation the degree to which staff felt defeated"*, describing the city which contained that school as *"twin cities, one of affluence, the other of poverty"* with the route to improvement much more straightforward in the first than the second. LL described the same

city as *"having inequality baked in"*. Interviewees characterised the problem variously as the greater distance to travel, the greater height and gradient of the climb, and the friction impeding movement (CC, HH, EE, FF).

Pressed on how improvement methods differed between the two settings, interviewees did not point to different actions, systems or processes: they spoke instead about the need for stronger leadership (SS, TT), for greater determination (LL, EE) and for greater levels of prescription for longer periods (TT, UU).

Systems in particular settings: systems needed to get a school off the bottom differ from those needed to make a sound school, good

Approximately half of interviewees raised this issue, the two most widely held views being that *"there are well-tried methods to get schools off the bottom"* (RR, also HH), and *"it is clear that moving a school from inadequate to sound, and thence to good and beyond, requires both more of the same and some new elements"*. (LL and others).

RR: *"it is not difficult to know what to do when a school effectively needs to start again"*. Several respondents referred to the need, in such circumstances, to *"manage behaviour provide staff with clear prescription and remove profoundly weak and blocking staff"* (RR), a sentiment repeated almost verbatim by HH, and in slightly different terms by CC, whose turn-round model comprised setting standards of practice for all staff to adhere to, *"monitoring staff against those standards, training those who can't and removing those who won't"*. The headteachers in the sample with experience of turning around several schools all had their own clear models, different in detail but with similar elements: using SOPs to manage teaching and student behaviour, setting clear standards, monitoring against them and so on.

It was a view commonly held within the sample that the adoption of a regime of 'non-negotiables' (a set of actions and behaviours required by leaders of all staff, what AA referred to as *"a militaristic regime"*) is temporary, at the early stage of the improvement process, in his words *"to get the school out of trouble"* (AA). In Collins's 'flywheel' model, this early stage is analogous to the hard labour necessary to get the wheel turning.

Eddie Jones (international rugby coach) was asked in a podcast interview whether he subscribed to a democratic style of leadership or one which is more dictatorial:

“When you inherit a team ... you want to be at the front of the room, guiding, directing, cajoling, setting the standards. By [the end], you want to be at the back of the room admiring what you’ve done.... There’s a progression from a smiling dictator to a fantastic democracy. That’s the evolution of the team that you want.” (Jones, 2023)

Interviewees described the same evolution: FF *“in the early stages of turn-round you insist on staff doing the same things; later in the process, you empower successful team leaders to adapt procedures to their own needs”*. In one school, weak or failing teams were placed

“not in special measures but in intensive care: leadership was transferred to a member of SLT, we insisted on particular SOPs in the team and maintained a monitoring regime to ensure these were carried out and that the team was improving. Then the intensive care was gradually removed” (VV).

LL, CC, RR and others described this process as ‘earning autonomy’: the power granted to effective teams/leaders to express individuality in curricula or teaching methods; basic procedures remained inviolate, and autonomy was subject to continuing success. Chairs and CEOs of MATs described an identical process with schools in their chains, those capable of self-management and -improvement being granted freedoms denied to the less successful.

There was a widespread belief in the sample that *“a combination of Ofsted and following the principles of school effectiveness has resulted in the worst schools being driven out and the baseline raised”* (FF). JJ agreed: *“Ofsted and school effectiveness ... have closed disgraceful schools down”*.

PP described the different skills needed by leaders at different stages of the improvement process: one school in his MAT had a headteacher *“who played a key role early in the process by winning the community to the school”*; when the need for improvement in teaching and curricula became obvious, that headteachers did not have the necessary skills and was replaced.

I reached the conclusion that different stages in the trajectory of improvement did not in fact have very different procedures; that the early stages were marked, in Eddie Jones’s words, by

a level of dictatorship, smiling or not; and that increasing democracy shifted the role of leadership to increased sharing of ideas and coaching of individuals, but the basic principles on which the school worked remain unchanged, that Collins's flywheel is a reality. YY made the point strongly that managing behaviour in his/her school with a tough intake was constantly top of the list of priorities: some improvement models suggest that, having got on top of poor behaviour at an early stage, the focus moves on to curriculum and developing leadership. YY takes the view, and others including WF, TT, FF agree, that *"behaviour has to be better than good, because these students need every possible moment to learn"* (YY).

The marked exception to the process of evolving procedures was BB. RR's view is that *"school improvement is predicated on a failing school in a deprived area, not on mediocrity in the unambitious suburbs, or a coasting school"*. Moving to his second headship, BB found exactly the latter situation, concluding that *"it would have been easy, and tempting, to see improvement as a set of actions, things to be done by everyone"*. S/he took the longer, harder route to *"sustainable, improvement, rather than short-term fixes"* by changing the culture, winning staff over to new ways of talking and thinking about what the school was for, its key values S/he worked on getting leaders *"out of 'fix-it' mode and into a deeper, more enduring understanding of change and the role of the school in its community"* (BB).

Sub-theme 4.3.4: People

To paraphrase Collins (2001), having got the driver (leader), get the right people on the bus, sitting in the right seats. DD gave particular emphasis within team building to recruitment of staff, especially leaders, describing the care s/he takes to find good people, not relying on advertisement and interview but seeking people out, visiting their schools, persuading them of the valuable role they will play and the rich experience they will gain. Because of a different school history, YY builds teams from within, recruiting new teachers, developing their skills and *"good habits"* and teamwork through systems which instil consistent practice.

Two final points in this general section on conditions for improvement. RR referenced the relevance of *momentum* by recalling the Collins flywheel. The beginning of the process of improvement in his/her school took enormous effort, like beginning to turn the heavy wheel. When the process has been going for some time, as long as the touches on the flywheel are

in the same direction, it spins rapidly with a light touch. *“No single action can necessarily be identified as vital, rather the cumulative effect”* (RR) of thousands of appropriate actions.

KK said that *“to change something, you have to understand it”*; PP developed the same idea: *“of the many attributes of good leadership, one key one is the ability to understand success and how it happened”*. I have termed this skill ‘the ability to deconstruct success’ and expand on it in the Discussion chapter {p.146}.

Sub-theme 4.3.5: External support

One aspect of managing improvement threaded through the data is the role of external support. AA describes one headteacher skill as *“organising external support and balancing it with within-school action”*. Support is categorised here under three headings: local, national government and Ofsted.

Local support

References to LA support were made mostly by those who had been LA officers (GG, LL, NN): they described LA-orchestrated project groups to support schools in trouble (NN, GG), including partnering weak schools with stronger neighbours (LL).

There were many more comments on the support role of MATs. SS asserted that *“MATs are capable of building the conditions for improvement in ways generally inaccessible to LAs”*; HH elaborated this, saying that *“MAT leaders know their schools better than LAs did, and have the power to make resource allocation decisions”*, an opinion shared in almost identical words by RR. Since the attributes of effective leadership apply equally to MAT leadership, I will not elaborate.

Many interviewees were keen to highlight the value of partnership between schools as means of helping constrained schools to improve: pairing of schools, such as in early federations or the London Challenge (LL, GG) or in MATs (HH, EE, PP, FF, RR). In support of partnership, EE saw *“the improvement journey as one from dependence (on the LA), through independence (grant-maintained or stand-alone academy) to inter-dependence (MAT)”*.

National government support

This was described extensively in the literature review. Interviewees, while commenting at length on the stress and turmoil of the era of rapid change between the late 1980s and about 2010, nonetheless acknowledged that the changes did bring about improvement. Several said that it would be hard to envisage a world without a National Curriculum; where resourcing, including staffing, decisions were made by the LA; or where there were rigid school catchment areas: they were emphatic that these were improvements, not just changes. SS referenced the significance of government in *“training leaders, greatly strengthened since the development of the National Qualification framework”* (SS).

There were a few references to the influence of Secretaries of State: *“those who were committed long-term to education”* (HH) *“big beasts, political heavyweights who used longevity in the role to convey a vision for education and bring investment”* (LL). LL cites particularly the combination of Blunkett and Balls, who together held the role for over half of Labour’s time in power from 1997 to 2010, and who provided vision and priority, as well as resources, to education. Others referred to Joseph and Baker, equally influential in the 1980s.

Support through Ofsted

“Only Ofsted really improved schools as a whole system, through a combination of defining quality and strong-arm tactics” (SS).

Although there were comments about the stresses caused by Ofsted inspection, such as: *“so much hangs on an Ofsted judgement: head or governors lose their jobs, the school can close or be transferred into a MAT”* (UU), many interviewees acknowledged Ofsted’s role in school improvement: *“a major source of evidence and a motive force for improvement”* (NN).

JJ listed what he saw as Ofsted’s main contributions:

“Providing a clear definition of quality; distinguishing between outcomes and contributory factors and linking the one to the other: inputs linked to outputs; clearing out grossly inadequate schools; re-inspecting systematically to track change”.

Others supported these views: many, including HH, RR and NN, referred especially to the original version of the Framework for Inspection (Ofsted, 1992), the so-called ‘buff folder’, as:

“a great contribution to school improvement; the legal framework; the definition of a ‘good’ school; the headings of quality and how they fit together: outcomes and contributory factors, and the significance of the different contributory factors”. (NN)

“Its definition of quality and its demonstration of how variables within a school can be organised, and their inter-relationships set out.” (HH)

RR added:

“not only do I mean the definition of quality through spelling out standards, but also what is good enough, better and not good enough an Ofsted ‘good’ defines the basic minimum acceptable standard that should be attainable by all, while ‘excellence’ means doing everything consistently well” (RR).

A few, including NN, LL and especially JJ, deplored the shift over time away from

“inspection as an extended professional dialogue to one-word judgements with the focus on information to parents: hence the use of lay language and very short reports: ‘what is it like to attend this school?’” (JJ).

Although several commented positively on the *Framework* revision in 2019, increasing the focus on curriculum in particular, JJ and NN were the only interviewees to comment on the profound, and largely unremarked, change in 2005-6, a concern shared by the author, and which features in the Discussion chapter [p.161 *et seq.*]. For reasons detailed in that section, the 2005-6 changes to Ofsted’s work abruptly changed inspection from a professional dialogue to a judgmental exercise with parents as the audience, a profound change largely unremarked in the literature.

A few interviewees (AA, BB, PP) commented on the specific effects of inspection on improvement in their schools: PP and AA spoke positively: *“it made us look hard at ourselves in the mirror and face up to things we didn’t do well”* (AA). BB, who was head of a school in which the judgement of a deeply flawed Ofsted inspection was over-turned, spoke in a more equivocal way: *“Having seen the worst that Ofsted can do, I am now clear about what and how I have to do to lead a school well”* (BB).

In summary, the section concludes that there are conditions for improvement, features of schools which are conducive. Since those conditions do not occur spontaneously, the focus

is on leadership: using data extracts from interviews, mainly with school leaders, the text analyses the attributes of leadership which has been successful in driving improvement. These attributes focus particularly on the building of a culture of high expectation and, because improvement is a whole-school, team enterprise, on building in consistency. The text considers improvement in schools in different contexts: those in more and less challenging settings, and those at different points on the trajectory of improvement. The significance of various forms of external support to the school is considered.

THEME 4.4: FUNDAMENTAL IMPROVEMENT HAS BEEN NEGLECTED

“Every measure of school improvement is a proxy for real school improvement” (PP). When pressed on what he meant by ‘real improvement’, PP talked of schools becoming better places for young people to grow in, better preparation for life after school, a view shared by TT, who held that it is the job of schools to *“enable students to take the next steps in their lives with confidence”*.

When conducting the literature review, I found little evidence of what I shall refer to as ‘fundamental improvement’: changes to the nature of schools and schooling that make them more humane places in which to live and learn. I contend that the study of fundamental improvement provides good evidence for improvement, evidence which is not as equivocal as that from more ‘technical’ measures [p.155 *et seq.*]. Chapter 1 provided personal, anecdotal versions of the argument; here, I outline the findings from interviews on this theme, while the Discussion chapter [p.155 *et seq.*] provides an analysis.

The surprising conclusion from the interviews is how little comment underlying improvement received. It could be argued that the way the discussion was introduced led interviewees into the technicalities of improvement: however, the opening question asked what the interviewee took school improvement to be about, an open-ended question. The penultimate question asked for an assessment of how successful school improvement had been, and each discussion closed with an invitation to add anything that the interviewee wished. I therefore believe that interviewees had opportunity to raise other matters if they wished. In the event, the number and variety of comments were small:

- a group, including FF, EE, HH, JJ, NN and RR, referred to the driving of the very worst schools out of the system, mainly by Ofsted inspection
- CC, BB, LL, NN referred to better communication with families and communities, and their greater involvement in the education of their children. BB spoke powerfully about the centrality of the school to its community, which *“lives or dies by what happens in this school”*
- NN and FF both mentioned *“self-evident improvements”* in *“the abolition of corporal punishment and the way less-able students are valued”* (FF), *“the way students are treated in school, in the focus on safeguarding and in the quality of teaching”* (NN). TT and UU contrasted *“the total lack of inspiration in the way I was taught at school”* (UU) with the lively, engaging practice s/he sees daily in school
- AA was the only interviewee to mention governance, *“its increased professionalism, greater presence in school and enhanced links with the community”*
- FF and LL both spoke about *“greater staff involvement in decision-making”* (FF) and *“more professional dialogue”* (LL)
- some (especially LL) acknowledged that performance gaps, such as those to the detriment of disadvantaged students, have been recognised latterly, with significant attempts to narrow them
- only LL cast doubt on some aspects of basic ‘improvement’, arguing that *“improvement in outcomes such as attendance and exam results has to some extent been at the expense of narrowing the curriculum and the richness of extra-curricular activities ... loss of freedom within the curriculum”*. S/he and HH talked of education being starved of resources¹², arguing that, *“overall, children are not having as good an experience in school than their predecessors”*.

However, a supplementary question such as “How were you taught at school?” prompted splenetic comments about rote learning and physical punishment.

In summary, neither literature nor interviews provided much comment on what I have termed underlying or fundamental improvement, changes to the nature of schools that make them more humane places in which to live and learn. The focus on achievement has been

¹² Radio 4’s *Today* programme, 10th January 2024, cites a 42% increase in health spending in the twenty-first century, while over the same period education spending increased by only 3%.

overwhelming; when prompted, however, school leaders talked with passion about improvements in teaching, in the education of pupils with learning difficulties and in relationships between teachers and pupil and their families.

THEME 4.5: REASONS FOR THE RAPID EXPANSION OF SCHOOL IMPROVEMENT ARE NOT FULLY EXPLAINED.

As the literature review explained, research, policy and legislation concerning school improvement all underwent a rapid, considerable expansion between 1980 and 2010. The review also explained that the reasons for that burgeoning interest were very varied, some from within education, others from the fields of economics, politics and sociology. My conclusion is that the accounts are scattered in the literature of different disciplines, and are incomplete, lacking a full, concise account.

The topic was the subject of a specific question in the interview schedule [Appendix III, question 3]: “what factors - social, economic, political as well as specifically educational - do you think made school improvement so important?” The question was phrased deliberately to leave open the widest possible range of contexts for answers; the earlier start date of 1970 was given in case anyone wanted to refer to the Ruskin speech of 1976 (which a small number did) or other earlier events.

Answers were surprisingly thin: the transcripts of answers to this question are brief in comparison with other questions. Most interviewees are mature people whose careers were under way during the period of most rapid change in the late 1980s and 1990s and yet very few gave detailed or sophisticated answers. Most responses rehearsed some of the events chronicled in the literature review; these comments are placed first in the analysis below. There were a small number of original ideas that extended the literature review, discussed in more detail at the end.

Most comments were about politics and legislation, not surprising given that the time of rapid legislative change was at the heart of many interviewees’ careers. One group focused comment on the political drive for a competitive environment: what AA described as “*the rise of individualism and a win-lose approach*”. They linked this to the 1980s approach of prime

minister Thatcher and her education minister and guru, Keith Joseph. Like JJ and NN, LL saw the breakdown of the post-war consensus in the late 1970s as *“leading to its replacement under Thatcher with an individualistic, competitive era”* in the 1980s and 90s. JJ linked the breakdown in consensus with *“dramatic changes in the nature of employment, with the collapse of traditional heavy industry and the reduced influence of trade unions”*, leading to uncertainties in society. UU described *“market forces at the heart of Thatcherism, expressed in ERA which is based on open enrolment and competition between schools”*. GG also referred to ERA and its fellow Act in 1992 as marking *“the beginning of accountability, prescription, inspection”* and to competition expressed in *“published inspection reports and league tables”*.

Others, especially LL, focused on the emphasis on education in the Labour years from 1997, which represented *“a real sense of the importance of education to everything”* and to the significance of *“a big beast Secretary of State for Education (David Blunkett) whose clout put education at the heart of government... [leading to] a golden age for education”*. NN contrasted education before-the 1976 Ruskin speech (*“education the sole property of the educators”*) with the 80s and 90s, when it became *“a government battleground”*.

The other repeated comments were about international comparisons: FF mentioned *“the rise of Asian tiger economies and adverse PISA scores for Britain”*; LL also referred to *“international comparison and the interest in apparently successful countries like Finland and Singapore”*.

There were three contributions which, in my view, add significantly to the literature:

- GG, from the perspective of a former director of education, linked the social upheaval caused by the high-profile deaths of children such as Maria Colwell and Victoria Climbié with the 2004 legislation relation to the ECM five outcomes and a permanent expansion of the aims of school improvement.
- SS pointed to the importance to the growth of school improvement of the rise of the comprehensive school in the 1960s and 70s. This is referred to in the literature, but SS makes a distinctive point by putting it at the heart of the changing role and position of education in the second half of the twentieth century. S/he links the spread of the comprehensive to growing public interest and, following high-profile incidents like William Tyndale [p.30], to disquiet about education.

- JJ is a sociologist and historian whose perspective is distinctive within the sample. In addition to his ideas mentioned above, he points to the significance of the raising of the school leaving age (ROSLA). Although a generic term for a change that happened several times in the twentieth and twenty-first centuries, the specific event took place in 1972 when the age of leaving school was raised from 15 to 16, and

“teachers were confronted with the need to teach older, less compliant teenagers ... part of a new world of consumer teenagers, unused to being at school and not in work, earning money” (JJ).

He links this with the breakdown of deference and traditional values, the changes in employment referred to above, and therefore to the changing role of the school. He claims that schools had previously been seen as a bulwark of traditional values and an antidote to the disintegration of society, a view shattered by events such as William Tyndale School and Pink Floyd’s “The Wall”; he sees these changes as pointing to the compulsion to change.

I believe the patchy nature of the evidence in the literature, coupled with the fresh points raised here, suggests the need for a re-appraisal. As well as lacking a comprehensive, succinct account, the literature fails to convey a sense of concatenation: simultaneous pressures to change from many directions. This is attempted in the Discussion chapter.

The previous paragraph indicates where this research seeks to add to existing knowledge. Each of the five themes drawn from the research data in this chapter does this, generating an idea(s) that adds to knowledge and is explored in the succeeding Discussion chapter.

In summary, the literature contains substantial evidence for the spread of school effectiveness and improvement activity during the research period, and for its causes. However, that evidence is *partial*, in that some causes are not referred to; and *patchy* in that it is scattered across sources. There is no single comprehensive account, nor a classification of factors, nor an account of how the various factors worked together.

4.6 SUMMARY OF FINDINGS

The chapter presents four themes, recurrent facets in the data (4.2 to 4.5), preceded by a section (4.1) which tests the thesis's underlying proposition, that effectiveness and improvement dominated thinking in English schools between about 1980 and 2010. Evidence from both the literature and interviews upholds the proposition, and shows that a second dominant theme, pupils' safety and wellbeing, emerged in the twenty-first century.

As a result of the findings, each of the other four themes is explored as follows in the next chapter:

4.2 Perceptions of 'what is to be improved?' have evolved and expanded: changes in how academic achievement is measured are examined, together with the various other measures of quality and improvement.

4.3 The text analyses the conditions for school improvement, especially for how it is led: this is developed [p.140 *et seq*] into a model for the leadership of school improvement, one which takes account of complexity and of school context. The meaning of key variables such as school culture, expectations and consistency is examined in detail.

4.4 Fundamental school improvement (changes to the nature of schools that make them more humane places in which to live and learn) has been neglected because of the overwhelming emphasis placed on achievement. It is rarely mentioned in the literature, nor by school leaders until prompted. The nature of those underlying improvements is explored [p.155].

4.5 Reasons for the rapid expansion of school improvement are not fully explained: evidence is partial and patchy. A comprehensive, classified analysis of the factors is presented with an analysis of how they work together to drive improvement.

CHAPTER 5: DISCUSSION

“We have learned much about how to improve individual schools” (Hopkins et al., 2014, p.257). “Much of what is claimed as school improvement is illusory” (Coe, 2013, p.i).

Each section analyses the contribution of one of the five themes generated in Findings [Table 3, p,77] to answering the research questions and to knowledge and practice; themes are re-ordered more logically.

5.1 Testing the basic proposition that improvement dominated schools’ thinking and action in the research period was largely resolved by a literature review which showed how research and legislation related to improvement burgeoned in the thirty years from 1980. Findings offered a closer definition of “dominated schools’ thinking” and corroborated the proposition through the evidence of school leaders, who added nuance by suggesting that the process was phased.

5.2 Considers gaps in accounts of the origins and early evolution of school improvement. Research evidence suggests that a range of societal factors interacted with others from within education to cause the expansion of improvement activity. No single source presents a comprehensive, concise account of the factors, nor classifies them.

5.3 Improvement in the context of unclear purposes for education. Interviews showed that perceptions of improvement evolved into a complex web of meaning. Contention over the purposes of education is not new but its relationship to improvement is, especially in the light of further expansion in the role of schools created by the pandemic.

5.4 Conditions for Improvement: how school improvement happens. The literature provides more information about the ‘what’ of improvement than the ‘how’. This section analyses the actions taken by school leaders to respond to their interpretations of improvement, focusing on the ‘Conditions for Improvement’, and proposes a model of improvement which acknowledges the complexity of the task, incorporates new elements and attempts to fix the role of consistency, which has hitherto remained unresolved.

5.5 Considers the impact of fifty years of effectiveness and improvement activity and prospects for its future.

5.1 TESTING THE BASIC PROPOSITION (that improvement dominated schools' thinking and action in the research period)

Analysis in the literature review, especially that based on Hallinger & Kovačević (2019), showed that research burgeoned following the Rutter paper of 1979; similarly, analysis of legislation showed rapid acceleration from the mid-1980s.

The basic proposition having been secured, research data show it to be nuanced: activity following 1979 was phased. The research community responded quickly, with an immediate flurry of research to extend Rutter. Government responded later: legislation through the early 1980s, although it improved schools as places to learn by improving provision for pupils with special education needs (1981 Act) and by abolishing corporal punishment (1986 Act), was generally unrelated to effectiveness research. The 1988 Act enacted fundamental change in schools, some of which (such as local management) linked with effectiveness research, but was driven more by the philosophy of the market. From the 1990s, effectiveness and improvement moved centre-stage in government thinking, with the introduction of Ofsted (1992) and its adoption of school effectiveness criteria (Sammons *et al.*, 1995), and the bringing of leading researchers into positions of political influence: David Hopkins (senior adviser in the DfE), Michael Barber (head of the government delivery unit) and David Reynolds (chair of the national numeracy task force).

The response of schools was slower, as it had been to the Ruskin speech; individual schools took the lessons of Rutter on board; some LA officers, like Brighouse in Oxfordshire, made use of the research in their professional development programmes, but wholesale change was driven first by the flurry of vocational programmes [p.112] of the mid-1980s, and then by ERA and the advent of Ofsted.

Although the pace of legislative change related to school improvement slowed after 2011, the drive for improvement was maintained through the following decade by MATs, which had been given the lead role by the Education Act 2011, by headteachers and especially by Ofsted.

Improvement in the 2010s was extended to give greater prominence to personal development issues: safeguarding, welfare and young people's mental health. Schools have always accorded these issues great significance, for their own sake and for their impact on pupils' academic development, but the demands became much greater in the 2010s. To

argue that these issues are separate from school improvement would be perverse: their interconnectedness was emphasised by almost all interviewees, most clearly by WW: *“unless young people are well fed, listened to, happy and secure in school, their propensity and ability to learn must be limited”*.

5.2 ORIGINS AND EARLY EVOLUTION OF SCHOOL IMPROVEMENT

The literature review made clear that diverse pressures for schools to provide a higher-quality, more relevant education characterised the last quarter of the twentieth century. Accounts in the various sources agree on main points but are partial, not joined up, and fail to convey the concatenation of simultaneous pressures from different directions, and how their mutual reinforcement accelerated change. Evidence from Interviews and personal experience added factors not mentioned elsewhere.

The following analysis lists and classifies the pressures for improvement without repeating detail from the literature review; the classification into educational, societal, economic and political pressures is helpful but simplistic: they overlap.

5.2.1 Educational pressures for change

Older interviewees raised several concerns about schools which they believed promoted the ‘improvement revolution’.

Disquiet in the profession and the public with aspects of education in the 1960s and 70s conflated the loss of grammar schools with the rise of so-called ‘progressive’ teaching methods, leading to the writing of the Black Papers. When the Tyndale affair in particular hit headlines, concern became profound and widespread. Tyndale was the only individual school named frequently (six interviews); interviewees with long educational memories described schools in the 1970s as *“sharing the turbulence of the age”* (JJ).

ROSLA from 15 to 16 in 1972 retained in school a group of *“older, less compliant teenagers who had expected to be at work and earning”* (JJ), at a time when teenage consumerism and distinctive ‘teenage culture’ were expanding.

Callaghan's Ruskin Speech (1976) launched the 'Great Debate' on education. Seven interviewees named the speech in discussions of the origins of the school improvement movement, surprising since only four were involved in education at the time. The term 'watershed' was applied by interviewees to three school-improvement-related events during the period: Ruskin, ERA and the creation of Ofsted.

There was a widespread belief in the profession (Hargreaves, 1982; Partridge; 1968; also interviews with NN, TT and others with special interest in less-able pupils) that pupils of lower ability had been ill served in the grammar school era. Until the introduction of CSE in 1965, they had left school without the qualifications that the job market increasingly required. Several interviewees are products of grammar schools: GG, NN, JJ, HH and RR referred to the social disadvantage conferred on friends or siblings who had failed the 11+ examination and were therefore *"relegated to the secondary modern school"* (NN). Many pupils had suffered *"a destruction of their dignity ... so massive and pervasive that few subsequently recover from it"* (Hargreaves, 1982, p.17).

Pupils with special needs had fared badly: the Education (Handicapped Children) Act, 1970, was the first to guarantee access to education for all children; those with severe learning difficulties, previously the responsibility of the health system, became part of education for the first time. NN, with a lifetime working with such pupils, was especially trenchant about how ill-served 'his pupils' had been.

Criticism of the country for failing to provide vocational and technological education to fit children for future jobs (the Sputnik argument) had existed for over a century. The accelerating pace of change in industry and employment, combined with ROSLA, brought this disquiet to a head in the 1980s with a flurry of vocational initiatives (TVEI 1983, CPVE 1983, BTEC 1984 and NVQ 1986), a significant response to the Great Debate.

5.2.2 Societal pressures for change

The literature review referred to many societal changes leading to turbulence and creating pressure for change in schools. Each older interviewee named one or more aspects of the 1970s as direct causes of the subsequent drive for school improvement; I am grateful to JJ for his overview of the pressures of the 1970s: he (and NN) claimed that *"schools had been seen as an antidote to the disintegration of society until they too shared in its turbulence"* (JJ).

These changes were not simultaneous but centred on the 1970s:

- the rise of youth sub-cultures in the 1960s: mods, rockers, skinheads
- 'liberalising' legislation passed by the Wilson government of the late 1960s contributed to the spread of the so-called 'permissive society'
- traditional notions of discipline were challenged, leading to the abolition of caning in English schools in 1985-6
- schools became directly involved in the permissiveness debate through *The Little Red Schoolbook* (Hansen & Jensen, 1969), which sought to inform young people about democracy, rights, sex and drugs. It was widely seen, including by the UK government which banned it until 2014, as subversive, encouraging young people to experiment and to challenge authority
- the breakdown in rigid class structure and deference to established authority: "*people no longer knew their place*" (JJ)
- family size decreased¹³ and families became more child-centred: parents spent more money and time on their fewer children, and legislation concerning child welfare and safeguarding expanded
- a society in which worth was based on social class was replaced by one which valued talent and achievement, as satirised in *The Rise of the Meritocracy* (Young, 1958). The rise of the comprehensive school was symbolic of meritocracy
- the first high-profile death through familial child abuse, that of Maria Colwell, took place in 1973. A series of such deaths, culminating in that of Victoria Climbié (2000), led to the 'ECM' legislation (2004) and to an expanded concept of school improvement
- Pink Floyd's *The Wall* ("we don't need no education") was issued in 1979.

These events bracketed a decade in which wholesale change in British society was perceived by many to amount to breakdown. When breakdown was seen to manifest itself in schools themselves, "*the pressure for change became irresistible*" (NN).

¹³ Office for National Statistics reports (ONS 2021) that average family size fell from 2.29 children per family for women born in the 1930s to 1.91 for those born in 1969, and that the level of childlessness among women born in 1969 (18%) is higher than for women born in 1942 (11%)

5.2.3 Economic and political pressures for change

When Andrew Adonis¹⁴ visited a school in Sunderland, his guide said: “Twenty years ago, boys leaving here walked down the hill and turned left to get a job in the shipyard or right to go down the mines. All those jobs have gone; they might as well walk straight on into the sea” (Adonis, 2012, p.22).

By 1980, it had been clear for some time that the future of the British economy and the nation’s competitiveness depended less on raw materials and more on the skills, and thus the education and qualifications, of its workforce. Thus the ‘sputnik argument’ led to the rise of ‘the needs of the economy’ within the government’s set of purposes for education, tracked in the literature review and developed below [p.119]; the relevance here is the considerable impetus given to school improvement.

Traditional working-class jobs had dwindled and there was a sharp increase in the proportion of the British population identifying as middle class; their increased prosperity was accompanied by increased aspiration, for property- and share-ownership, and for better education and futures for their children. The new jobs required qualifications, putting pressure on schools to deliver better results for more pupils.

The rise of neoliberalism under Joseph and Thatcher and the breaking of the post-war consensus took education out of the welfare state and into the world of choice-individualism-market competition. The government saw this not only as a matter of ideology but also the means of bringing about improvement. Almost all interviewees saw ERA as a watershed.

The theme of international competition included, from the 1990s, educational tests such as PISA, TIMSS and PIRLS. Data from these studies have been used by governments to guide policy for school improvement.

The final factor promoting improvement was that it was contemporaneous with the quality movement. The dramatic rise of technological industries in Japan in the 1960s and 70s was accompanied by a revolution in the quality and reliability of products, so-called ‘total quality

¹⁴ Minister for Schools in the Labour government, 2005 - 2008

management'. This concept spread across the western world in the 1970s and 80s, reaching education in the 1980s with the two key texts published in the early 1990s: *Total Quality Management in Education* (Sallis, 1992) and *Total Quality Management and the School* (Murgatroyd & Morgan, 1993).

The existing literature does not identify all these causes of the rise of improvement. There is very little mention of *The Little Red Schoolbook*, ROSLA, child-centredness and family size, and none of the rise of meritocracy or TQM. More significant to clarifying this period in educational history is the failure to describe the convergence of factors to drive improvement: push factors from 'society' simultaneous with pull factors from within education, nor of the simultaneity of the zenith of those pressures for change in the early 80s with the 'enabling mechanism' of change and improvement: the shift from a 'schools make no difference' to a 'schools make a significant difference' paradigm, hinged at the paradigm shift in 1979. I have sought to show how these factors interact and mutually reinforce. The rapid flowering of effectiveness research in the 1980s, analysed below, showed that improvement was possible and pointed its direction. This is, I believe, why the educational changes of the late twentieth century in England were of two types: structural reform, breaking the local government monopoly of schooling, establishing parental choice and a marketplace; and improvement, linked to the reform changes but emphasising quality: national systems of curriculum, school inspection, tests and examinations and the publication of their results.

The second part of this section explains the rapidity with which school improvement activities grew and spread, another phenomenon attracting very little comment in the literature. The literature review described the speed with which research and government action spread following the birth of school effectiveness in 1979. In my interview with HJ, a biologist, we came to liken the accelerating change to adaptive radiation, defined in evolutionary biology as:

"a process in which organisms diversify rapidly ... into a multitude of new forms, particularly when a change in the environment makes new resources available ... or opens new environmental niches" (Schuiter, 2000, pp.10-11).

Schuster goes on to show that the opening of a new ecological niche provides opportunities which a diverse range of organisms can exploit, and the rapidity with which this happens. The mass extinction of the dinosaurs opened a range of new environments and resources for mammals, which exploited them rapidly, flying, climbing, hopping, running and swimming within a (geologically) short time (Feduccia, 1995).

The 'new environments' created by the school effectiveness paradigm led in the 1980s and 90s to the burgeoning of research into the shared characteristics of more and less effective schools, into the various lists of correlates [p.36], and into the distribution of the effects: whether they are the same in primary and secondary schools, in subject departments as well as schools, and in different countries. When effectiveness evolved into improvement, governments and schools saw the means to bring about rapid change. Within two decades, research was taking place world-wide, and governments were legislating change, not just in the early research foci of the UK and US, but widely through Europe, Canada and East Asia.

The final part of this historical section concerns the validity of the term 'paradigm' as applied to school effectiveness and improvement. Several leading authors in the fields used 'paradigm' to describe effectiveness, improvement or both and the use of the term has persisted: Reynolds and Hopkins (1993), Doyle & Wells (1997), Sammons (2007), Hallinger & Kovacevic (2019) and Hopkins (2022) all claim it. As the literature review concludes, in no case is the term defined, nor its appropriateness examined. This section therefore tests the evidence concerning whether the term is justified in this context.

Despite being hi-jacked into vacuous 'management speak' and thus falling into disrepute (McFedries, 2001; Khan, 2002), there remains a place in scholarship for the term 'paradigm' as part of the suite of terminology used to categorise the translation of ideas into scientific laws. A hypothesis is a starting point, a supposition based on limited evidence, while a theory is better established, with more supporting evidence. Terms are often used hierarchically, with paradigm at the top, the implication being that it describes processes of considerable significance, longevity and/or wide applicability.

Kuhn, originator of the term, defined paradigms as "universally recognized scientific achievements that, for a time, provide model ... solutions" (Kuhn, 1970, p.10). Paradigms come to dominate the landscape to such an extent that they determine what constitutes evidence and preclude dissent; contrary evidence is rationalised away and the belief becomes 'taken for granted'. Paradigms are the spectacles through which we look but of which we are barely aware. Contrary evidence is sometimes ignored or twisted to meet the existing paradigm. Eventually, the weight of contrary evidence becomes so great that the paradigm yields and is replaced by another.

Some so-called educational paradigms more closely resemble models, providing different simplifications to provoke discussion about the nature of education. Thus constructivism, often described as a learner-centred 'paradigm', is contrasted with instructivism, which tends to the teacher-centred (Huitt, 2019, p.4). These are alternative ways of thinking, not a lens through which generations of scholars view education and which, following new evidence, yields to another.

School effectiveness and improvement began as fields of study through a clear paradigm shift: the 'schools make a difference' shift in 1979, described earlier. The previous 'schools make little difference' view meets the paradigm test: it framed thinking and controlled policy. Hallinger's & Kovacevic's (2019) bibliographic investigation into education research themes identifies 1979 as the point of a paradigm shift, with effectiveness and improvement as one of the four main growth themes in their period of study (1960 - 2018). The more significant shift is, in my view, that from 'no effect' to 'significant effect', since the latter leads directly to a concept of the effective school and what makes it.

According to Kuhn, it is characteristic of the period following a paradigm shift that there is a diversification of research related to the new paradigm (see adaptive radiation, p.115), part of which was the emergence of school improvement, a natural extension of the static notion of effectiveness into a dynamic process that applied the effectiveness correlates to make more schools more effective. As early as 1989, Creemers, Peters and Reynolds were claiming in the introduction to their book summarising the field that:

"school effectiveness and school improvement have different origins: ... However, in their orientation to outcomes, input, processes, and context in education, they also have much

in common (they) were combined in a framework of effective school improvement.”
(Introduction, p.i)

To apply the term ‘paradigm’ to improvement does not seem justified, for four reasons:

- effectiveness elided into improvement without any abrupt change
- research into improvement began only a few years after that into effectiveness
- many of the same authors were involved
- although sometimes described as separate ‘schools’, effectiveness and improvement use the same data and, by 1989, were being harmonised.

It seems sounder to describe effectiveness and improvement as a twin or joint paradigm. The literature review [p.12] examines the dissenting evidence from authors like Coe (2013) who, after only about 20 years of the so-called improvement paradigm, begin to cast doubt on its impact, and indeed existence; they generally do not dispute effectiveness.

5.3 IMPROVEMENT IN THE CONTEXT OF UNCLEAR PURPOSES FOR EDUCATION

A clear theme to emerge from interviews was that improvement is hampered by the lack of a clear statement of the nation’s aims for education (UU, PP, TT *et al.*). It was argued that the result is a Professor Joad¹⁵-like “it depends what you mean by ...”: without clear objectives, the idea of improvement lacks meaning and direction, “*elusive and slippery*” (PP).

The literature records the various formulations and discussions of aims over time, to which the current work adds three new elements:

- by charting changes in government statements of aims, an analysis of the direction of travel of political thinking about the purposes of education
- an analysis of how the lack of nationally agreed aims affects school improvement
- since it involves the same principles, an exploration of the issue raised by several interviewees: the relative merits of pursuing ‘improvement’ or the ‘good school’.

¹⁵ Joad was a twentieth century English philosopher, renowned for beginning his answer to any question with “it depends what you mean by ..”

5.3.1 The recent convergence of government statements of educational aims

By the mid-twentieth century, three main traditions of the purposes of education existed:

- essentialist: emphasising content and knowledge, what Peters (1983) referred to as the acquisition of worthwhile knowledge
- utilitarian: education as preparation for adult life and employment
- child-centred: a concern for what education should develop in young people, rather than what it should seek to pass on to them.

Interviewees referred to all three traditions; most wove them together to create a pragmatic compromise. RR typified this: *“reading, writing, basic maths and knowledge of the world ... careers education and employability as key components of a great education”*, while describing how schools *“seek to develop basic values of decency”*.

To track government statements of aims through the research period is to witness a drift toward education as the engine of the economy. This analysis is not repeated here; key moments are identified.

At a conference in Paris in 1970, Margaret Thatcher, Secretary of State for Education, had been asked:

“what are our educational goals in Britain and how we set out to achieve them. I have no ready formula at my fingertips. There are certain fundamental aims which are as permanent as they are far-reaching. These are: to afford to ... every man and woman, the best possible opportunity to discover and to achieve his or her full potentialities for self-fulfilment, not as an individual over against society, not as a social or economic unit subordinated to society, but as a “person-in-community” ... [and] to do this in such a way as to offset so far as possible cultural and environmental handicaps to equality of achievement to do it in such a way as to make the best possible match between the aspirations of each person and the manpower needs of the community; and, last but not least, to do it in such a way as to give the greatest possible scope to the participants in the educational process ... to contribute their energy, their wisdom, their ideals, and their skills to the success of this complex social and political enterprise.’ (Thatcher, 1971, my emphasis)

I call attention to the italicised highlights: the lack of a “ready formula”; the frequent references to society, not just the economy; the pervasive notion of personal development and fulfilment; an early reference to closing the disadvantage gap; the scope for participation; and the permanence of these aims. In the same speech, she set out what she describes as “the central controversy of education ... the individual, child-centred and the subject-centred theories” (*ibid*) and acknowledges the tension between them.

In his 1976 speech launching the Great Debate, James Callaghan said that:

“The goals of our education ... are clear enough. They are to equip children to the best of their ability for a lively, constructive, place in society, and also to fit them to do a job of work. Not one or the other but both.” (Callaghan, 1976)

This represents a distinct shift towards the utilitarian.

Conservative government guidance issued in 1981 contained six aims, acknowledging all three traditions. Two of the six are specifically utilitarian, including “to help pupils to acquire knowledge and skills relevant to adult life and employment in a fast-changing world” (DES, 1981, para. 11). Essentialist ideas are also represented: “to help pupils to understand the world in which they live”: worthwhile knowledge. A child-centred approach is also threaded through, including “to instil respect ... and tolerance ...” (*ibid*).

A notable feature of this aim set is the tentative way it is introduced: the aims are prefaced by the statement that they are “offered”, “individual authorities and schools might refer [to them] in drawing up their own lists” (*ibid*). The tone of subsequent versions is markedly different: just six years later, in its Education Reform Bill (1987), the caution of 1981 and the variety of aims and tensions outlined by Thatcher in 1971 have been abandoned. The non-interventionist approach to schools and curricula that had prevailed throughout the century gave way: instead of “offering” suggested aims, ERA introduced an extensive, compulsory National Curriculum, shaped around subject knowledge, with an underlying philosophy combining essentialism with utilitarianism, leading Jack Straw, Shadow Secretary of State, to describe the draft legislation as the “Education Reform (Central Control) Bill” (Hansard, 1987b).

ERA itself describes the aim of the curriculum only in essentialist terms, the knowledge young people should acquire; a later formulation states:

“[to prepare] pupils for the opportunities, responsibilities and experiences of adult life” (DfE, 2014, para.2.1) and to “provide[s] pupils with an introduction to the essential knowledge they need to be educated citizens. It introduces pupils to the best that has been thought and said, and helps engender an appreciation of human creativity and achievement.” (*ibid* para.3.1)

The first statement is the clearest expression of utilitarianism, the latter equally clearly essentialist.

The section of the Queen’s Speech presaging ERA made clear that it was to be a vehicle for raising standards:

“My Government will take action to raise standards throughout education and to extend parental choice. Legislation will be introduced to provide for a national curriculum for schools ...” (Hansard, 1987a).

- a programme predicated on a belief in the power of market forces to drive improvement; ERA represents the victory of both the essentialist-utilitarian subject curriculum by enshrining it in legislation, and that of market philosophy by ensuring a competitive environment.

Government statements of aims continue this utilitarian-essentialist model: in 2015, Minister of State Nick Gibb spelled out the government’s three purposes of education:

“Education is the engine of our economy, it is the foundation of our culture, and it’s an essential preparation for adult life ... introducing [pupils] to the best that has been thought and said, and instilling in them a love of knowledge and culture for their own sake. But education is also about the practical business of ensuring that young people receive the preparation they need to secure a good job and a fulfilling career” (DfE, 2015).

In 2022, the DfE’s website ‘front page’ (DfE, 2022b) lists the government’s education priorities, beginning with “drive economic growth through improving the skills pipeline, levelling up productivity and supporting people to work”, and referring to early education and childcare as “help[ing] parents to work”.

The emphasis on education driving the economy has strengthened progressively in government literature and is now always placed first in the list of objectives. In the run-up to the 2019 general election, the education section of the Labour party manifesto (Labour Party, 2019, p.38) began with the words: “Education makes our economy stronger”. Likewise, the first reference in the education section of the Conservative manifesto in the same year (Conservative Party, 2019, p.13) is to ensuring that “every pupil gets the qualifications they need for a prosperous future”. Primacy in both documents goes to the economy.

5.3.2 How the lack of clear aims affects school improvement

To argue paralysis in schools because the nation lacks an agreed statement of aims would be specious; the pragmatic solution, as interviewee RR argues above, is that schools help young people become better qualified, literate, numerate, mature, decent and employable; and schools get on with this job based on their own aims.

Interview evidence suggests lack of aims causes one significant problem for schools and their improvement: lack of definition makes aims *elastic*. The list of plausible knowledge and attributes which schools should aim to develop in their pupils, and the number of social issues given to schools to tackle, can be extended infinitely.

Crucially, the point is made by many interviewees (BB, DD, HH, PP, VV) that during the last decade, and especially since the pandemic, headteachers are increasingly held accountable for the safety, welfare and mental health not just of their pupils but also of their families.

Interview evidence concerning the evolution and extension of what constitutes ‘improvement’ was discussed at length in the Findings chapter. The importance of safeguarding and child protection cannot be disputed. Unlike other Ofsted judgements, graded on a four-point scale, safeguarding is binary: effective or ineffective. It is also a limiting judgement: if judged ineffective, the school will almost certainly be placed in special measures regardless of its other qualities. The range of requirements on teachers set out in the key safeguarding manual, *Keeping Children Safe in Education* (KCSIE) (DfE, 2023a, revised annually) is very large and the penalties for non-observance, harsh. KCSIE is a 178-page document; as a governor, I had to sign annually that I had read it and pass a test to check my

knowledge. As an example, all teachers “should be particularly alert to the potential need for early help for a child who ...”, followed by 15 categories of vulnerability, including “at risk of so-called ‘honour’-based abuse such as Female Genital Mutilation or Forced Marriage” (p.9). Teachers failing to report evidence of FGM may be liable to disciplinary action (p.154).

Like the claims to inclusion in the curriculum referred to above, it is hard to object to any of the many extensions of the responsibilities of teachers and schools in the cause of protecting children. And yet there was genuine alarm among some of the headteachers interviewed lest the substantial extension of accountabilities acquired during the pandemic become permanent. Every current headteacher interviewed both accepted the importance of child protection and expressed concern about their, and their schools’, vulnerability in this area. BB [p.79.] pointed to the tension between the competing demands placed on him to raise achievement in the school and to support the mental health and welfare of the community, and his inability to undertake both. By contrast, interviewee WW described the one as a pre-requisite of the other: *“unless young people are well fed, listened to, happy and secure in school, their propensity and ability to learn must be limited”* (WW).

The National Health Service Survey (NHS, 2023) of mental health among young people reports that 20% of school-age children have a probable mental disorder, and an eighth of 17-year-olds have an eating disorder. The daily experience of teachers therefore involves attempting to raise the standards achieved by many young people whose prior issues (physical, mental, social and economic) impair, and frequently preclude, such progress.

Loosening this Gordian knot cannot be a matter of either/or: schools’ resources of time, energy and money cannot embrace all the demands on them, and it is increasingly clear that those resources must either be enlarged and/or harnessed to those of other agencies, such as health and social care. The Children’s Commissioner’s report *Big Ambition* (2024) summarises young people’s views concerning how the government could improve their lives and responds with detailed proposals. The section on education (p.38 *et seq.*) identifies the acute need for children’s mental health and support services; the following suggestions are made, among others, which would go a considerable way to resolving the tension described (pp.45 – 46):

- co-locate mental health services in every school

- deliver children’s support services, including “educational psychologists, speech and language therapists, Children and Adolescent Mental Health Service practitioners, social workers, youth workers, and family liaison officers” (*ibid* p.45) on school sites
- a school nurse in every school.

5.3.3 The pursuit of school improvement or the ‘good school’

Several interviewees, notably JJ but also FF, RR, TT and others, raised the relative merits of pursuing school improvement or the ‘good school’, presenting strong, varied arguments for the latter.

JJ does not oppose the pursuit of school improvement but rather its limitlessness. This is partly a philosophical objection: since Ofsted for most of its existence relied heavily on test and examination results to steer its judgement of schools, and since those results are norm-referenced, school improvement is limitless and there will always be first and last carriages on the train. His is also an emotional argument, the effect of endless pressure for improvement being oppressive of teachers’ morale and thus counterproductive. “44,000 teachers left the state-funded sector in 2021/22¹⁶, up by 7,800 since last year. This represents 1 in 10 ... of all qualified teachers” (UK Parliament, 2024). This is the highest proportion since records began. The General Secretary of the Association of School and College Leaders listed causes of this drain including “an eye-watering system of performance tables and Ofsted inspections, all of which is deterring recruits and driving out teachers” (Barton, quoted in *The Independent*, 8th June 2023).

FF and BB reason differently, believing that school improvement is associated with an excessive focus on a narrow set of measurable outcomes, especially test and examination results, and that this both distorts the purpose of schooling and fails to engage all those who work in schools. FF believes that “schools prosper when they have an ideology, a belief system, a philosophy.” FF’s leadership of schools focused on what s/he called “doing the right

¹⁶ in-career, not retirements

things rather than obsessing about outcomes”: the consistent pursuit of a small set of key attributes of a good school. Therefore, when FF inherited a school with examination results in the bottom 5% in the country, rather than emphasising improving those results, s/he gained the school community’s commitment to the purpose, values and priorities it should pursue, such as valuing everyone and never wavering from a focus on the quality of lessons. BB also discussed changing the culture, *“everyone needs to be able to articulate what the school is about, what it believes in”*. BB was one of the many who emphasised building the conditions in which improvement can take place, rather than improvement itself: improvement as a by-product.

RR emphasised the deleterious effects of a narrow focus on improvement: various types of gaming the system, not only by the extreme ways described on [p.86], but also by narrowing the curriculum, losing sight of what constitutes an excellent education.

It is, I believe, possible to achieve a reconciliation between these two views: both employ the same effectiveness principles, but FF’s approach to a school in deep difficulty emphasises the *how* rather than the *what* of improvement. A way forward requires a blend of approaches; attempted by the proposed school improvement model [p.140 *et seq*].

5.4 CONDITIONS FOR IMPROVEMENT: HOW SCHOOL IMPROVEMENT HAPPENS

“We believe that educators, policymakers, and the public cannot be expected to do what works until they actually know what works.” (Kidron & Darwin, 2007, p.24)

“School improvement requires remorselessly bloody hard work” (JK).

From a discussion of this theme, a new school improvement model is built, drawing on interviewees’ views of the *conditions* and *skills* necessary. Interviewees frequently referred to conditions propitious, or pre-requisite, to school improvement; BB described his remarks as *“an analysis of how to bring about the conditions for improvement”* while SS described the MAT as *“expert in creating the conditions for improvement”*.

This section first explores whether users mean the same thing by, and envisage the same purpose(s) for, the term ‘model’. A working definition of ‘school improvement model’ is proposed.

Existing improvement models are then reviewed; they tend to be simple, one function of a model being to simplify reality; but MacGilchrist's & Mortimore's (1997, p.213) research into improvement plans and models concluded: "The more effective the plan the more complex are its characteristics". The proposed model acknowledges complexity, which comes chiefly from aspects of school context, three of which are considered:

- different points on schools' improvement trajectory
- interviewee SS's concept of "*constrained and unconstrained schools*", the added difficulty of bringing about improvement in schools serving areas of significant disadvantage
- individual context specificity, the suggestion by several interviewees (AA, HH, FF, WW, PP, RR) that, at least in the detail, school improvement is idiosyncratic.

The discussion concludes that the new model should deal in principles while acknowledging that schools are individual.

In the final sub-section, conditions and skills are summarised and the new model is propounded.

5.4.1 Interpretations of the term 'model'

A major article on school improvement models (Kidron & Darwin, 2007) uses the term over 200 times without defining it. Every interviewee used the term 'model'; gentle pressure to explore their meaning elicited terms like "*components*" (CC), "*system*" (LL, PP) and "*inter-connectedness*" (JJ). Several implied diagrammatic. The following text is a distillation of evidence from interviews, together with that from a search based on 'models of social systems'.

'Model' carries overtones of idealisation (model pupil, fashion model), simplification (model aeroplane) and of a version of a machine (latest model). The more arcane use of 'model', describing systems in social science disciplines, contains the same notions: such models are simplified and idealised versions of reality.

Schools are complex social systems since their detailed behaviour cannot be predicted. To create an improvement model must therefore be an attempt first to *clarify* components and inter-relationships by being logically displayed, perhaps diagrammatic; second, to *explain* the

relationships or interactions between those components; and third to *predict* broad patterns of behaviour. It is an abstraction of reality, not perfectly ‘accurate’, because the complexity of schools precludes detailed prediction; thus the model leaves space for co-construction to meet local circumstance.

5.4.2 Existing school improvement models

RR described *“existing ideas of school improvement [as] predicated on a failing school in a deprived area, not on mediocrity in the unambitious suburbs or the coasting school”*. Text on [p.62] describes how headteacher interviewees were selected to include a high proportion who had led both a failing and a good school; although no-one described it as easy, there was widespread agreement that the process of turning round the dysfunctional school is tried and tested: *“it isn’t difficult to know what to do when a school needs to start again”* (RR, also PP, CC, AD, NN. TT), while bringing about the sustained improvement of a sound school is not as well documented.

Headteachers who had led the improvement of a school in difficulties (AD, CC, LL HH, FF, TT, BB) all described a ‘turn-round model’. The best-known current formulation is David Carter’s four-stage version (reproduced in DfE, 2016, p.28): Stabilise, Repair, Improve, Sustain, a clear analysis of the process for dealing with a ‘broken’ school. That the focus is on the dysfunctional school is shown by the greater specificity of the leadership qualities in the ‘Stabilise’ than in the ‘Sustain’ phase (*ibid*). This is one major principle underlying the notion of the MAT: that the take-over of a weak by a strong school/trust will lead to the former’s improvement, a process identified in early research for the DfE (Potter, 2004). Since 2010, this has been the government’s main school improvement mechanism.

But these headteachers’ turn-round processes varied: YY described maintaining the school’s focus on behaviour, even though the improvement has been continuous for 12 years, because *“these students need every minute to learn ... so ‘good’ behaviour isn’t enough”*. FF focused rather on teaching quality, treating behaviour as a by-product. But both describe *“a relentless focus on classrooms”*. TT’s, LL’s and CC’s turn-round models for schools which had been profoundly weak over several years were based on *expectations*, articulated by CC as follows:

- 1 *“Set standards, make expectations of teachers and leaders clear.*

- 2 *Hold everyone to account to those standards.*
- 3 *Identify and remove people who won't.*
- 4 *Improve people who can't but who want to.*
- 5 *Build on people who can."*

HH has similar experience but employs a different turn-round model, focusing on the installation of systems:

1. *"Install technology, staff and student portals which rapidly enable them to understand aims, values and processes, while encouraging collaboration and communication."* S/he describes IT as *"a key part of changing the culture"* and says the organisation *"hangs improvement efforts on technology"*.
2. *Install new meeting structures for professional development, communication and joint planning. Teams meet within the working day.*
3. *Install the MAT curriculum and assessment structure.*
4. *Remove existing financial structures and replace."* (HH)

FF, XX and BB described an improvement process with some similarities to the above but emphasising **culture**. BB described the process of building *"an agreed sense of what it means to work here ... the foundation for staff buy-in"*. FF described these broken schools as *"racked with self-doubt spending too much time agonising rather than doing"*; key to FF's process is *"changing the culture, creating a belief in the possibility of change and improvement"*. In a telling phrase, FF believes that *"school improvement is intimately linked with institutional confidence"*.

"What is missing from the strategy is any notion about *school or district culture*. If theories of action do not include the harder questions – 'Under what conditions will continuous improvement happen?' and, correspondingly, 'How do we change cultures?' – they are bound to fail" (Fullan, 2006. p.4).

FF, XX, HH and BB all linked culture with **confidence**. This was particularly true of XX, who reflected on work in two schools, one of them broken, the other in a state of *"tick-over"*, with staff who were *"a disempowered group of people who turned up to work with little sense of engagement or agency"* (XX). In the profoundly weak school, XX found that advice for

improvement had come from a succession of headteachers, LA advisers and other 'experts'; in common with the experience of many broken schools, that advice differed, sometimes conflicted, leaving staff disempowered and renouncing responsibility: *"tell me what to do and I'll do it"* (XX).

Recognising that this disempowerment mitigated against improvement (*"unless there is trust in the leadership, staff will not take the risk to innovate"*), XX set about changing the culture with an emphasis on the building of trust through openness. By explaining to staff how the budget worked, XX removed the threat of redundancy; by encouraging staff, not just teachers, to visit each other's classrooms in a low-stakes way, s/he transformed staff attitudes to lesson observations from threat into investment in improvement. The 'tick-over culture' became one in which staff wanted to come to work. Confidence comes, says FF, especially from *"implementation and feedback concerning progress and next steps"*.

It is noteworthy that all headteachers cited the need to shed those staff who could or would not measure up, but they described the process differently. Some emphasised the setting of standards and moving staff out; others described the building of a culture which draws staff in and leaves others isolated, who are shed. These discussions led in part to the see-saw analogy [p.143], the former prioritises lightening the negative end of the see-saw, the latter sliding as many as possible from the negative to the positive. YY describes going to considerable lengths to bring staff onside: a lengthy process of discussion and tailored professional development.

I asked those who emphasised the building of a new culture (FF, XX, HH, BB, YY) first in what terms the new culture was expressed, and second, how it raised confidence. All five gave an account of how they developed the school's values and how these were *"communicated, pursued and reinforced"* (FF). Values were developed with a contribution from the school community (BB, FF), the process led from the top, and articulated in language pupils and parents can understand (*"kidspeak"* FF) and communicated to all. These interviewees' views matched Fullan's (2007) definition of school culture, the guiding beliefs and values of the school; but Fullan goes on to emphasise that those beliefs should be evident in the way the school operates. Interviewees stressed that values are reinforced constantly in meetings, newsletters, one-to-ones with staff, pupils and parents. They spoke about the courtesy with which staff and pupils would address each other (FF); that all staff and pupils would be valued

equally (FF, BB, HH); staff would have *“clear responsibilities within the school improvement plan and be held accountable for them”* (BB, AA); that the school would be open, sharing views and *“challenging each other”* (BB).

Although they describe them in different ways, these leaders’ approaches have the same concept at their heart: the establishment of a set of values and beliefs around which the school community will coalesce, and which defines those attributes around which the school will build consistency: a framework of expectations.

Professional development was described as a key part of a culture, building consistency and valuing staff. XX emphasised professional development as *“investment in people, with benefits not just in the knowledge gain but the sense of being a significant part of the team”*.

Several interviewees (especially XX, YY) mentioned the **pacing of change**. Both describe subtle leadership, balancing a sense of urgency with avoidance of frenetic activity. Again, openness is described as crucial: the community understands what needs to be done and that there is a clear plan for its completion over time (ZZ, BC, JK). YY described at length how s/he sought to pace change *“so that staff aren’t distracted from their focus on teaching and supporting pupils as well as possible”* (YY).

The development of a culture implies **reliability**, a term (or its synonym) used in almost all interviews, sometimes the high-reliability approach (QQ), more often in the commonplace sense of ‘doing what one says’. Reliable implementation (fidelity, [p.141]) led to a discussion of the role of consistency in improvement. Consistency has been much disputed in education; locating exactly where it is crucial to improvement is built into the model [p.140]; adherence to core values is, from the message of most interviewees, one such aspect.

All interviewees took it as read that improvement efforts must involve the **whole school**: all staff, governors and, in their own ways, pupils and parents. In an early study, MacGilchrist & Mortimore (1997, pp.207-209) identified four types of improvement plan in schools, which they termed rhetorical, individual, co-operative and corporate. They found that

“the corporate plan, characterized by a united effort to improve and a strong sense of shared ownership and involvement by the teaching staff” (p.208) as having “the most positive impact” (p.209).

Compare this with the author's reflections on the school in which he began teaching, in which isolated individuals and departments rendered school improvement an impossibility [p.32].

The models considered so far have been generated by experience, although Carter's is also based on his extensive work with the executive leaders of MATs. Greany's model is based on research: his detailed study of sustainable improvement in 31 MATs established what he termed "five strategic areas for improvement" and "five improvement fundamentals" (Greany, 2018, chart p.11). His strategic areas (vision, values, strategy and culture; people, learning and capacity; assessment, curriculum and pedagogy; quality assurance and accountability; and "a sustainable learning organisation" (*ibid*)) form a helpful analysis of a school's fields for improvement.

Finally in this analysis of existing models, the American concept of 'model' (Kidman & Darwin, 2007; Graczewski *et al.*, 2007) means the range of commercial improvement programmes that a school or district might buy into: 'Success for All' and 'Direct Instruction' are two of the best known. These are highly prescriptive programmes involving curriculum, pedagogy, organisation, assessment and family involvement. Analysis of them by Reynolds, Hopkins, Potter & Chapman (2002) suggests that whole-school fidelity of implementation of a programme is a more significant factor in success than choosing programme A rather than B. No comment is made here on the effectiveness of such models, but the approach has rarely formed part of the British tradition. The importance of consistent implementation is however a critical dimension of a new model.

In summary, the following points, distilled from this study of existing models, will guide the framing of the new model:

Improvement is a **whole-school** endeavour and involves consistent implementation. The new model, as well as emphasising whole-school implementation, will circumscribe where consistency is necessary.

- Corporate endeavour and consistency do not happen spontaneously, so the model concerns the **leadership** of improvement.
- Since the turn-round methodology is well documented, the focus should be particularly on **sustained improvement** and therefore (see comments by HH, FF and BB above) on securing culture and values.

- It should address all aspects of school performance (Greany's five strategic areas).

5.4.3 The significance of school context

Many (AA, HH, FF, WW, PP, RR) expressed the view that *"school improvement is contextual, serving the needs of the actual children with their distinctive needs"* (RR): 'context specificity'. A model, because it is a simplification, deals in principles, with the detail co-created by the school. To leave the principles 'bare' risks stating only the obvious; the opposite *reductio ad absurdum* is that improvement is unique to the school, after which nothing can be said. *"Somewhere along the spectrum between one size fits all and total context specificity ... lie some common-sense principles"* (PP). A new model should reject 'one size fits all', offer some depth related to context and acknowledge that schools are individual.

Interview evidence suggests that two dimensions of context be considered in the model:

The school's position on the trajectory of improvement: the turn-round and sustained improvement phases have already been mentioned. RR added the *"coasting ... unambitious"* school; in the 2010s, Ofsted used the term 'stuck' schools. This section considers whether these are legitimate separate categories for inclusion in the model.

A coasting school was initially defined as "one that over time does not support its pupils to fulfil their potential" (DfE, 2016b, p.4), a useless criterion (all pupils? how is an individual's potential assessed?) which could be alleged of any school. The detailed criteria (*ibid* p.6) describe aspects of pupils' attainment and progress in tests and examinations. 'Coasting' implies leisurely progress, not performing as well as it should (OED: "to travel without thrust from the engine"). Later, the DfE changed their definition to schools "not making necessary improvements" (Ofsted, 2024, Introduction). The 2022 Coasting Schools Regulations (DfE, 2022c, Section 4) set out criteria describing schools which have not been judged 'good', and which have flatlined or declined between inspections.

The category stuck schools was used by Ofsted and defined (Ofsted, 2024, Introduction) as those which had "consistently poor inspection outcomes throughout the last 13 years": schools that had received at least four Ofsted reports between 2007 and 2020 describing a flatlining school that was less than 'good'. This category therefore came to merge with coasting. Research into 580 such schools inspected between 2003 and 2018 (Munoz-Chereau, Hutchinson & Ehren, 2022) found that Ofsted had judged 'stuck-ness' to signify that

their outcomes and quality of education needed to improve. This unremarkable conclusion is another example of circularity, since these aspects drive the Ofsted overall grade. The study states that “stuck schools face a combination of unusually challenging circumstances” (*ibid* p.6) such as higher levels of mobility among teachers, pupils and governors; and of pupils with SEND and from poor backgrounds, than ‘not-stuck’ schools.

If coasting = stuck = constrained schools, it is not necessary to include a separate strand in the model. The potential for useful analysis suggested by the vernacular meaning of ‘coasting’ has been lost. The model [p.140 *et seq*] attempts to build in a concept of ambition.

Schools in areas of significant disadvantage: almost all interviews referred to the added difficulties of bringing about improvement in schools serving areas of significant disadvantage. SS used the phrase “*constrained and unconstrained schools*” to contrast schools in which “*the application of straightforward school improvement knowledge is likely to bring about improvement*” with those in which high levels of demand among the pupils and community add complexity to the task. Interviewees EE, CC and SH all described the major city in which they all worked as “*twin cities, one of affluence, one of poverty, the two cheek by jowl*” (CC). The model must accommodate levels of complexity, as shown in this text:

“Many schools labelled “failing” or “low performing” ... are serving highly diverse student populations—diverse cultures, languages, religions, and economic circumstances. Such diversity is challenging on two fronts—the sheer range of educational needs that schools must take into account, and the presence of significant numbers of children whose needs typically exceed the capacities of many schools to adequately address.” (Leithwood et al., 2006, p.441).

When asked to identify differences in improvement programmes in constrained schools, no interviewee described different processes; instead they referred to “*greater intensity*” (XX, NN); “*more of the same*” (EE); “*a regime of tight prescription*” (NN); “*“a longer period of standardisation, before freedoms are granted to teams and individuals”* (AA, also PP); “*frequent review and reformulation of plans*” (RR). When a school in RR’s MAT got into difficulties, s/he put in place a six-week recovery plan, with fortnightly reviews of progress carried out by an external consultant working with members of the leadership team.

The conclusion of this part of the research was therefore that constrained schools do not need different procedures, but rather intensified implementation of the same procedures.

It was shown [p.127] that Carter's four-stage process is a turn-round version. for 'broken' schools. The proposed model focuses on the conditions under which the 'turn-round' (Carter's stages 1 – 3) happens and, because of the emphasis on continuous improvement, it particularly expands on the 'sustain' phase. The 'turn-round' phases represent *transformational* change; in term of Collins's flywheel, the hard work of getting the heavy wheel to begin to turn and accelerate. The 'sustain' phase is not steady state but rather sustained improvement through *incremental* change, the basics having been established, through lighter touches to the wheel adding the energy needed to avoid entropy.

Analysis of interviewees' responses revealed [p.124] a tension between the pursuit of improvement and that of excellence (or the 'good school'). The new model embraces both, arguing that the former is represented by the turn-round phases, the latter by the sustain phase. When the school has reached the standard of an Ofsted 'good', it may continue to improve because of momentum but the focus often shifts to providing an excellent all-round education. Interviewee RR suggested this, claiming that *"an Ofsted 'good' is the basic minimum acceptable standard that should be attainable by all. Excellence means doing everything consistently well"* (RR). When the schools in the MAT were all 'good', RR shifted the emphasis on to defining and pursuing *"excellence and a great education for all"*.

5.4.4 A proposed model for school improvement

The model has five components, derived mainly from the views of interviewees:

- a. An analysis of the leadership skills necessary to bring about improvement, in constrained and unconstrained schools.
- b. The see-saw concept: staff changes in a school during the early phase of transformational improvement.
- c. A precise definition of consistency and where it aids improvement.
- d. The inverted pyramid: how each tier of management in a school creates the conditions for optimum performance in the tier it supports.

e. Intelligent accountability: how feedback through the organisation delivers formative and collegial accountability, leading to incremental improvement.

a. The **leadership skills for improvement** (summarised in Table 4) derive from the verb ‘to improve’, exemplified through pupils’ school attendance. To ‘improve’ attendance implies first an understanding that attendance is important, and second that the present situation is sub-optimal: attendance is lower than it should be, implying that comparable schools do better. This analysis involves evaluation, not just a piece of monitoring data like a percentage. To ‘improve’ attendance further suggests change toward a target, chosen to be above the present figure but below 100%, since pupils have legitimate reasons for absence. Improving attendance therefore implies that school leaders:

- understand that attendance is significant to the school’s success
- have aspiration (attendance should be better) and agency (it can be made so)
- establish procedures to evaluate, as well as monitor, performance: our data acquire meaning by being compared with those of similar schools
- based on that evaluation, set targets that are achievable and yet represent a stretch.

Having chosen the targets, leaders then set about achieving them, implying a further set of processes. Leaders must:

- design methods likely to lead to improvement
- monitor in-flight progress, to ensure that improvement is taking place
- because staff will be responsible for carrying out the plan, win their support to the cause by communicating the rationale, target and progress.

TABLE 4: FRAMEWORK OF SKILLS FOR THE LEADERSHIP OF IMPROVEMENT

Sub-themes 4.3.1 – 4.3.4 [p.88 *et seq.*] suggested four sets of skills to create the conditions for improvement. They and their components are re-stated here, with a key data extract for each.

Sub-theme (Findings 3]	Components	Key data extract
3.1 Leadership	Facing up to the need for change	<i>“The school will make no progress until it faces up to the realities of its position”</i> (EF).
	Setting (resetting) school culture and values	<i>“there is a temptation to see school improvement as a list of jobs, when what really matters is the creation of a culture”</i> (BB).
	Focusing on the right things	doing the right things rather than obsessing about outcomes: <i>“do the right things and outcomes will follow”</i> (FF).
	Building teams (see also People)	School improvement is impossible in an individual culture: it must be a team game.
	Ability to deconstruct success	<i>“of the many attributes of good leadership, one key one is the ability to understand success and how it happened”</i> . (KK)
3.2 Principles	School effectiveness principles	<i>“School improvement works by good leaders applying the principles of school effectiveness”</i> . (KK)
	Focus on classrooms/teaching	<i>“never waver in the focus on classrooms”</i> [FF]; <i>“school improvement is about pedagogy and practice”</i> (EE)
	Inclusivity and equal value	<i>“all staff and pupils will be valued equally”</i> (HH)
3.3 Systems	Consistency	<i>“strong systems to save time and bring about improved outcomes, not consistency for consistency’s sake”</i> (DD).
	School context: point on improvement trajectory, constrained or unconstrained school	<i>“schools are victims or beneficiaries of the houses that surround them”</i> . (SS) <i>“it is not difficult to know what to do when a school effectively needs to start again”</i> (RR)
3.4 People	See-saw	<i>“achieving critical mass buy-in”</i> (TT) <i>“getting rid of wrong ‘uns and recruiting good ‘uns”</i> (DD)
	Teamwork	See above

Expanding the goal from a single variable to whole-school improvement is exponentially more complex but involves the same processes. The evidence from interviews puts flesh on this skills skeleton, summarised from the Findings chapter into Table 4.

One leadership skill described by interviewees, lying behind the work of experienced leaders, is the ability to deconstruct success: to identify in the turn-round of a school which elements were crucial and replicable, and which peripheral. An interviewee described how the deputy head of a very successful school gained the headship of a failing school: s/he introduced various features of the previous school, such as two long lessons per day. They had worked in the previous school because they were built on the back of consistent teaching and behaviour management over many years. Introduced 'naked', they proved disastrous.

- b. The **seesaw effect** is a simple analogy describing the early stages of transformation: Several headteachers who took over broken schools described classifying staff into supporters, opponents and those in between; TT spoke of *"winning staff over or removing them"* and *"achieving critical mass buy-in"*. FF's *"relentless focus on classrooms ... [leading to] ... losing one third of the staff in the first twelve months"*. DD described a process of *"getting rid of wrong 'uns and recruiting good 'uns"*.

These data extracts conveyed the image of an unbalanced seesaw: when these headteachers took up post, they described isolation, dangling, holding on at the upper end of the see-saw, while a group on the other side of the pivot contained members both 'heavy' (by dint of status and/or personal influence) and far 'off-centre'. The physics of moments multiplies their mass by their distance from the pivot to measure their force. Removing a small number of high-negative-force individuals, and replacing them with high-status, high-influence individuals at the 'positive' end of the see-saw not only re-balances it, but the tilt also causes some in-betweeners to slide toward the positive.

- c. The **inverted pyramid** (Diagram 1, p.145) is a visual representation of the philosophies of schools that seem to have achieved improvement sustained over a substantial period.

Traditional organisation charts (organograms) are drawn with the CEO at the top, branching downwards to a line of senior executives, then middle leaders, with the workers at the bottom. This is a convenient guide to the structure of the organisation and its line management but may not convey its spirit or purpose. Those who live near the bottom of the tree may perceive it as bearing down on them.

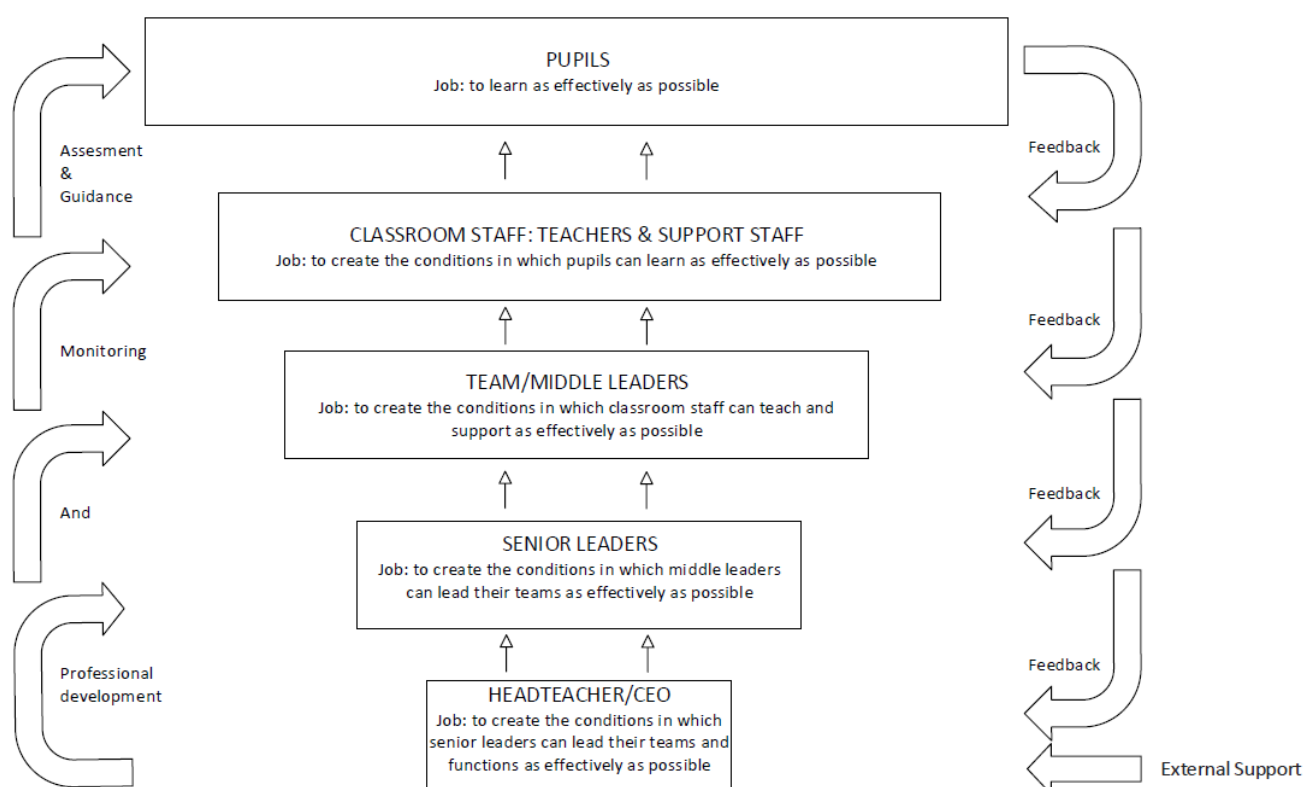
Every interviewee put the pupils at the heart of their improvement efforts: leaders referred to *“the purpose is a better environment in which students can live and learn”* (LL); *“dedication to the students, not the job”* (YY).

The Timpson high-street company describes its approach as ‘upside-down management’ while YY’s team described their highly successful school as organised around *“servant leadership”* (CD). Both concepts involve inverting the conventional organisation: first, ideas for doing the day-to-day job better come mainly from the workers; and second, the job of classroom staff is to teach and support pupils as well as they can; leadership’s job is to create the conditions which enable them to do so. Leaders, who must be visible around the store or school, ensure that issues raised by teachers are followed up, and provide feedback to them.

In the diagram, the upward arrows represent support for the tier above through the conditions for improvement, which are culture/values, policies, expertise, resources and professional development (support, challenge and training)

These ideas led to the creation of an inverted pyramid model (Diagram 1), with pupils at the top, whose learning is the ultimate purpose; successive tiers of leadership (teachers, team leaders, senior leaders, the head/CEO) whose job is to support the tier above by creating conditions for them to do their job as well as possible. Ideas for organisational improvement are fed downwards through the diagram; feedback for personal development flows upwards; in both cases, the feedback provides the energy to maintain incremental improvement, keeping the wheel spinning.

DIAGRAM 1: THE INVERSE PYRAMID MODEL OF SCHOOL IMPROVEMENT



In the diagram, the upward arrows represent support for the tier above through creating the conditions for improvement: culture/values, policies, expertise, resources and professional development (support, challenge and training)

As well as being a visual representation of a philosophy, the diagram displays two key aspects of the model:

- what 'supporting the next tier' means. The job of pupils is to learn, that of classroom staff to create the conditions in which they can do so most effectively; subject leaders create the conditions in which teachers can teach best, the headteacher creates those in which senior leaders can best lead and manage their teams and functions
- what is meant by 'conditions' in this context. The answer is the same at each interface: each tier provides the culture/values, policies, expertise, resources and professional development (support, challenge and training) which will enable the layer it supports to do the best possible job, whether this is the head to her/his staff, or the class teacher to the pupils.

Adding detail to the 'conditions' from interview evidence:

- the culture/values by which the school will run, the components of which might include commitments to, for example:
 - value all pupils and staff equally
 - the continuous learning and development of all
 - open communication
 - challenge, facing up to the need for change
 - high expectations for and from all community members
 - polite communication among all members of the school community
 - aspects of health, safety and wellbeing
 - risk-taking
 - teamwork ...
 - 'expertise' includes teachers' knowledge of subject and how to teach it most effectively, and leaders' understanding of effectiveness principles and how to manage people and organisations: for example, XX emphasised how s/he "*managed the pace of change: brisk but never hectic*". Key leadership skills for improvement have been listed [p.142]
 - 'policies' are the guidelines by which the school runs, the purposes being to ensure safety and consistency, thereby maximising the time and opportunity to learn. The school may have policies for learners, such as for equipment and homework; for teachers, such as for lesson organisation; and for leaders, such as high visibility around the school and a focus on teaching and learning
 - 'resources' include time, equipment and accommodation as well as finance
 - 'professional development' in the model is continuous for all staff, in the form of feedback from monitoring of all kinds leading to support, challenge and training.
- d. **Consistency.** As text above shows, headteachers successful in leading both turn-round and sound schools describe how they went about it differently, but frequently with a common core of key ideas. It seems to matter less which of a range of sensible strategies or priorities are chosen than the consistency with which they are pursued. This has been the finding of other studies (Stringfield & Teddlie, 1988; Reynolds, Hopkins & Potter,

2002) leading to the term ‘fidelity of implementation’, borrowed from management theory to describe “the degree to which an intervention or programme is delivered as intended” (Carroll *et al.*, 2007, fig.1). Reynolds confirms the importance of this view, claiming that school effectiveness research

“suggests that gains vary more within improvement projects than between them, and that schools achieve greater gain by pursuing any project thoroughly rather than by choosing one project over another, the ‘fidelity of implementation’ issue” (Reynolds, 2007, p.4).

The proposed model therefore emphasises consistent implementation rather than particular content. Unfortunately, the term ‘consistent’ is both divisive (by sometimes being caricatured as straitjacketing) and lacking in precision. I have written elsewhere (Potter, unpublished EdD assignments) that the term is over-simplistic and needs unpacking. Underpinning the proposed model are five dimensions of consistency:

- A shared culture or adherence to a set of values: what it means to be a member of this school. Most headteacher interviewees referred to the centrality of (re-) establishing a school culture, components of which have been suggested.
- Coherence: decisions, policies and actions should be consistent with the culture, so that the reality is harmonious with the rhetoric. Coherence builds community (staff, pupils’, parents’) trust and confidence.
- Consistently good teaching: the well-researched merit of all teachers adhering to a small number of well-researched characteristics of good teaching (Corallo & McDonald, 2002; Chapman & Allen, 2006), such as those of Rosenshine (2010), Perkins (2008) or a similar set co-constructed within the school.
- Decontextualised consistency: aspects of the school in which different approaches can be equally effective, but in which reducing within-school variation maximises time to teach and learn and builds the confidence of staff and students. These aspects generally concern the organisation of the school and of learning, such as managing students’ behaviour, assessing their work or circulation in corridors. It would be possible to include teaching within this category; I have separated it because of the consistently pre-eminent importance given to teaching quality in all sources, both literature and interviews.

- Continuity or longevity: consistency of culture over a substantial period. Interviewee XX described the stultifying effect on staff morale of frequent changes in direction. James Timpson, CEO of a major service company, asserted in a televised interview that “a culture takes years to develop; a lot of CEOs aren’t around long enough to really develop a culture of complexity that works” (Timpson, 2024). The sustained improvement of YY’s school over a long period is, in the view of the SLT, attributable in large part to the continuity of direction set and reinforced by the long-term stability of the leadership team, all of whom have grown into post in the school. They have built the culture by “*recruiting junior staff and developing good habits*” in them.

e. Intelligent accountability.

YY’s team described a dynamic, three-fold process to ensure continuous improvement; these are the touches on the flywheel described above [p.100] and are represented in the model (Diagram 1) by the arrows on each side:

- monitoring by leaders to ensure that systems are operational and deliver results
- continuous (daily) feedback to individuals and teams from the monitoring, leading to individual professional development
- feedback through the organisation of ideas for incremental changes to bring about continuous improvement.

The Inverse pyramid defines how each layer is accountable to those adjacent to it, and thus provides supporting conditions upwards and feedback downwards, in a system of **intelligent accountability** which drives improvement. Didau’s (2020) work applies the principle of intelligent accountability to schools in a similar way.

The term ‘accountability’ tends to convey a negative slant: ‘being called to account’ suggests that leaders mistrust team members and must police their work. Such an approach fits better in the top-down than the inverted pyramid model.

The inverted pyramid shows accountability is shared: those who lead teams or functions are responsible for how well they work, and senior managers are responsible for providing the conditions under which the team leaders can succeed.

Herzberg's (1959) work on motivation teaches that people are motivated at work by a sense of achievement, by recognition for the work done and by the taking of responsibility. Accountability through performance management has the potential to push these motivator buttons but equal potential to press demotivator buttons like being monitored. Intelligent accountability describes systems at the point of balance, at which the benefits of accountability (recognition, achievement, growth) are gained but the distrustful, policing style is avoided. The two adjectives that describe intelligent accountability systems are *collegial* and *formative*. Here, 'collegial' means mutual and team accountability: the team leader is accountable for performance, but the senior manager is equally accountable for ensuring that the team leader is skilled and resourced to do the job properly. This mutuality builds trust and teamwork. Judgement without support is a poor model of accountability. 'Formative', means developmental, building a sense of professional responsibility or ownership. Professionals are self-evaluating and -regulating, so performance management systems should encourage staff to self-assess and propose their own improvements.

The improvement model includes **external support** for improvement. All headteachers interviewed made use of expertise external to the school, some from within the MAT, some independent. KK was critical of headteachers in the MAT s/he chaired who "*bought in consultants to tell them about their own schools*" and embraced the idea of building expert teams within the MAT to work across schools. HH, RR and CC bought in trusted external consultants to support schools in their MATs, while AA spoke strongly in favour of links with expertise across the LA. Whatever the mechanism, external support brings fresh perspectives informed by experience, helping frame the issues for the school. A thorough survey is provided by Spouse (2001); the topic was heavily researched in the 1990s (Rudduck, 1991; Ainscow & Southworth, 1996), less so in the twenty-first century, with the rise of the MAT.

5.5 THE IMPACT OF FIFTY YEARS OF EFFECTIVENESS AND IMPROVEMENT ACTIVITY

Beginning with an attempt to evaluate the hard evidence for school improvement, this section then highlights two facets which add to existing knowledge:

- analysis of improvement has focused on such ‘technical’ aspects of schooling as attainment and teaching quality, and has largely ignored fundamental or underlying improvement (Findings Theme 4)
- changes to Ofsted’s role in school improvement (Theme 3), especially the significance for improvement of the largely unheralded but pivotal change in practice in 2005. This change lies behind many interviewees’ comments about Ofsted.

5.5.1 HARD EVIDENCE FOR SCHOOL IMPROVEMENT

A YouGov Poll in 2021 showed that 35% of adults believed that UK schools were worse than they had been ten years previously; but that 70% of parents interviewed believed that their children received a good education (YouGov 2021).

Therein lies the problem with seeking a clear answer to the ostensibly simple question “have schools improved?”. The answer is far from straightforward and the question of attributability even more complex: the unknown proportion of any improvement attributable to school improvement (such as better teaching or a better-designed curriculum), to improvements in young people’s ability caused by better health and nutrition, to harder work by pupils, or to grade inflation. Most relevant data pose more questions than answers: for example, whether the 35% of the adults questioned in the YouGov poll were parents with direct experience of schools or random individuals echoing social media, is unknown, nor which aspects of schooling shaped their judgement: whether, for example, they believed standards to be low or the building to be collapsing through faulty concrete.

I sought hard evidence of improvement in the two most promising sources: Ofsted inspection reports and test and examination results. Even here evidence proved elusive.

Ofsted inspection reports deliver an unequivocal judgement as to whether a school is good enough (good or outstanding) or not (requires improvement or inadequate). Ofsted has been

inspecting schools since 1993; most have received at least five inspections, some many more, so changes in quality should be clear.

Recent evidence is that the proportion of schools judged good or better is high and rising: 79% in 2021, 88% in 2022 (Ofsted, 2022, para.1.2) and 90% by late 2023 (Ofsted, 2024). The proportion judged to require special measures seems to indicate a small but significant group of schools resistant to improvement. An early annual report (Ofsted, 1996, p.35) stated that 2% of schools had been placed in special measures; in 2005-06 (Ofsted, 2006, p.24) “2.7% were made subject to special measures, the same proportion as in 2004/05 and smaller than the 3.7% in 2003/04.” The most recent figure is between 3 and 4% (Ofsted, 2023b).

However, the data are slippery, and the degree to which the ‘good schools’ of 1995 and 2020 are similar is open to question.

- Grading has changed: until 1996, schools were judged against a 5-point scale, with grade 3 ‘satisfactory’. From 1996 – 2005, this was replaced by a 7-point scale, grade 4 being ‘satisfactory’; in 2005, the scale was condensed to 4 points, with grade 3 ‘satisfactory’. In 2012, ‘satisfactory’ was changed to ‘requires improvement’. Like Humpty Dumpty¹⁷, Ofsted decides the meaning of words: dictionary definitions of ‘satisfactory’ include “fulfils needs or expectation”, ‘good enough’, but Ofsted decided it means ‘not good enough’. This is not the place to argue the merits, but to note that for 20 years the number of ‘good enough’ categories exceeded ‘not good enough’, but since 2012 they have been equal. What is ‘good’ has therefore changed.
- Definitions have changed: at various times, Ofsted has employed categories such as ‘notice to improve’, ‘serious weaknesses’ and ‘requires improvement’ to describe schools with significant weaknesses not severe enough to require special measures. In her 2005-06 report, HMCI states “some schools that might formerly have been made subject to special measures received notices to improve instead” (Ofsted, 2006a, p.24).
- Quality benchmarks have risen: in 2009 there was a spike in schools falling into special measures, up to 8%. Ofsted and the government made clear that the spike was the result of “the higher expectations of the new framework.” (Ofsted, 2010, p.55). Thus the

¹⁷ “When I use a word,” Humpty Dumpty said in rather a scornful tone, ‘it means just what I choose it to mean, neither more nor less’ “ (Alice in Wonderland).

apparent 'rump' of 3% does not exist: the proportion in special measures rises and falls with the annual sample of schools and with the criteria applicable at the time.

- Annual samples of schools inspected vary: each year HMCI reports the quality of schools inspected in the previous year, and they do not constitute a random sample. Secretary of State Balls made clear (BBC, 2010): "Weaker schools are being specifically inspected more regularly ... to turn them around - while outstanding and good schools are now inspected every five years and satisfactory schools every three". Between 2012 and 2020, Ofsted did not inspect schools judged 'outstanding' in their previous inspection. At no point, not even in 1998 when the first full cycle of inspection was complete, has there been a 'state of the nation' summary.
- Criteria have changed: in Ofsted's 30 years, its *Framework for Inspection* (aspects inspected and the criteria by which they are judged) has undergone at least six revisions, some very substantial. Safeguarding did not enter Ofsted inspection judgements until 2005; from 2007 to 2010, schools were judged on their contribution to community cohesion, not since.
- The institutions inspected have changed: initially secondary schools, then primary and special; later pupil referral units, nurseries and early years settings. Care needs to be taken over which types of institution are included in the total.

I led Ofsted inspections from the first month (September 1993) until 2006. I am convinced that inspectors improved: better versed in the criteria, more thorough and perceptive; our writing more precise, in terms of both compliance and the quality of judgements. These changes in inspectors' accuracy and consistency constitute another, unquantifiable, variable.

The apparent steady rise in the proportion of 'good' schools and the stubborn group of 'failing' schools disappear, to be replaced by fluidity. Having sought precision, I am left with an impression, gleaned from reading the forewords to successive HMCI Annual Reports that, overall, schools are improving. Those forewords adopt a similar, truistic tone: things are getting better but there's a lot left to do. The 2007-08 report is typical: "one thing is clear: for most children and learners, things are getting better [list follows] ... But such undoubted improvements do not mean that all is well ... [list follows]" (Ofsted, 2008, p.5), as is the headline of the 2013 report (Ofsted, 2013b, p.6): "Grounds for optimism – but more needs to be done".

Test/examination results: Interviewees made clear that these provide the ‘default setting’ for government’s evaluation of improvement and thus should provide clear evidence. I refer to three well-established datasets: GCSE and A-levels, and international comparison tests (PISA, TIMSS, PIRLS).

GCSE data appear to show substantial improvement: the proportion of students who achieved the historic benchmark of five or more passes at grade C or above rose from 10.7% in 1954 to over 80% in 2012 (House of Commons Library, 2012, Table 1 p.19). Change over such a long period is mainly attributable to the considerable rise in the proportion of students taking examinations. The proportion achieving the newer benchmark of five or more grades A*- C including English and mathematics rose from 35% in 1996 to nearly 60% by 2012 (Stubbs, 2016; DfE, 2016) but, as Coe (2013) points out, changes in measurement blur the issue: at various times, the results of vocational qualifications and other ‘equivalent’ examinations, such as the International GCSE, were included in the GCSE scores but later excluded. A new grade of A*, indicating exceptional performance, was introduced in 1994; between 2017 and 2020, the system was changed from the eight grades A*-G to the nine grades 9–1 without clear equivalence (Ofqual blog, 2018, table 2). Whereas grade C formerly represented a ‘pass’, now there is ambiguity, grade 4 representing a pass, 5 a good pass.¹⁸

In the face of a near-doubling in the proportion achieving the benchmark, Coe concludes “the question is not whether there has been grade inflation, but how much” (Coe, 2013, p.vii). It is not possible to attribute proportions of the increase to school improvement and to grade inflation, an unknown mix of alleged easier papers, more lenient marking and lowered grade boundaries.

A-level results, which have not been subject to the same rule changes (except the addition of an A* grade in 2010), have improved steadily. In 1993, there was a 75% pass rate; since 2008, the pass rate has exceeded 95%. In 1993, 46% of results were graded A - C; in 2023 grades A*-C accounted for two-thirds of all results (Stubbs 2023). Press debate rages about whether A-levels have got easier, marking more lenient or students work harder; rarely is the

¹⁸ I have avoided reference to results during and immediately after the pandemic, which were subject to different rules.

possibility of school improvement (better teaching, better curriculum design, better leadership of schools and subjects) considered.

Hanging over the debate is the role of the Office of Qualifications and Examinations Regulation (Ofqual), which includes ensuring that grades do not change much year to year: “the adjustment of the grade boundaries so that, overall across England, the proportion of grades achieved is similar to the same subject in previous years” (Tennison, 2020). Schools seek to improve grades which are designed to remain steady.

Neither Ofsted nor examination results therefore provides unequivocal evidence of school improvement. Coe (2013, Table 1, p.iii) maps the results of international comparative studies for England between 1995 and 2011, and concludes that they vary by study, by subject, by students’ age and over time, but overall are roughly stable: “the pattern of results from the different international surveys is actually fairly consistent: not much change between 1995 and 2011” (*ibid*).

I sought evidence for school improvement from the study of teaching quality over time. I have been evaluating lesson quality for over 40 years: the total of my observations must exceed 5000, and I have no doubt that teaching is considerably more engaging, varied and tailored to individuals’ needs than when I began observing, and unrecognisable compared with my experience as a schoolboy. The only large-scale source of data concerning teaching quality is Ofsted inspection. Again, the data are not straightforward: to conclude that teaching had improved would involve demonstrating that outcomes have improved, and that has been shown above to be elusive. Having graded teaching quality in each lesson observed, Ofsted initially published (in HMCI’s annual report) the proportion of *lessons* which achieved each grade. The proportion judged good or better rose rapidly as teachers learned the grade criteria. Ofsted switched to publishing the proportion of *schools* in which teaching was good or better and, in 2014, ceased grading lessons altogether.

In his 1996-97 Annual Report, HMCI Woodhead demonstrated another problem with using Ofsted data: circularity. “Teachers are ... teaching better. In primary schools we are seeing more challenging direct teaching” (Ofsted 2007 p.13): teaching is good when it meets Ofsted criteria or, in this case, HMCI’s preferences (having endlessly stated that they had no

preferred teaching methods). Circularity works in two ways: first, if one criterion for judging teaching is the teacher's planning of lessons, then it is unsurprising that good teaching is characterised by good planning. Second, for many years Ofsted judged the quality of teaching by the results achieved: if output variables such as attendance, examination results and behaviour were good, then the input variables like teaching must also be good; the reverse was also assumed to be the case. Again, this is not to dispute circularity, but to note that it weakens Ofsted as a source of valid, reliable data on teaching quality over time.

My final attempt at hard evidence comes from data on pupils' school attendance. In a House of Commons debate on truancy, an MP described an attendance rate of 85% as "normal or satisfactory" (Hansard, 1971, col.1553). A chart in a parliamentary research paper (House of Commons, 2024, p.14) shows the steady improvement of attendance from 93% in 2006 to 95% in 2012, followed by 8 years at the same figure, since when the pandemic changed patterns. Improvement is undoubted, but whether attributable to better monitoring and sterner penalties, including fines, or to improvement making it both more enjoyable and productive to attend school, cannot be determined.

5.5.2 THE NEGLECT OF UNDERLYING IMPROVEMENT

Having failed to establish hard evidence for improvement, I have brought the autoethnographic dimension of this study into play: I am certain, from nearly 60 years working in them, that schools are better places for young people to live and learn in than in my early days as a teacher, since I led a school, and since I became an inspector. In focusing on hard outcomes, the literature largely neglects fundamental improvements.

Interviewees, conditioned to discuss improvement in terms of hard outcomes, rarely raised underlying improvement spontaneously; however, when prompted ("How has pupils' experience of life in school changed?"; "how well were you taught at school?") they immediately began to reflect on the monotony of their experience of rote teaching (UU, TT, RR); friends and siblings failing the 11 plus and being separated (LL, TT); physical punishment (NN); variability in curriculum within and between schools (BB); pupils as passive recipients (JJ); and the poor treatment of less-able pupils (CC, NN).

In Chapter 1, I set out evidence, anecdotal and eye-catching, to support my belief in schools as incomparably better places; I now pursue this theme in a more analytical way.

Attitudes to young people in school

Foucault argues that “schools serve the same social functions as prisons and mental institutions - to define, classify, control, and regulate people” (Foucault, 1975, p317). Although their insistence on rules, good order and uniform may support Foucault’s view of their control function, schools have, over the research period, become more humane places. They may play the social function of a prison but are not prison-like. Legislation has changed issues of access and human rights, and pupils are treated more as participants, less as recipients.

There is now a pervading concern for young people’s welfare and wellbeing in school, enshrined in law. Unkind individuals will, sadly, always exist, but the system’s callous disregard described by Justin Webb [p.1] has largely been driven out. Corporal punishment, the deliberate infliction of pain using a variety of implements, had been a method of discipline in British schools for centuries until, following a European Court of Human Rights decision (1982), all forms of corporal punishment were banned in England in 1986; no serious discussion of its restoration has taken place since.

The pupil’s voice is heard more clearly than was previously the case. Article 12 of the United Nations Convention on the Rights of the Child (1989) states that young people should have a say in decisions that affect them. Most schools have some kind of student council, advocating for people and issues, organising events, communicating between staff and students, and promoting leadership development and political education. It is routine for student councils to play a role in the appointment of staff including the headteacher (DfE, 2023, para. 228), and to present issues to governors.

My own experience was that more parents approached me and my school in anger or in awe than did so in partnership. Although the first two emotions have by no means disappeared, in recent years I routinely observed parents and schools working together in the best interests of pupil and family.

JJ said: *“Today’s adolescents are welcomed at the gate by staff who enjoy the relationship they have with them. We are at ease with our students ... and they with us.”* I have stood with senior staff at the gates of a dozen schools, welcoming pupils at the start of the day or bidding farewell at the end. Of course this plays a monitoring function, checking behaviour and appearance but, in every school, I have seen cordiality on both sides, with several pupils taking the chance for a confidential chat.

Access and inclusion

The history of English education is one of increasing participation, from a privilege of the rich to a right for everyone. During the research period, “It wasn’t until 1970 that disabled children – previously classed as being unable to be educated – were given a right to participate in education.” (Mencap, n.d.). Only in the 1970s did children with severe learning difficulties become the responsibility of education; previously deemed ineducable, they had been the responsibility of health services. The school leaving age rose to 16 and then, effectively, to 18. The proportion of young people aged 17 in full-time education had risen to nearly 92% by 2018 (DfE, 2023b), and that attending university from around 3% in the early 1960s to 14% by 1980, to September 2019, when the BBC website headline reported “The symbolic target of 50% at university reached.”

Inclusion means more than participation: the system’s widespread disregard for less-able pupils can be evidenced by qualifications: GCE O level, targeted at the top 20% of the ability range, roughly equivalent to the grammar school population, was introduced in 1951; only one in ten students in secondary modern schools took any O levels (Gillard, 2008). CSE was introduced in 1965 to enable the next 40% of the ability range to leave school with some qualifications. Nothing was designed for the lowest 40% of the ability range, with all the inaccuracies and potential biases that figure represents. From 1986, O levels and CSEs were merged into a single examination, GCSE. In 2022, around 94% of all 16-year-olds took at least five GCSE examinations (Ofqual, 2022).

At the time of writing (spring 2024) the burgeoning demand for tailored provision for pupils with SEND is greatly outstripping supply and resources. Ironically, these often unresolved, bitter disputes between families with needy children and authorities with inadequate

resources evidence the greater concern for these pupils, whose needs were, in the past, frequently disregarded (NN, PP, CC).

A guarantee of curriculum breadth and depth

Although several interviewees criticised its scale (LL) and “*over-engineering*” (HH), the advent of a National Curriculum provided an entitlement for every pupil: subjects, content and yearly progression, with a measure of continuity should they change school. The randomness of the previous situation was indefensible, especially in primary schools, where the curriculum in subjects such as physical science and technology was, at best, inconsistent within and between schools (JJ, BB).

The secondary school technology curriculum was, until the 1970s, usually gender-determined, with male students given wood- or metalwork, females home economics or needlework. My school removed that now unthinkable gender barrier in 1981. I believe schools now show genuine commitment to gender equality of opportunity, with girls outperforming boys on many measures. This may be linked with increased female representation in secondary school leadership; according to Coleman (2002), women made up just under 30% of secondary school headship at the millennium, while DfE figures for 2018 show the proportion to have risen to close to 40%. In 2018, nearly half of deputy heads, and over half of assistant heads, were women (DfE, 2018 Tables 2.8 and 2.9, p.32).

Teaching

In the Introduction, I gave vignettes of how I was taught. Interviewees, and not only older members, described the predominance of copying from the board or from dictation, and answering questions from the textbook (TT, UU, HH, NN, JJ). In seven years of secondary education, I was never asked to carry out a classroom task with another student: tasks were conducted individually and in silence. The proportion of lesson time spent in pair and group work, in discussion, collaboration and research, is now much higher.

Much of current teaching literature concerns meeting the needs of individuals. All the teaching I received, and much of that I observed as a school leader and adviser, was undifferentiated, whole-class teaching. Only in the last 30 years has teaching geared to individuals and groups within the class become commonplace.

The feedback pupils receive on their work is greatly improved, from a mark and a judgement to dialogue about strengths, weaknesses and how to improve. Carol Atherton, lifelong English teacher, describes getting an essay back when a pupil in the 1980s:

“mine had ticks in red ink, a mark of 18 out of 20 and a short comment: *V. good*. I had no idea why it got 18 out of 20 and not 17 or 19, or what I would have needed to do in order to improve it. It didn’t occur to me to ask, either. *We didn’t have those kinds of conversation back then*” [my emphasis] (Atherton, 2024, p.10).

The significance of this quotation lies not only in how it evidences frequent lack of thoroughness and helpfulness in the assessment of pupils’ work, but also in the last, italicised sentence: discussion and questioning did not routinely form a part of teaching, even in the 1980s.

A previously unidentified corollary of the ‘schools make a difference’ paradigm is the greater responsibility that teachers and schools now take for pupils’ results of all kinds, especially academic. By showing us that schools have a significant impact on pupils’ achievement, Rutter and colleagues changed teachers’ mindsets about examination results. A new headteacher joined a school in my LA: in his discussions of A-level results with each subject leader, the head of a major subject explained the poor performance of his department by saying “sir, the department doesn’t ‘perform’, the pupils do”. This attitude has almost disappeared in the last 40 years.

The criteria for judging teaching included (Ofsted, 2006) *inter alia* teachers’ subject knowledge, expectations of pupils, planning of lessons to deepen pupils’ knowledge and understanding, skilful questioning to re-shape learning, the creation of a positive climate for learning. In this demanding context “Teaching and learning are good or outstanding in 61% of all schools satisfactory in 35% ... inadequate in very few” (Ofsted, 2007, p.30).

Leadership and the potential for school improvement

Miles’s (1967) work on ‘organisational health’ in English schools, and their unpreparedness for change has been quoted [p.32]; the book was published in my second year of teaching and described precisely the culture of my school, one of professional isolation inimical to the whole-school priorities necessary for school improvement.

Before the arrival of the National Professional Qualification for Headship (NPQH) in the 1990s, there was no approved programme to develop leadership skills for headship: senior leaders learned mainly by apprenticeship. By 2020, there was a joined-up professional development framework for schools: eight National Professional Qualifications (NPQs) for aspects of senior leadership in schools, including the leadership of specialist areas like teaching and literacy; together with formal sets of standards for teachers and headteachers, an induction and promotion structure and an annual appraisal scheme for all. This stands in stark contrast to an interviewee's (NN) account of becoming a school leader [p.2] and the mistakes he made.

I accept that some of the above may be liable to the charge of rose-tinted spectacles, and that there will be dispute as to whether some of the changes constitute 'improvement'. It is arguable whether, for example, 50% of school leavers truly benefit from university. Provision for pupils with SEND remains inadequate; the National Curriculum is widely criticised for excess and over-specification; there are still disproportionately few women in secondary school leadership; attempting to meet the needs of 28 individuals, five hours per day, 190 days per year is beyond the ability of mere humans; the gaps in attainment between disadvantaged children and their peers, and between different ethnic groups, are intolerable.

But there can be no doubt that the 14% who went to university in the early 1980s represented a huge waste of talent, as did the exclusion of many pupils with learning difficulties and disabilities from education; nor that the great variations in curriculum within and between schools, and the bias in outcomes in favour of more affluent, white students, were unacceptable.

Neither is it the case that everything has improved: I would cite three aspects:

- the extra-curricular diet was arguably richer when I was a young teacher, with abundant sports fixtures and practices, drama, art and musical performances, trips, visits and societies. The sustained increase in teachers' workloads has led to some activities being curtailed, and aspects of health and safety have limited some activities
- since the pandemic in particular, pupils' attendance and behaviour have deteriorated. Teachers' professional associations have conducted surveys, with alarming results, one reporting that:

“37% of respondents have experienced physical abuse or violence from pupils in the last 12 months, while 90% report verbal abuse or violence from pupils ... Being sworn at is cited by 81%” (NASUWT, 2023)

- two interviewees (LL, WW) challenged the assertion that schools are better, happier places for young people. When pressed to say why, they cited societal ills such as drugs and the pressures of social media and cyber-bullying, including sexual harassment. When a website (‘Everyone’s Invited’) was established on which women and girls could record incidents of misogyny and harassment, nearly 3000 UK schools and colleges were cited. When Ofsted investigated the issue, their study revealed:

“how prevalent sexual harassment and online sexual abuse are for children and young people ... for some children, incidents are so commonplace that they see no point in reporting them” (Ofsted, 2021).

WW and LL also called attention to the current mental health crisis affecting schools. Future research will undoubtedly illuminate the scale, duration and causes of some of these problems, and how they link with the pandemic.

5.5.3 CHANGES IN OFSTED’S ROLE IN SCHOOL IMPROVEMENT

From its inception, Ofsted placed school improvement at the heart of its work, its initial strapline being “Improvement through Inspection”. The organisation persists with this priority: “We are a force for improvement” (Ofsted, 2022, p.4); “We are required to carry out our work in ways that encourage the services we inspect and regulate to improve” (Ofsted, 2023). In its response to a House of Commons investigation into its work, Ofsted stated: “We are committed to doing good as we go” (House of Commons Committee Special Report, March 2024), a phrase described by interviewee GH as “*a powerful mantra for inspectors*” and one s/he was taught on joining HMI.

I contend that much of Ofsted’s potential to support improvement has been lost. Ofsted dropped ‘improvement’ from its strapline in the early 2000s, changing it to “Better Education and Care” (Ofsted, 2006, front cover) and then to “Raising Standards, Improving Lives” (Ofsted, 2007, front cover).

Interviewees identified four positive impacts of Ofsted inspection on school improvement:

- Ofsted added compulsion to the drive for improvement, mentioned by several interviewees: HH, FF, and especially SS: *“Only Ofsted really improved schools as a whole system. How? By definition [quality standards] and by strong-arm tactics.”* ERA had established a market-force-driven education system; the customers (parents) needed information about schools’ quality, in a form they could compare, hence the publication of league tables and Ofsted inspection reports, both from 1993.
- The first Ofsted Handbook (1992), the ‘buff folder’, was widely praised for being *“the most effective school improvement manual of its generation”* (NN, also HH, JJ *et al.*): by separating outcomes from contributory factors, putting the latter into a logical structure and establishing the criteria by which each would be judged, Ofsted effectively defined and structured school quality.
- In the early days of inspection, several profoundly poor schools were identified: most had their headship and governance replaced, some federated with strong schools, others were closed. All six interviewees who had been Ofsted inspectors referred to this as a positive.
- The 2019 revision to the inspection framework was welcomed, especially by interviewees currently in headship, for reducing the dominance of attainment data in the overall judgement of the school, focusing more on how the curriculum is organised and delivered.

A government-commissioned review of Ofsted’s work by NFER in 2009 found that:

“Schools perceive Ofsted inspections to have a positive, if mainly indirect, impact on school effectiveness and improvement The main contribution of the inspection process is seen to be in confirming, prioritising and clarifying areas of improvement, rather than in identifying new areas to address.” (House of Commons, 2010, Executive Summary)

This seems scant justification for such a costly (in every sense) system.

Interviewees saw the impact of inspection on the improvement of their schools not as direct but because they shaped their practice and priorities to Ofsted’s changing emphases.

The most profound change to Ofsted practice came in 2005 and received almost no attention in the literature; when it is mentioned, the significance of the change is not captured. Until 2005, the length of an inspection, measured in inspector days, was determined by Ofsted

based on the type and size of the school (Ofsted, 1992, Table p.9). An average-sized secondary school of 1000 students attracted 49 days of inspection, an average primary around 20. Lead inspectors deployed their forces to discharge the requirements of the inspection schedule in the given time. Most conducted a secondary school inspection over four or five days; a core team of about three including the lead inspector stayed throughout, inspected aspects such as leadership and finance and collated the team's evidence about teaching and the curriculum. A group of subject inspectors attended for one to three days each to inspect English, history, science and so on. Each inspector had to be endorsed by Ofsted to inspect each subject s/he covered.

Daily during their period in school, subject inspectors met with their head of subject and, at the end of it, gave feedback to her/him, normally accompanied by a member of the school leadership team. Teachers could, and usually did, request individual feedback on each lesson observed. The lead inspector met at least daily with the headteacher to brief him/her about emerging issues and questions and, at the end of the inspection, gave detailed feedback to the leadership team and subsequently to governors. Reports were substantial, detailed documents, and a brief summary was prepared for parents. Practice varied between teams, but most of the above was required by Ofsted (1992, p.6 *et seq.*).

Thus until 2005, Ofsted inspections were detailed and expert, involved extensive professional dialogue with the school and a detailed report with recommendations for improvement to which the school had to respond with an action plan, also sent to parents. The potential for supporting school improvement was therefore extensive.

In 2005, the Ofsted methodology changed radically: "a major programme of change that will allow us to meet our efficiency targets by the end of March 2008" (Ofsted Departmental Report, 2006b, p.1). The motivation for the change was financial and to reduce the burden on schools which, especially in small schools and departments, was considerable.

The NFER report to parliament judged that:

"the main changes [in 2005] were: less advance notice of the inspections (usually two days), more frequent inspections, smaller inspection teams, increased emphasis on the school's own self-evaluation evidence and shorter reports with fewer, clearer recommendations for improvement" (House of Commons, 2010).

All true but missing the point: that the secondary school that previously attracted 49 inspector days now received four to six: two or three inspectors over two days. And not two full days, since inspectors withdrew on the afternoon of the second day to consider their findings, fed back at the end of that day. While small tweaks to the number of inspector days could be managed, a cut from 49 to 5 is not the same job done more efficiently, it is clearly **a different job**.

From 2006, inspectors became generic, licensed to inspect any school or subject. Individual subjects were no longer inspected, although in 2019, Ofsted instigated 'deep dives' into subjects. An interviewee's school, inspected in 2024, received five 'deep dives', carried out by three inspectors in the one-and-a-half days of inspection. The 'depth' of these dives is questionable when conducted by non-specialists in an already crowded inspection schedule.

I contend that post-2006 inspections, though undoubtedly less burdensome, lost the specialist professional dialogue with teachers, senior and middle leaders that provided both the expertise and the space for supporting improvement. Replacing this dialogue is a brief (in a small primary school, perhaps only one inspector for a day and a half), superficial treatment resulting in a report written in language designed for parents: inspectors are specifically enjoined not to use the language of education but parent-friendly language: the first heading in a report containing only a couple of sides is "What is it like to attend this school?", information clearly targeted at parents. The guide to writing reports (Ofsted, 2014, p.9) makes frequent reference to the parent audience: "Parents need to know if this is a school they can send their children to ... If you answer these questions in the report in a way that a lay person can understand...".

Whatever functions Ofsted may now serve, its role in improvement has clearly shrunk significantly. The sharp reduction in the number of observations carried out too easily means that the report's conclusions are heavily influenced by "*a single event which easily becomes interpreted as a pattern*" (JJ). This is not a climate in which professional dialogue supporting school improvement is likely to flourish.

Not only did the 2006 changes diminish Ofsted's role in improving individual schools, they also reduced inspection's role in promoting research into effectiveness and improvement. Previous inspections generated a mass of supporting data, including numerical judgements of

every aspect of the school and each subject in it, so-called 'judgement recording forms'. These provided huge databases within which to correlate variables, providing insights into such questions as: "which factors have the greatest influence on achievement?". No such ability exists within subsequent frameworks.

On the day I write (25th April 2024), in the face of considerable opposition including from the all-party education select committee, Ofsted is to retain its one-word grading system for schools¹⁹, arguing that "the grades, such as 'outstanding' or 'inadequate', had significant benefits for parents choosing a local school for their children" (Local Government website, 25th April 2024). Parents, not schools, are the intended audience.

¹⁹ Postscript: in September 2024, a new government reversed the decision about Ofsted grading and initiated a school report card. However, its purpose remains "to provide parents with a complete picture of how schools are performing" (Government Education Hub blog, 2nd Sept.2024).

CHAPTER 6: CONCLUSION

The publication of Rutter's *Fifteen Thousand Hours* represented a paradigm shift, a watershed in schooling, demonstrating that it has a significant effect on outcomes for pupils. It is a measure of the rapidity of change that in 1979 this finding ran counter to the prevailing sociological *zeitgeist* but now feels self-evident, almost banal.

Although the term 'school improvement' had yet to be coined, it was evident to interviewee NN, a rising school leader, and to me, being interviewed for a deputy headship in 1980, that if we did more of the things Rutter cited, the outcomes for our pupils were likely to improve. This research confirmed our agency and, over the following few years, gave an increasingly refined understanding of what we should do to bring improvement about.

My life since gaining that deputy headship has been devoted to promoting school improvement. As the period of intense school improvement activity subsided and I began to reduce work toward retirement, I had time to reflect on the evolution of that work. Reflections fell into five groups, which became my research questions. They are re-presented below with a summary of the key findings and possible implications of each; since this thesis is part of a *professional* doctorate, I have tried to put myself back into school or LA leadership and indicate where the findings would help practice.

6.1 TO TEST THE EXTENT TO WHICH SCHOOL IMPROVEMENT DOMINATED EDUCATIONAL THINKING AND ACTION IN THE PERIOD

The evidence of both literature and interviews supported the view that issues of improvement did indeed dominate, but the interviews added nuance, mainly by demonstrating the shift in focus in the forty years since Rutter: his impact was phased, a phenomenon not described in the literature, which is clear that there was an immediate surge in academic research into effectiveness and, slightly later, improvement. Brighouse (Brighouse & Waters, 2021), then CEO in Oxfordshire, and the small number of interviewees (plus the author) who were active in education at the time agreed that, although they had been excited by Rutter's message, it and the Great Education Debate launched in 1976 resulted in "*very little happen[ing] until the*

mid-80s” (JJ), but that the following period was characterised by an intense drive for improvement.

This period of wholesale change in schools began when ministers, guided first by political ideology and, as KK indicated [pp.80-81], later by the persuasiveness of effectiveness and improvement research, launched a legislative programme for education that drove the period of true dominance of the improvement agenda in schools between the late 1980s and the early 2010s. Almost every interviewee referred to the impact of the Education Act, 1988, and all highlighted that of Ofsted inspection (Education Act, 1992) on driving improvement. Most credited Ofsted with the lead role, CC pointing out that headteachers may not have effectiveness research at their fingertips *“but all know the details of the Ofsted framework”*, while PP and SS made clear that *“only Ofsted brought about real change”*.

Legislative and quasi-legislative changes (such as the detail of the National Curriculum and its tests, and various iterations of the Ofsted Framework for Inspection) were near-continuous through the 1990s and 2000s; interviewees repeatedly described the pace of change and the pressure for improvement as *‘relentless’* or *‘remorseless’*.

In 2010 – 11, the Coalition government removed many of the structures put in place by Labour to support improvement (such as the requirement for schools to have a SIP and a SEF), after which the pace of legislative change slowed. However, headteachers described continuing pressure to improve; many of them were part of a MAT, and the Education Act 2010 focused on them as the means to bring improvement about. Some heads in the sample (YY, XX, AA) were not part of a MAT but nonetheless described the same pressure from the LA and especially Ofsted as the driving forces, together with ambition for their pupils.

Although legislation and guidance concerning pupils’ welfare and safeguarding had existed since the late nineteenth century, their significance increased considerably in the twenty-first; latterly they have come to sit alongside academic achievement as the major focus of schools’ attention, HH describing *“the monster safeguarding”*. Current headteachers and MAT leaders stressed that safeguarding and welfare dominated the budget as well as schools’ attention. This thesis asserts that welfare and safeguarding have become embraced within the definition of school improvement, avoiding an unhelpful dichotomy like that between ‘curriculum’ and ‘pastoral’ issues in schools. Most interviewees were clear that *“unless young people are well*

fed, listened to, happy and secure in school, their propensity and ability to learn must be limited” (WW).

6.2 TO ASCERTAIN WHY THIS HAS THIS BEEN THE CASE: WHAT COMBINATION OF EXTERNAL (SOCIO-ECONO-POLITICAL) AND INTERNAL (EDUCATIONAL) FACTORS LED EDUCATION TO BE DOMINATED BY SCHOOL IMPROVEMENT DURING THIS 50-YEAR PERIOD?

The literature review presented evidence of a considerable range of factors leading to the period of intense change and drive for improvement described above. Interviews reinforced the public disquiet concerning education in the 1970s, with references to the Ruskin speech, the Black Papers and William Tyndale from those with long educational memories (NN, JJ in particular). Two interviewees (AA, LL) linked the disquiet with the transition from grammar and modern to comprehensive schools, while JJ linked it to the 1970s foment and impassioned debate about the social and economic fabric of Britain. S/he also contributed a set of social and economic factors driving school improvement, such as changes in patterns of employment and the rise of an aspirational middle class leading to the demand for better-qualified school leavers. Some of these points, such as the rise of meritocracy and the impact of ROSLA, are not to be found in the literature; s/he and I discussed the significance of family size and child-centredness, also not found in the effectiveness literature but nonetheless factors contributory to an increased societal focus on schooling.

This set of findings contributed three facets to knowledge of the evolution of school improvement in England:

- it brought the lists of factors promoting the drive for improvement, which exist piecemeal across numerous sources, together in one place, and extended the list as described
- my notes of a discussion with JJ in March 2024, during which we listed factors, read: “Basis for a classification? Internal (educational) factors, external (societal)?”. SE/SI research = enabling factors”. Hence the following, original classification of the factors leading to educational change into three inter-acting sets.

Internal factors: from within education, such as Tyndale, and the Black Papers

External factors: from society beyond the school:

- social factors: changes in family size and structure; the decline of deference and the rise of youth culture; increased affluence and ambition

- political factors: the breaking of the post-war consensus, the rise of neoliberalism and the application of market economics to social institutions including education
- economic factors: changes in employment patterns leading to the demand for a better-qualified workforce; the quality movement (TQM).

Enabling factors: research showing first that schools make a difference to pupils' outcomes, and subsequently what good schools do/have/are, making improvement conceptually straightforward.

- Concatenation among the factors, how they played out across a relatively short period of time, together creating irresistible pressures for change and improvement.

6.3 TO ANALYSE AND EXPLAIN THE RAPID EVOLUTION OF THE SCHOOL IMPROVEMENT PARADIGM, IDENTIFYING HOW IT RELATED TO THE PRE-EXISTING SCHOOL EFFECTIVENESS PARADIGM.

The second question sought to identify and classify the factors leading to the drive for improvement; the third attempts to explain why effectiveness and improvement evolved so rapidly, and the relationship between them.

Chapter 1 described the burgeoning of improvement activity as remarkable for its pace and breadth, commenting on the speed with which research and legislation related to improvement spread geographically and through different types of school.

Literature, and to some extent interviews, makes clear that effectiveness is a thoroughly researched concept leading to a broadly agreed set of correlates. Interviewees, especially those who had worked to support school improvement, referred frequently to key researchers in the field, to the features of effective schools and how they sought to propagate them. The relentless focus on classrooms and teaching described by many interviewees, most notably FF, PP, LL, HJ and YY, and the pursuit of consistent values emphasised by many (HH, BB, HJ, EE) are good examples, confirming effectiveness as a research paradigm, and its inception (1979) as a paradigm shift.

A list of 'what makes a good school' characteristics translates readily into a 'to do list' for improvement. Improvement research began as a series of case studies of the application of effectiveness principles, but time has given it the ability to mature into a well-researched

field. This thesis argues that effectiveness and improvement should not be identified as separate paradigms, but as a twin or joint paradigm.

The literature remarks on the speed of propagation of effectiveness/improvement activity but does not account for it. Discussion with JJ and HJ of the post-paradigm explosion of research, led to the proposal that the rapidity can be explained by likening it to the adaptive radiation observed in evolutionary biology following the opening of a new ecological niche. As lifeforms evolve rapidly to exploit opportunities created by the new environment, so research and legislation burgeon to create new knowledge and embed it in the context of the new paradigm.

6.4 TO IDENTIFY WHAT KEY GROUPS OF EDUCATORS HAVE UNDERSTOOD BY 'IMPROVEMENT', AND THE ACTIONS THEY HAVE TAKEN AS A RESULT.

Interviewees made clear that the meaning of improvement evolved and expanded over the forty plus years of its existence. Rutter, in his analysis of effectiveness, embraced pupils' personal, as well as their cognitive, development, but headteachers frequently referred to examination results as governments' default measure of effectiveness in the 1990s. Shifts in the technology of test data meant that, over the 1990s and 2000s, headteachers described how raw attainment, value-added, contextual value-added and progress were emphasised in turn to capture effectiveness and improvement in ways fairer to more pupils. Interviewees described how they were led to slice the data to reveal the differential attainment and progress of groups within the cohort: first, gender and ethnic groups and then, especially from 2010, between disadvantaged pupils and their peers. The 'disadvantage gap' is now perhaps the key theme in school and government thinking. A focus on sub-groups of challenged students, such as those in alternative provision or who are young carers, may be expected.

Many interviewees identified the *Every Child Matters* Green Paper (DCSF, 2003) and the tragic deaths that led up to it, as a validation of the personal development strand of improvement that may have taken a back seat in government priority, but which had never been lost from schools' priorities. Previous text has described how welfare and safeguarding have first been embraced, and then risen to the top, within the improvement agenda. Links between the priority given to welfare with that on the low attainment of disadvantaged pupils are clear.

Current headteachers interviewed pointed to the dramatic impact of the pandemic on both priorities: the stresses placed on pupils and families by lockdown, illness, loss of relatives and isolation on their mental health and wellbeing; and the adverse effect of online learning on disadvantaged pupils. There was genuine alarm among some interviewees at the further expansion of what constitutes improvement in the wake of the pandemic, and its embedding in headteachers' roles.

This analysis of the evolution of the meaning of school improvement is new to the literature, but the greatest weight of interview evidence concerned how improvement takes place:

*"School improvement works by good **leaders** applying the **principles** of school effectiveness (KK)." "School improvement is about having **systems** built around the right **people**" (DD) (my added emphasis).*

Interview data show frequent allusion to conditions for improvement. Interviewees delved into the 'how' of unspecific phrases in the effectiveness literature like 'purposeful leadership' and 'consistency among teachers' (Mortimore *et al.*, 1988) and spelled out those conditions under the four headings highlighted in the data extracts above.

The four-fold classification of conditions (leadership, principles, systems, people) shapes an original approach to improvement. Since interviewees took the view that conditions are not spontaneous but created, they described the key to improvement as the leadership that shapes them. Interview data are therefore used to classify key leadership characteristics and, in the 'systems' sub-section, to elaborate the idea of *consistency*. This has been a contested idea within school improvement: failure to define the concept has led to its being caricatured as straitjacketing. The text proposes a new classification of consistency and a definition of the contexts within which it is important, such as those in challenging circumstances and those at different points on the improvement trajectory. Interviewees' views on the importance of external support, including that from Ofsted, are summarised.

A new school improvement model is offered, based on the research data, with five components:

- the expanded list and classification of leadership skills for improvement
- the seesaw model to describe the isolation of new headteachers in struggling schools and how, by a combination of offloading the 'heavyweights' from the negative end while

building capacity at the positive, they caused intermediate bodies to slide toward the positive. This is a model I have toyed with for many years, but the moment in the interview with TT when s/he described shedding and developing staff and thus **achieving critical mass buy-in** was an epiphany marked with the word “SEESAW!!!” in my notes

- the Collins flywheel, to describe how large inputs of energy are needed to get an organisation moving (transformational change) but, once moving rapidly, light touches (incremental change) keep it moving. This process is matched against the conventional four-stage model, which focuses on the turn-round of weak schools and less on sustaining improvement. A conversation with YY, head of a school which has shown sustained improvement over a long period, called to mind the Collins flywheel; YY spoke of “*constant fine-tuning of the system, constant nudges*” to keep improvement happening. Here, my interview notes have the word “FLYWHEEL” in the margin
- the inverted pyramid model, which inverts the conventional hierarchic organogram to propose that each ‘tier’ supports the one above, support taking the form of the conditions to enable the latter to do its job most effectively. The job of a pupil, at the top of the pyramid as the system’s ultimate beneficiary, is to learn as effectively as possible, and thus that of the teacher is to create the conditions in which s/he can do so. Conditions are defined as vision, policy, resources and professional development (feedback, support, challenge)
- intelligent accountability, to demonstrate how feedback flows through the pyramid to involve members of the school community in refining improvement; while feed forward provides support, guidance and professional development through the organisation.

Whereas the findings of the first three questions concern establishing a clear account of the evolution of school improvement, this question provides practical guidance for leadership. Understanding that there are conditions for improvement, what they are and how they interact as components of the model would have given shape and clarity to improvement activity in my school. Just like the headteacher who used TQM to transform the school [p.88], the ‘conditions’ notion would have shaped in-school professional development for classroom staff and leaders, and the inverted pyramid would have framed the school culture. For a team of LA school improvers, the model would be a toolkit with which to guide school leaders.

6.5 TO IDENTIFY THE IMPACT OF FIFTY YEARS OF SCHOOL IMPROVEMENT, AND IMPLICATIONS FOR THE FUTURE.

This research attempts to answer the ostensibly simple question “have schools improved?” but is thwarted by elusive data:

- examination grades which are managed by government to remain stable, and which are simultaneously subject to disputed levels of inflation
- Ofsted inspection reports whose criteria and gradings have changed repeatedly; inspections which sample weak schools more frequently than strong
- measures of teaching quality which are either highly subjective or unhelpfully circular (the only effective judge of teaching quality is the results achieved by pupils; hence if outcomes are not improving, neither is teaching).

As well as analysing the slipperiness of data, the thesis offers an analysis of improvement to be found piecemeal but not comprehensively elsewhere, and adds two original points:

- the low priority accorded in the literature, and indeed in the interviews, to what I have termed underlying or background improvement. Interviewees, having defined school improvement broadly, embracing both a personal and a cognitive component, went on to focus almost exclusively on what they described as the government’s “*default setting*” (UU and others): examination results. However, when prompted (“how were you taught when you were at school?”), they launched into an impassioned critique of rote learning, mass medication, narrow curriculum and physical punishment. The research offers an analysis of ways in which schools are better places to learn in than they were, without being blind to some ways in which they are less happy, for example the pressure of social media, cyber-bullying and sexual harassment
- the marked reduction in the improvement role played by Ofsted since the profound, largely unheralded, change in inspection arrangements in 2005-06. Inspections before this date were long, detailed and expert, with compulsory, extensive feedback to teachers, team and senior leaders. The huge cut in the number of days mandated to inspections, carried out over a day and a half by inspectors who are not specialists, limits or precludes the feedback and discussion which promote improvement. Likewise, the detailed reports which followed earlier inspections were designed to help improvement; modern reports are brief, in lay language and targeted at parents and their choice of

school for their child. The current (September 2024) debate about Ofsted grading centres on clarity for parents, not on what would help the school.

The research also reveals that the 2006 change removed much of Ofsted's data for research into effectiveness and improvement.

The point about Ofsted was made in interviews by inspectors who worked in both pre- and post-2006 regimes (five, including the author), and to a lesser degree by headteachers. The latter, while critical of the lack of Ofsted support to improvement, nonetheless understandably favoured the less intrusive, current version.

I have separated the last part of this question (implications for the future) into an Epilogue.

6.6 UNEXPECTED FINDINGS

Most research findings confirmed and extended previous knowledge and expectations. Some research findings were, however, unexpected:

- several interviewees raised the tension between the pursuit of the 'good school' or 'good education' and that of school improvement. As discussion on [p.130] shows, practitioners arrived at this tension from different perspectives, but the resulting debate has genuine philosophical merit. For the purposes of the improvement model, this research proposes that the 'turn-round' phases of improvement should be interpreted as the pursuit of *improvement*, whereas the 'sustain' phase should be considered the pursuit of *excellence*, the 'good education'
- I assumed that interviewees, school leaders in particular, would be anxious to highlight the improvement in schools as humane, dignified and enjoyable places in which to learn and to teach, what I have called 'underlying improvement'. However, with very few exceptions, they did not do so unless prompted, at which point they gave vivid descriptions of the monotony of teaching, the physical punishment and the neglect of individual learning needs which characterised earlier times. The 'blinkers' which interviewees wore, constraining their view to narrow interpretations of outcomes, were a surprise
- although I knew that hard evidence of the impact of 50 years of school improvement activity would be elusive, I expected to be able to produce more and better evidence of

improvement than proved possible. Ironically, the best evidence comes from the neglected theme of ‘underlying improvement’ described in the previous paragraph.

6.7 SUMMARY OF CONTRIBUTION TO KNOWLEDGE AND PRACTICE

This research contributes to knowledge and practice in three main areas.

1. The evolution of school improvement

Existing accounts are *partial*, not enumerating the full range of factors involved; and *patchy*, in that the factors are scattered across the literature. Not only does the account in this thesis extend the range of relevant factors, it also gathers them together, classifies them into push, pull and enabling factors, and shows how they worked together to promote the expansion of school improvement activity in the last quarter of the twentieth century.

The account also proffers an explanation for the rapidity with which school improvement activity spread, both geographically and through different types of school, and highlights the neglect of underlying improvement in schools.

2. Ofsted’s role in school improvement

The research shines a light on the significant erosion of Ofsted’s focus on support for improvement since radical changes made in 2005-06, a change largely unreported in the literature. Detailed, expert-led inspection, providing extensive developmental feedback including a comprehensive report, has been replaced by brief inspection conducted by generalists, resulting in a short report, the target audience for which is parents.

3. The leadership of school improvement

Existing models have two main limitations: they focus on the turn-round of weak schools; and they are simplified to a point which makes them descriptive rather than analytical guides to practice. This thesis offers a new model focused on the *leadership* of school improvement through an analysis of the skills required. The model acknowledges complexity, taking account of aspects of school context and focusing on sustained

improvement rather than rapid turn-round, and is based on the views and lived experience of practitioners.

All three points above represent contributions to *knowledge*, the third also has the potential to contribute to *practice*. Subjecting the school improvement model to critical evaluation and refinement (as described on [p.177]) has the potential to create a set of tools of real value to school leaders and those advising them.

Some of the academic discussion about context-specificity is unprofitable: that principles must be tailored to unique contexts and internal dynamics is self-evident and does not undermine the validity of those principles.

The richness of the evidence from interviews highlights the importance of practitioner perspectives in shaping effective and sustainable school improvement initiatives.

6.8 LIMITATIONS OF THE RESEARCH

Chapter 1 recognised three limitations:

- those inherent in an inductive, interpretivist study: conclusions drawn from a modest sample
- a set of interviewees the great majority of whom are known to me, threatening confirmation bias
- although referenced globally, a study based in the English education system.

The Methodology chapter explains how I sought to mitigate the risks under the first two of these headings.

Having conducted the study, I would add two more.

- The diversity of the sample of interviewees, which is skewed toward males, and intensely skewed ethnically, with only one of the 32 interviewees from a Black, Asian or Minority Ethnic (BAME) background, which significantly under-represents the population at large. The study also draws more heavily on respondents from the south-centre of the country than the north. I cannot know whether more representative gender, ethnic or geographic

samples would have changed the findings, but I am conscious of the discourtesy implicit in the imbalance.

- My concerns over the appropriateness of autoethnography to the study led to my limiting its role to validation and triangulation.

6.9 FUTURE RESEARCH

The limitations above suggest the potential value of replicating the work with a larger, more diverse sample to test the durability of the findings.

To realise the potential value of the research to praxis, I intend to take the section concerning the Model for School Improvement (5.4, pp.125 *et seq*) and, with the guidance and support of more experienced authors, refine and polish the text and present it to a peer-reviewed journal, so that it can be subjected to the criticism of the research community. If it survives this review, then to refine it further and share with agencies such as the Education Endowment Foundation for possible use in their publications and training.

The research ends around the time of the pandemic. The latter seems to have caused, or is at least correlated with, profound changes in aspects of education such as behaviour and attendance.

EPILOGUE

This takes the form of four proposals to help the future of school improvement excite and engage the profession rather than oppressing it, as often happens now, evidenced by the drain of teachers [p.124]. The reflections are based on this study and on Ofsted's 2023 'Big Listen': this survey showed that Ofsted inspections always affect staff wellbeing, "especially acutely in schools (72%)" (IFF Research, 2024, p.4), and "only a third ... of schools said Ofsted was effective [in supporting improvement]" (*ibid* p.45).

I have argued that although effectiveness represents a true paradigm, and its emergence in 1979 a true paradigm shift, school improvement has not achieved paradigm status. However, from around 1990, governments have treated improvement as a paradigm with rules, principles and processes: a 'blind faith' improvement replacing an evidence-based

effectiveness. Perhaps we have stopped asking difficult questions about the validity of one-day inspections, the value of school performance tables, the MAT as the single tool of improvement and the efficacy of centrally imposed programmes such as approved systems of phonics. Whatever their possible merits, they are not kept under meaningful review but are part of the 'taken-for-granted' apparatus of government; I propose that their continuing effectiveness be questioned, continuously.

It is increasingly clear that improvement is not a matter for schools alone: as the Children's Commissioner's (2024) report indicates, it needs the whole panoply of services - school placement, mental health, nutrition, speech and language therapy, social care and welfare - to work under the same roof, with their budgets harnessed, to make the school component possible.

As I write (September 2024), a new government is promising to introduce regional improvement teams to support struggling schools, perhaps marking a change from the hands-off, MAT-monopolistic approach of its predecessor. If this is part of an approach which privileges support and guidance over naming and shaming, leading to the extinction of brutish slogans like 'zero tolerance', then I look forward to a school improvement future which stands a chance of engaging the profession.

As Ofsted conducted its Big Listen, two former senior members of HMI conducted the Alternative Big Listen, asking questions that the official survey did not. They found that 89% of respondents agreed with the statement "Ofsted does NOT sufficiently consider the context of a school's local area in the judgments they make" (Richards & Norris, 2024, p.3). Former Ofsted head of inspection Sean Harford memorably said that an Ofsted inspection grade is not an effort grade: perhaps it should be, since there are schools with challenges so great that inspectors, including myself, have been unable to suggest a single improvement strategy that the school has not already tried. The threat I see is that, since the pandemic, changes in pupils' behaviour and attitudes are bringing to many schools the overwhelming challenges previously faced only by the few.

REFERENCES

Accella, I. (2011) 'The Focus Groups in Social Research: Advantages and Disadvantages.' *Quality Quantity* Vol.46(4), pp.1125–36.

Adams, W.(2015) *Conducting Semi-structured Interviews* Chapter 19 of Newcorner K, Hatry H & Wholey J *Handbook of Practical Program Evaluation* 4th Edition. Hoboken, NJ: Jossey-Bass.

Adonis, A. (2012) *Education, Education, Education: Reforming England's Schools*. London: Biteback Publishing.

Aikin, W.M. (1942) *The Story of the Eight-year Study*. New York: Harper and Brothers.

Ainscow, M. and Southworth, G. (1996) 'School improvement: a study of the role of leader. *School Effectiveness and School Improvement*. Vol. 7(3), pp.229-251.

Allen, R., Evans, M. & White, B. (2021) *The Next Big Thing in School Improvement*. Woodbridge: John Catt Educational.

Ambition Institute (2023) *National Professional Qualifications*. Available at: <https://www.ambition.org.uk/npqs-national-professional-qualifications/#:~:text=There%20are%20eight%20NPQ%20programmes,by%20the%20Department%20for%20Education> (Accessed Sept 2023).

Anderson, V. (2017). 'Criteria for evaluating qualitative research', *Human Resource Development Quarterly*, Vol.28(2), pp.125–133.

Andrews, J., Robinson, D. & Hutchinson, J. (2017) *Closing the Gap? Trends in Educational Advantage and disadvantage*. Education Policy Institute. Available at: https://epi.org.uk/wp-content/uploads/2017/08/Closing-the-Gap_EPI-.pdf (Accessed Nov.2023).

Atherton, C. *Reading Lessons: the books we read at school*, London: Penguin Books.

Auld Report (1976) *The William Tyndale Junior and Infants Schools: Report of the public inquiry conducted by Mr Robin Auld, QC* Inner London Education Authority.

Ball, S.J. (2021) *The Education Debate* 4th edition. Bristol: Policy Press, Bristol University.

Barber, M. (2009) in A. Hargreaves M. Fullan (Eds) *Change Wars*. Bloomington, Indiana: Solution Tree.

Barber, M., & Mourhsed, M. (2007) *How the World's Best-Performing School Systems Come out on Top. The McKinsey Report*. New York: McKinsey.

Barton, G. (2023) Quoted in *The Independent*. Article entitled "Increase in number of teachers leaving profession 'hugely concerning'". 8th June 2023.

BBC (2010) News Channel *More schools are failing Ofsted checks*. March 10th 2010. Available at: <http://news.bbc.co.uk/1/hi/education/8559402.stm> (Accessed March 2024).

BBC (2017) News Website: Transport section. Dec. 22nd 2017. Available at: <https://www.bbc.co.uk/news/uk-42182497> (Accessed April 2023).

BBC (2019) News Website: Family & Education Sept 28th 2019. Available at <https://www.bbc.co.uk/news/education-49841620> (Accessed March 2023).

Bellei, C., Morawietz, L., Valenzuela, J. & Vanni, X.. (2019). *Critical issues for the study of school improvement: Contributions from a research program in Chile*, in *The SAGE Handbook of School Organization* (pp.650 – 665). London: SAGE Publications.

Bennett, A. (2004) *The History Boys*. London: Faber & Faber.

Benson, J. (1994) *The Rise of Consumer Society in Britain, 1880 – 1980*. Harlow: Longman.

BERA (British Educational Research Association) (2018). *Ethical Guidelines for Educational Research, 4th Edition*. Available at: <https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018-online> (Accessed June 2022).

Berg, L. (1974) *Risnghill: Death of a Comprehensive School*. London: Penguin Books.

Berger, P.& Luckman, B. (1969) *The Social Construction of Reality*. London: Penguin Books.

Bernstein, B. (1970). 'Education cannot compensate for society', *New Society*, Vol.38, pp. 344-347.

Best, G. (2009) 'The Father of the Permissive Society', *History Today*, Vol 59(6).

Betts, J. R. (1995). Does school quality matter? evidence from the national longitudinal survey of youth', *The Review of Economics and Statistics*, Vol.77(2), pp.231 - 250.

Blair, Tony (1996a) Speech to the Labour Party Conference, Blackpool, October 4th.

Blair, Tony (1996b) 'The agenda for a generation', speech at Ruskin College, Oxford, 18th December 1996.

Bogdanor, V. (2018) *Britain in the Twentieth Century: Thatcherism, 1979 – 90*, Public lecture at Gresham College. Available at: https://www.youtube.com/watch?v=IOiJnNN8bmc&list=PL98A2AB26E26E27F0&index=11&ab_channel=GreshamCollege (Accessed February 2023).

Bolton, P. (2012) *Education: Historical Statistics*. House of Common Library Research Briefing Ref: SN/SG/4252. Available at: <https://researchbriefings.files.parliament.uk/documents/SN04252/SN04252.pdf> (Accessed March 2023).

Braun, v., Clarke, V., Hayfield, N., Terry, G. (2019). 'Thematic Analysis', in Liamputtong. P. (ed). *Handbook of Research Methods in Health Social Studies*, pp 843-860. California: Sage Publications.

Brighouse, T. (1993) Private discussion with the author, Keele University.

Brighouse, T. (2020) Foreword to NAHT (National Association of Headteacher) (2020) *Improving Schools: A Report of the School improvement Commission*. Available at: <https://features.naht.org.uk/improving-schools-report/index.html> (Accessed Jan 2023).

Brighouse, T. & Waters, M. (2021) *About Our Schools*. Carmarthen: Crown House Publishing.

Brown, R. & Kulik, J. (1977) 'Flashbulb memories', *Cognition*. Vol.5(1), pp.73–99.

Burušić, J., Babarović, T., & Velić, M. Š. (2016). 'School Effectiveness: An Overview of Conceptual, Methodological and Empirical Foundations', *School Effectiveness and Educational Management* (pp. 5-26). Springer International Publishing.

Callaghan, J. (1976) *A Rational Debate based on the Facts*. Speech at Ruskin College, Oxford, October 16th, 1976.

Campbell, A. (2001) "The day of the bog-standard comprehensive is over", Press Briefing 12 February 2001. Quoted widely, including in the Guardian; available at: <https://www.theguardian.com/education/2001/dec/24/schools.uk1> (Accessed June 2023).

Campbell, S., Greenwood, M., Prior S., Shearer, T., Walkern, D., Young, S., Bywaters, D & Walker, K. (2020a) 'Purposive sampling: complex or simple? Research case examples', *Journal of Research in Nursing* Vol,28(8), pp.652 – 661.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2020b). 'Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement.' *Sociological Methods & Research*, 50(3), 837-865.

Card, D., & Krueger, A. B. (1996) 'Labor market effects of school quality: Theory and evidence', in G. Burtless (ed.), *Does Money Matter? The Link Between Schools, Student Achievement and Adult Success*, Washington, DC: Brookings Institute, pp.97-140.

Case, P., Case, S. & Catling, S. (2000) 'Please Show You're Working: a critical assessment of the impact of Ofsted inspection on primary teacher'. *British Journal of Sociology of Education*, vol.21(4), pp.605 – 621.

CBI/Pearson (2014) *Gateway to Growth: Education and Skills Survey 2014*. London: Confederation of British Industry.

Chapman, C., Allen, T. (2006) 'Collaborative reform for schools in difficulty', *Improving Schools*. 9 (3), pp.291-301.

Chapman, C., Armstrong, P., Harris, A., Muijs, D., Reynolds, D., & Sammons, P. (Eds.). (2011). *School Effectiveness and Improvement Research, Policy and Practice*. London: Routledge.

Children's Commissioner (2024) *The Big Ambition: Ambitions, Findings and Solutions*. Available at: <https://www.childrenscommissioner.gov.uk/resource/the-big-ambition-ambitions-findings-and-solutions/> (Accessed March 2024).

Clough, P. (1999) 'Crises of schooling and the 'crisis of representation'', *Qualitative Inquiry* Vol.5(3) pp.428–48.

Coe, R. (2013) *Improving Education: A triumph of hope over experience*. Text of inaugural professorial lecture, CEM University of Durham, 18th June 2013. Available at: https://www.academia.edu/5218549/INAUGURAL_LECTURE_of_Professor_Robert_Coe_Director_of_CEM_and_Professor_of_Education_at_the_School_of_Education_IMPROVING_EDUCATION (Accessed May 2023).

Coe, R. (2019) *What Makes Great Teaching?* Presentation at the Cambridge School Conference, Cambridge, UK. September 14th 2019. Available at: <https://www.cambridgeinternational.org/Images/557877-rob-coe-what-makes-great-teaching-.pdf> (Accessed Sept. 2023).

Cohen, D. & Crabtree, B. (2006): 'Qualitative Research Guidelines Project. Semi-structured interviews' New Jersey: Robert Wood Johnson Foundation Available at: <http://www.qualres.org/HomeEval-3664.html> (Accessed: March 2022).

Coleman, J. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Coleman, M. (2002) *Women as Headteacher: Striking the Balance*. London: Trentham Books.

Collins, J. (2001) *From Good to Great: Why Some Companies Make the Leap and Others Don't*. London: Harper Collins.

Conservative Party (2019). *The Conservative Party Manifesto*. Available at: <https://www.conservatives.com/our-plan/conservative-party-manifesto-2019> (Accessed: October 2023).

Corallo, C. & McDonald, D.H. (2002) *What Works with Low-Performing Schools*. Lanham, Maryland: Rowman & Littlefield Education.

Cox, C.B. & Dyson, A.E. 'Comment', *Critical Survey*, 4: 1 (1969), pp.2– 6. First in the series of 'Black Papers', published in the same journal, various dates to 1974.

Creemers, B.; Peters, T.; Reynolds, D. (eds) (1989) *School Effectiveness and School Improvement*. London: Routledge.

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge.

Cresswell, J. (1998) *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Crisp, R., Macmillan, R., Robinson, D., & Wells, P. (2009). 'Continuity or Change: considering the policy implications of a Conservative government', *People, Place and Policy*, 3(1), pp. 58-74. Available at: <https://doi.org/10.3351/ppp.0003.0001.0006> (Accessed Jan 2024).

Critical Quarterly website *History of the Critical Quarterly* Available at: <https://web.archive.org/web/20070929022136/http://www.criticalquarterly.com/history.asp> (Accessed April 2023).

Cuttace, P. (1982) 'Reflections on the Rutter Ethos: A Professional Researcher's View of Fifteen Thousand Hours', *Urban Education* Vol.16(4), pp.483-492.

Davies, Howard (1991) Speech to Council of Local Education Authorities (CLEA) conference, London, 18th July 1991.

Davies, Hunter. (1976) *The Creighton Report: A Year in the Life of a Comprehensive School*. London: Hamilton House.

Davis, J. (2002) The Inner London Education Authority and the William Tyndale Affair, 1974 - 1976 *Oxford Review of Education* 28.2/3, *A Century of Local Education Authorities* (June–September 2002) 275–98.

Department for Children, Schools and Families (DCSF) (2003) *Every Child Matters* Green Paper London: TSO. Available at: <https://www.gov.uk/government/publications/every-child-matters> (accessed April 2023)

Department for Children, Schools and Families (DCSF) (2007) *Guidance on the Duty to Promote Community Cohesion*. DCSF Publications, Nottingham. Available at: <http://www.tedcandle.co.uk/publications/029%20Guidance%20on%20duty%20to%20promote%20community%20cohesion%20in%20school.pdf> (Accessed June 2023).

Denshire, S. (2013) 'Autoethnography', *Sociopedia.isa*. Available at: <https://pdf4pro.com/amp/view/autoethnography-sage-publications-9d74c.html> (Accessed Feb 2022)

Department for Education (DfE) (1992) *The Education (Schools) Act*. Available at: <https://www.legislation.gov.uk/ukpga/1992/38/contents> (Accessed March 2021).

Department for Education (DfE) (2010a) *The Case for Change*. Education White Paper, November 2010.DFE-00564-2010 Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/526946/The_case_for_change_The_importance_of_teaching.pdf (Accessed May 2023).

Department for Education (DfE) (2010b) *Academies Act*. Available at: <https://www.legislation.gov.uk/ukpga/2010/32/contents> (Accessed May 2023).

Department for Education (DfE) (2010c) *The Importance of Teaching*. Education White Paper. Available at: <https://www.gov.uk/government/publications/the-importance-of-teaching-the-schools-white-paper-2010> (accessed May 2023).

Department for Education (DfE) (2014) *The National Curriculum in England: Framework for Key Stages 1 – 4*. Available at: <https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4/the-national-curriculum-in-england-framework-for-key-stages-1-to-4> (Accessed March 2024).

Department for Education (DfE) (2016c). *Revised GCSE & Equivalent Results in England, 2014- 15. Published 21st January 2016*. Available at: https://assets.publishing.service.gov.uk/media/5a80042ce5274a2e8ab4daf0/SFR01_2016.pdf (Accessed March 2024).

.

Department for Education (DfE) (2013b) *The National Curriculum* DfE: Reference: DFE-00178-2013. Available at: https://assets.publishing.service.gov.uk/media/5a81a9abe5274a2e8ab55319/PRIMARY_national_curriculum.pdf (Accessed March 2023)

Department for Education (2016a) *Multi-academy trusts: Good practice guidance and expectations for growth*. Available at: <https://dera.ioe.ac.uk/id/eprint/28029> (Accessed Sept 2024)

Department for Education (2016b) *Coasting Schools*. Available at: https://assets.publishing.service.gov.uk/media/5a759ffbed915d506ee8040d/SFR62_KS2_2016_Coasting_text.pdf

Department for Education (DfE) (2018). *School leadership in England 2010 to 2016: characteristics and trends*. DfE Ref. RR812

Department for Education (DfE) (2022c) *The Coasting Schools (England) Regulations 2022*. Available at: <https://www.legislation.gov.uk/ukxi/2022/720/regulation/4> (Accessed June 2024).

Department for Education (DfE) (2023b). *Participation in education, training and employment age 16 to 18*. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/participation-in-education-and-training-and-employment> (Accessed April 2024).

Department for Education (DfE) (2024) *Schools Causing Concern: Guidance on Intervention*. Available at: https://assets.publishing.service.gov.uk/media/659d064e614fa20014f3aa63/Schools_causing_concern_guidance.pdf (Accessed June 2024).

Department for Education & Employment (DfEE) (1996) *Self-government in Schools*. White Paper, Cmnd 3315. London: HMSO.

Department for Education & Employment (DfEE) (1997) *Excellence in Schools*. Government White Paper, HMSO. Reproduced in full in Gillard *Education in England* and available at: [http://www.educationengland.org.uk/documents/wp1997/excellence-in-schools.html#:~:text=Excellence%20in%20Schools%20\(1997\)%20%2D%20full%20text%20on](http://www.educationengland.org.uk/documents/wp1997/excellence-in-schools.html#:~:text=Excellence%20in%20Schools%20(1997)%20%2D%20full%20text%20on)

line&text=This%20was%20New%20Labour's%20first,established%20Education%20Action%20Zones%20etc. (Accessed January 2023).

Department for Education & Employment (DfEE) (2000) Learning and Skills Act. Available at: <https://www.legislation.gov.uk/ukpga/2000/21/contents> (Accessed June 2023)

Department for Education and Science (DES) (1965) *The Organisation of Secondary Education. Circular 10/65*. Available at <http://www.educationengland.org.uk/documents/des/circular10-65.html> (Accessed March 2023).

Department of Education and Science (DES) (1982) *Study of HM Inspectorate in England and Wales*. London: HMSO.

Department of Education and Science (DES) (1970) *HMI Today and Tomorrow*. London: HMSO.

Department for Education & Science (DES). (1981) *The School Curriculum*. London: HMSO.

Department of Education and Science (DES) (1988) *Education Reform Act (ERA)* Available at: <https://www.legislation.gov.uk/ukpga/1988/40/contents> (Accessed June 2022).

Department for Education and Skills (DES) (1981) *New Training Initiative: An Agenda for Action* (White Paper). Cmnd 8455. London: HMSO.

Department for Education (DfE) (2011a) *The National Strategies 1997–2011 A brief summary of the impact and effectiveness of the National Strategies*. DfE.Ref Number 00032-2011PDF-EN-01.

Department for Education (DfE) (2013) *Reform of the National Curriculum in England Consultation Paper*. Available at: <https://ico.org.uk/media/about-the-ico/consultation-responses/2013/2134/consultation-document-reform-of-the-national-curriculum-20130207.pdf> (Accessed May 2023)

Department for Education (DfE) (2015a) *The Purpose of Education* Speech given by the Rt Hon Nick Gibb MP to the Education Reform Summit, London, 9th July 2015. Text available at: <https://www.gov.uk/government/speeches/the-purpose-of-education#:~:text=Schools%20Minister%20Nick%20Gibb%20addresses%20the%20Education%20Reform%20Summit.&text=Education%20reform%20is%20the%20great,people%2C%20regardless%20of%20their%20background> (Accessed Jan 2023)

Department for Education (2016b) *Coasting Schools: Provisional Data*. Available at: https://assets.publishing.service.gov.uk/media/5a809e78ed915d74e33fb764/coasting_schools_note.pdf. (Accessed April 2024)

Department for Education (DfE) (2017) *House prices and schools: do houses close to the best performing schools cost more?* London: DfE Publications, Reference number:00098-2017.

Department for Education (DfE) (2022a) *Guidance: Sponsor an Academy*. Available at <https://www.gov.uk/guidance/sponsor-an-academy> (Accessed Jan 2023).

Department for Education (DfE) (2022b) *Opportunity for all Strong schools with great teachers for your child*. Government White Paper London: HMSO CP 650. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063602/Opportunity for all strong schools with great teacher for your child print version .pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063602/Opportunity_for_all_strong_schools_with_great_teacher_for_your_child_print_version.pdf) (Accessed Sept 2022).

Department for Education (DfE) (2022b) *Priorities for Education* DfE website front page, available at: <https://www.gov.uk/government/organisations/department-for-education/about> Accessed Jan 2023).

Department for Education (DfE) (2022a) *Schools, Pupils and their Characteristics (Official Statistics)*. Available at: <https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2022> (Accessed: April 2023)

Department for Education (DfE) (2011b) *Review of Vocational Education: The Wolf Report*. Available at:

<https://assets.publishing.service.gov.uk/media/5a7a38c4ed915d1fb3cd6520/DFE-00031-2011.pdf> (Accessed July 2024).

Department for Education (DfE) (2016a) *Multi-academy trusts: Good practice guidance and expectations for growth*. London: DfE, Ref. DFE-00334-2016.

Department for Education (DfE) (2023a) *Keeping Children Safe in Education*. Available at: <https://www.gov.uk/government/publications/keeping-children-safe-in-education--2> (Accessed June 2024).

Department for Education (DfE) (2024) *Schools, pupils and their characteristics. Academic year 2023/24*. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics> (Accessed November 2024)

Department for Education and Skills (DfES) (2004) *Memorandum by the Department for Education and Skills Select Committee Enquiry into Every Child Matters November 2004*. Available at: <https://publications.parliament.uk/pa/cm200405/cmselect/cmeduski/uc40-vi/memo58.pdf> (Accessed October 2023)

.

Department for Education and Skills/Office for Standards in Education (DfES/Ofsted) (2004) *A New Relationship with Schools* London: HMSO.

Didau, D. (2020). *Intelligent Accountability: Creating the conditions for teacher to thrive*. Woodbridge: John Catt Educational.

Doyle, J., & Wells, S. (1997). 'Social Class and the effective school paradigm', *International Journal of Educational Management*, Vol.11(4), pp.145-148. Available at: <https://www.emerald.com/insight/content/doi/10.1108/09513549710186245/full/html> (Accessed Jan 2023).

Drucker, P. (1966) *The Effective Executive*. London: Heinemann.

Dutton, D. (1997). *British Politics Since 1945: The Rise, Fall and Rebirth of Consensus* (2nd ed.) London: Blackwell.

Eden Sir A. (1956) Speech at Bradford, Jan 18th 1956. Available at: https://www.youtube.com/watch?v=Yb3c7JaB9es&ab_channel=BritishMovietone Accessed March 2023.

Edmonds, R. (1979) 'Effective schools for the urban poor', *Educational Leadership*, Vol.36, pp.15-17.

Education Executive (2018). 'Survey shows inconsistency in school inspections' *Education Executive*. 18th April 2018. Available at: <https://edexec.co.uk/survey-shows-inconsistency-in-school-inspections/> Accessed March 2018.

EEF (The Education Endowment Foundation) (2017) *The Attainment Gap* London: The Sutton Trust. Available at: https://educationendowmentfoundation.org.uk/public/files/Annual_Reports/EEF_Attainment_Gap_Report_2018_-_print.pdf (Accessed March 2023).

Ehren, M., Hutchinson, J. & Munoz-Chereau, B. (2023) 'Place-based Disparities Faced by Stuck Schools in England'. *School Effectiveness & School Improvement*, Vol.34(4), pp.401 – 418.

Ellis, C., Adams, T. & Bochner, A. (2011) 'Autoethnography - an overview', *Forum for Qualitative Social Research*, Vol.12,(1). Available at: <https://www.qualitative-research.net/index.php/fqs/article/view/1589/3095> (Accessed February 2022).

Feduccia, A. (1995). 'Explosive evolution in Tertiary birds and mammals', *Science*, 267 (5198) pp.637-638.

Finkenauer, C., Luminet, O.; Gisle, L.; El-Ahmadi, A., Van Der Linden, M. and Philippot, P. (1998). 'Flashbulb memories and the underlying mechanisms of their formation: Toward an emotional-integrative model', *Memory & Cognition*. Vol.26(3), p.516 – 531.

Finlay, L. (2002). "'Outing" the researcher: The provenance, process, and practice of reflexivity.' *Qualitative Health Research*, 12(4), 531-545.

Fores, M. (1971) Britain's Economic Growth and the 1970 Watershed *Lloyds Bank Review*, Jan 1971

Foucault, M. (1975) *Discipline and Punish: the Birth of the Prison*. (English language edition 1977) New York: Random House

Friedman M, (1955) *The Role of Government in Education in Economics and the Public Interest*, ed. Robert A. Solo New Brunswick, NJ; Rutgers University Press.

Fullan, M. (2006) *Change theory: A force for school improvement* Centre for Strategic Education, Seminar Series Paper No 157, Nov 2006. Victoria, Australia.

Fullan, M., (2007) *The new meaning of educational change*. New York: Routledge.

Fullan, M. (2008) From School Effectiveness to System Improvement: An Inevitable Conceptual Evolution Paper prepared for *Journal für Schulentwicklung (Journal for School Development)* Available at: https://michaelfullan.ca/wp-content/uploads/2016/06/Untitled_Document_25.pdf (Accessed Nov. 2022)

Galvan, J. & Galvan, M. (2017) *Writing Literature Reviews*. London: Routledge.

Giles, C., & Hargreaves, H. (2006). 'The Sustainability of Innovative Schools as Learning Organizations and Professional Learning Communities during Standardized Reform', *Educational Administration Quarterly*, Vol.42, pp.124-156.

Gillard, D. (2018) *Education in England: a history of our schools* Legislation section. Available at: <http://www.educationengland.org.uk/documents/acts/index.html> (Accessed Jan 2023)

Given, L. (2008). "Convenience Sample". In Green, L. (ed) *The SAGE Encyclopedia of Qualitative Research Methods*. Los Angeles: SAGE Publications.

Goldstein, H. (1997) 'Methods in school effectiveness research', *School effectiveness and School Improvement*. Vol 8, pp.369-95.

Gorard, S. (2010). 'Serious doubts about school effectiveness', *British Educational Research Journal*, Vol.26(4), pp.745 – 766.

Gorard, S. (2011). Doubts about school effectiveness exacerbated – by attempted justification. *Research Intelligence*, No.114, pp.24 - 26.

Gove, M. (2012) Speech *How are the children? Achievement for all in the twenty-first century*. Spectator Conference, London, 18th June 2012.

Hallinger, P., & Kovacevic, J. (2019). 'Mapping the intellectual lineage of educational management, administration and leadership, 1972–2020'. *Education Management, Administration & Leadership*. Volume 50, Issue 2

Hansard (1971) *Schools (Truancy)*. Debated 23rd June 1971. Volume 819 Columns 1553-62. Available at <https://api.parliament.uk/historic-hansard/commons/1971/jun/23/schools-truancy> (Accessed March 2024).

Hansard (1987a) *Queen's Speech* HC 25th June 1987. Vol 118 cc38-40

Hansard (1987b) *Education Reform Bill*. Debated 1st December 1987. Vol. 123, Column 78. Available at: <https://hansard.parliament.uk/commons/1987-12-01/debates/f4935e71-5df7-48fb-9aa3-943231f8c6ca/EducationReformBill> (Accessed March 2024).

Hansen, S. & Jenssen, J. (1969) *The Little Red Schoolbook*. (First publication in the UK by Handyside and Collins, London, 1970).

Hargreaves, D. (1982) *The Challenge for the Comprehensive School*. London: Routledge, Kegan & Paul.

Harris, A. Jamieson, I. & Ross, J. (1995) 'A Study of 'Effective' Departments in Secondary Schools', *School Organisation*, Vol.15(3) pp. 283-299.

Hart, C. (1998) *Doing a Literature Review*. London: Sage Publications.

Herzberg, F. (1959), *The Motivation to Work*, New York: John Wiley and Sons.

Hines, B. (1968; filmed in 1969) *Kes: A Kestrel for a Knave*. London: Michael Joseph

Hodge, G. (2007). 'The origins of competency-based training', *Australian Journal of Adult Learning*, Vol.47(2), pp.179-209.

Holt, N. (2003) 'Representation, legitimation, and autoethnography: An autoethnographic writing story', *International Journal of Qualitative Methods* Vol.2(1): 1–22.

Hopkins, D. (Ed) (1987) *Improving the Quality of Schooling*. Lewes: Routledge Falmer Press.

Hopkins, D. (2022) *School Improvement: Precedents & Prospects*. Leading Excellence Series no.12. Melbourne: Centre for Strategic Excellence.

Hopkins, D. & Reynolds, D. (2001) 'The Past, Present and Future of School Improvement: towards the Third Age', *British Educational Research Journal*, Vol27(4).

Hopkins, D., Stringfield, S., Harris, A., Stoll, L. & Mackay, T. (2014) 'School and system improvement: a narrative state-of-the-art review', *School Effectiveness & School Improvement*, Vol 25(2) Available at: <https://www.tandfonline.com/doi/full/10.1080/09243453.2014.885452> (Accessed Jan 2023).

Hopkins, E, with Garrod, F., McClare, S., Pettit, D., Smith, L. and Temple, J. (2016) 'Teacher views of the impact of school-evaluation and external inspection processes.', *Improving Schools*, Vol.19(1) pp. 52–61.

Horrell, S. (2000) 'Living standards in Britain 1900–2000: Women's century?' *National Institute Economic Review* No. 172 (April 2000), pp.62 – 77. Cambridge University Press.

House of Commons (1999) *A Century of Change : UK Statistics since 1900*. Research Paper 99/111. Available at: <https://researchbriefings.files.parliament.uk/documents/RP99-111/RP99-111.pdf> (Accessed March 2023)

House of Commons (2010) *The Role and Performance of Ofsted*. NFER Report to the Education Select Committee. Available at: <https://publications.parliament.uk/pa/cm201011/cmselect/cmeduc/writev/ofsted/64.htm> (Accessed March 2024).

House of Commons Library (2012) *Education: Historical Statistics*. Available at: <https://researchbriefings.files.parliament.uk/documents/SN04252/SN04252.pdf> (Accessed March 2024).

House of Commons (2020) *Grammar School Statistics: Briefing Paper 1398* London, HMSO. Available at: <https://researchbriefings.files.parliament.uk/documents/SN01398/SN01398.pdf> (Accessed March 2023).

House of Commons Library (2024) *Pupil Attendance in Schools*. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/pupil-attendance-in-schools> (Accessed March 2024).

Hudson, L. A., & Ozanne, J. L. (1988). 'Alternative ways of seeking knowledge in consumer research', *Journal of Consumer Research*, 14, pp.508-521.

Huitt, W. (2019). 'Analyzing paradigms used in education and schooling', *Educational Psychology Interactive*. Atlanta, GA: Available at: <http://www.edpsycinteractive.org/topics/intro/analyzing-paradigms.pdf> (Accessed Jan 2023)

Hume, D. (1739) *A Treatise of Human Nature*. London: Penguin Classics.

IDACI (Income Deprivation Affecting Children Index) (2020) *Ofsted Inspection Data & IDACI*. Available at: <https://daisi.education/ofsted-inspection-judgements-and-idaci/#:~:text=Overall%20Effectiveness,-Firstly%2C%20for%20these&text=There%20is%20a%20correlation%20between%20Deprivation%20and%20Requires%20Improv> (Accessed September 2023).

IFF Research (2024) *Ofsted Big Listen research report: findings from professionals*. Available at: <https://assets.publishing.service.gov.uk/media/66d06d137c42acbece502c8b/ofsted-big-listen-research-report-findings-from-professionals-IFF-Research.pdf> (Accessed Sept. 2024)

ISC (Independent Schools Council) *Census and Annual Report 2022*. Available at: https://www.isc.co.uk/media/8421/isc_census_2022_final-v2.pdf (Accessed March 2023).

Jencks, C. (1972). *Inequality: A reassessment of the effects of family and schooling in America*. New York: Basic Books.

Jenkins, P. (1987) *Mrs. Thatcher's Revolution*. London: Jonathan Cape.

Jones, E. (2023) Podcast interview with Simon Jordan 27th October 2023. Available at: https://www.youtube.com/watch?v=uHjj5bd92m0&ab_channel=UpFrontWithSimonJordan (Accessed November 2023).

Jones, K. (2016) *Education in Britain, 1944 to the Present*. Cambridge: Polity Press.

Joseph, Sir K. (1974) 'Our Human Stock is Threatened', Speech at the Grand Hotel, Edgbaston, 19th October 1974. Available at <https://www.margaretthatcher.org/document/101830> (Accessed April 2023)

Joynson, R. B. (1989). *The Burt Affair*. London: Routledge.

Kelly, S. (2010) *Qualitative interviewing techniques and styles* In Bourgeault, I., Dingwall, R. and de Vries, R. (eds) *The Sage Handbook of Qualitative Methods in Health Research*. Thousand Oaks, CA: Sage Publications.

Khan, Haider. (2002). *On Paradigms, Theories and Models*. CIRJE, Faculty of Economics, University of Tokyo, CIRJE F-Series. 34.

Kidron, Y. & Darwin, M. (2007) 'A Systematic Review of Whole School Improvement Models' *Journal of Education for Students Placed at Risk*. Vol 12, issue 1.
Available at:
<https://www.tandfonline.com/doi/full/10.1080/10824660701247226> (Accessed May 2024).

Knowles, E. & Evans H. (2012) *PISA 2009: How does the social attainment gap in England compare with countries internationally?* London: DFE Ref: DFE-RR206.

Kuhn, Thomas S. (1970) *The Structure of Scientific Revolutions*, 2nd edition. Chicago: University of Chicago Press

Kyriakides, L. (2004) 'Differential School Effectiveness in Relation to Sex and Social Class: Some Implications for Policy Evaluation', *Educational Research and Evaluation*, Vol 10(2), pp. 141-161.

Labour Party (2019) *It's Time for Real Change: The Labour Party Manifesto*. Available at <https://ucrel.lancs.ac.uk/wmatrix/ukmanifestos2019/localpdf/Labour.pdf> (Accessed June 2024)

Laslett, B. (1999) *Personal Narratives in Sociology*. New Brunswick, NJ: Rutgers University Press.

Leckie, G. & Goldstein, H. (2016) *The evolution of school league tables in England 1992-2016: 'contextual value added', 'expected progress' and 'progress 8 '*. University of Bristol: Working Paper #02.2016.

Leithwood, K., Jantzi, D. & McElheron-Hopkins, C. (2006). 'The Development and Testing of a School Improvement Model'. *Journal of School Effectiveness and School Improvement*, Vol 17(4), pp 441 – 464.

Lester, J., Cho, Y. & Lochmiller, C. (2020) 'Learning to Do Qualitative Data Analysis: A Starting Point', *Human Resource Development Review* 2020, Vol.19(1), pp.94–106.

Lezotte, L.W. (1991) *Correlates of Effective Schools: The First and Second Generation*.

Okemos, MI: Effective Schools Products Ltd.

MacGilchrist, B. & Mortimore, P. (1997). 'The impact of school development plans in primary schools'. *School effectiveness and school Improvement*, Vol.8, pp.198 – 218 .

Macmillan, H. (1957) Conservative rally speech (untitled), Bedford, July 20th 1957.

Major, Sir J. (2011) Ditchley Annual Lecture Speech (Ditchley Foundation), Saturday 9th July 2011, Ditchley Park, Oxfordshire.

Marx, S., Pennington, J. & Chang H (2017) 'Critical Autoethnography in Pursuit of Educational Equity', *International Journal of Multicultural Education*, Vol 19(1), pp.1-6.

Marzano, R., Warrick, P. and Simms, J. *A Handbook of High Reliability Schools*. Bloomington: Marzano Research.

McCrone, T., Coghlan, M., Wade, P. & Rudd, P. (2009) *Evaluation of the Impact of Section 5 Inspections*. Slough: NFER (National Foundation for Educational Research).

McFedries, P. (2001) *The Complete Idiot's Guide to a Smart Vocabulary* New York: Alpha Publishing.

Mencap website. <https://www.mencap.org.uk/> (Accessed March 2023).

Migration Watch UK (2021) *The Impact of Immigration in Changing the UK Population*. Available at: <https://www.migrationwatchuk.org/briefing-paper/492/impact-of-immigration-in-changing-the-uk-population#:~:text=The%20ethnic%20minority%20component%20of,funded%20primary%20and%20secondary%20schools>. (Accessed April 2023)

Miles, M. B. (1967). *Some properties of schools as social systems*, in G. Watson (Ed.), *Change in School Systems*. Washington, DC: National Training Laboratories.

Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Ministry of Education (1944) *Education Act (The Butler Act)*. Available at: <https://www.parliament.uk/about/living-heritage/transformingsociety/livinglearning/school/overview/educationact1944/> (accessed June 2023).

Ministry of Education (1956) *White Paper: Technical Education*. February 1956. Available at: <http://filestore.nationalarchives.gov.uk/pdfs/small/cab-129-79-cp-56-40-40.pdf> Accessed March 2023.

Mittchell, R., Sung, J., Zhang, S., Mendiola, B. and Tarler, J. (2015) *School Effectiveness: a Meta-analysis Review of Published Research*. In M. diPaolo and W. Hoy (Eds.) *Leadership for School Quality*. Charlotte, North Carolina: Information Age Publishing.

Montacute, R. & Cullinane, C. (2018) *PARENT POWER 2018 How parents use financial and cultural resources to boost their children's chances of success*. London: Sutton Trust. Available at: <https://www.suttontrust.com/wp-content/uploads/2019/12/Parent-Power-2018.pdf> (Accessed Jan 2023).

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988) *School Matters*. London: Open Books.

Muijs, D., Campbell, J., Kyriakides, L., Robinson, W. (2005) 'Making the Case for Differentiated Teacher Effectiveness: An Overview of Research in Four Key Areas', *School Effectiveness & School Improvement*, Vol.16, (1), pp. 51 – 70.

Munoz-Chereau, B., Hutchinson, J., & Ehren, M. (2002). '*Stuck' schools: Can below good Ofsted inspections prevent sustainable improvement?* London: University College & Institute of Education. Available at: <https://epi.org.uk/publications-and-research/stuck-schools-can-below-good-ofsted-inspections-prevent-sustainable-improvement/> (Accessed Sept.2024).

Murgatroyd, S. & Morgan, C. (1993) *Total Quality Management and the School*. London: Open University Books.

NAHT (National Association of Headteacher) (2020) *Improving Schools: A Report of the School improvement Commission*. Available at: <https://features.naht.org.uk/improving-schools-report/index.html> (Accessed Jan 2023).

NAO (National Audit Office) (1991) *The Implementation and Development of the Technical and Vocational Education Initiative* London, HMSO

NAO (2018) *Ofsted Inspection of Schools*. Available at: <https://www.nao.org.uk/reports/ofsteds-inspection-of-schools/> Accessed May 2023

NASUWT (National Association of Schoolmasters Union of Women Teacher) (2003). *Behaviour in Schools*. Available at: <https://www.nasuwt.org.uk/static/357990da-90f7-4ca4-b63fc3f781c4d851/Behaviour-in-Schools-Full-Report-September-2023.pdf> (Accessed April 2024).

National Transport Survey 2021 (2022) Available at:
<https://www.gov.uk/government/statistics/national-travel-survey-2021/national-travel-survey-2021-household-car-availability-and-trends-in-car-trips> (Accessed April 2023).

NCEE (National Commission on Excellence in Education) (1983) *A Nation at Risk: The Imperative for Educational Reform A Report to the Nation and the Secretary of Education United States Department of Education*. Washington D.C., US Government Printing Office.

NHS (2023). *Mental Health of Young People in England, 2023*. National Health Service, November 2023. Available at: <https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2023-wave-4-follow-up> (Accessed March 2024).

Nickerson, R.S. (1998). 'Confirmation Bias: A Ubiquitous Phenomenon in Many Guises'. *Review of Psychology*. Vol 2(2), pp.75 – 220.

Nunnery, J A (1998) 'Reform ideology and the locus of development problem in educational restructuring', *Education and Urban Society*, Vol.30(3), pp. 277–295.

OECD PISA (Organisation for Economic Cooperation and Development Programme for International Student Assessment) (2023) Website Available at: <https://www.oecd.org/pisa/> (Accessed March 2023).

Office for Standards in Education (Ofsted) (1992) *Handbook for Inspection*. No longer in print; available from Ofsted as a downloadable file.

Office for Standards in Education (Ofsted) (1995-96) *The Annual Report of Her Majesty's Chief Inspector of Schools*. London: HMSO/TSO. Available at:
<https://www.gov.uk/government/publications/the-annual-report-of-her-majestys-chief-inspector-of-schools-199596> (Accessed July 2024)

Office for Standards in Education (Ofsted) (1998) *The Annual Report of Her Majesty's Chief Inspector of Schools*. London: HMSO/TSO. Available at: <https://assets.publishing.service.gov.uk/media/5a7c768fed915d48c240fe72/0453.pdf> (Accessed May 2024).

Office for Standards in Education (Ofsted) (2006a) *The Annual Report of Her Majesty's Chief Inspector of Schools, 2005-6*. London: HMSO/TSO. Available at: <https://www.gov.uk/government/publications/hm-chief-inspector-of-education-childrens-services-and-skills-annual-report-2006-to-2007> (accessed May 2024).

Office for Standards in Education (Ofsted) (2006) *Departmental Report, 2005-06*. Available at: <https://assets.publishing.service.gov.uk/media/5a7b8844ed915d4147620fc7/6813.pdf> (Accessed June 2024).

Office for Standards in Education (Ofsted) (2007) *The Annual Report of Her Majesty's Chief Inspector of Schools, 2006-7*. London: HMSO/TSO. Available at: <https://www.gov.uk/government/publications/hm-chief-inspector-of-education-childrens-services-and-skills-annual-report-2006-to-2007> (accessed May 2024).

Office for Standards in Education (Ofsted) (2008) *The Annual Report of Her Majesty's Chief Inspector of Schools, 2007-8*. London: HMSO/TSO. Available at: <https://assets.publishing.service.gov.uk/media/5a7c7fa7e5274a559005a493/1114.pdf> (Accessed May 2024).

Office for Standards in Education (Ofsted) (2010) *The Annual Report of Her Majesty's Chief Inspector of Schools, 2009-10*. London: HMSO/TSO. Available at: <https://assets.publishing.service.gov.uk/media/5a7c27b240f0b645ba3c70a9/0559.pdf> (Accessed May 2024).

Office for Standards in Education (Ofsted) (2013a) *The Annual Report of Her Majesty's Chief Inspector of Schools, 20012-13*. London: HMSO/TSO. Available at: <https://assets.publishing.service.gov.uk/media/5a7c9181e5274a0bb7cb7ea6/0855.pdf> (Accessed May 2024).

Office for Standards in Education (Ofsted) (2013) *The Framework for Inspection* Available at: <http://www.educationengland.org.uk/documents/pdfs/2013-ofsted-inspection-framework.pdf> (Accessed May 2023).

Office for Standards in Education (Ofsted) (2014) *Writing the Inspection Report*. Available at: <https://assets.publishing.service.gov.uk/media/5a7c768fed915d48c240fe72/0453.pdf> (Accessed August 2024).

Office for Standards in Education (Ofsted) (2015) *The Common Inspection Framework*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/828112/Withdrawn_common_inspection_framework.pdf (Accessed June 2023).

Office for Standards in Education (Ofsted) (2018) *Deprivation, ethnicity and school inspection judgements* (June 2018), blogpost written by Harford, S. Available at: <https://educationinspection.blog.gov.uk/2018/06/22/deprivation-ethnicity-and-school-inspection-judgements/> (Accessed June 2023)

Office for Standards in Education (Ofsted) (2020) *Fight or Flight? How Stuck Schools are Overcoming Isolation: evaluation report*. Available at: <https://www.gov.uk/government/publications/fight-or-flight-how-stuck-schools-are-overcoming-isolation/fight-or-flight-how-stuck-schools-are-overcoming-isolation-evaluation-report#:~:text=This%20report%20draws%20on%20research,state%2Dfunded%20schools%20in%20England.> (Accessed September 2024)

Office for Standards in Education (Ofsted) (2023a) *Education Inspection Framework*. Available at: <https://www.gov.uk/government/publications/education-inspection-framework/education-inspection-framework-for-september-2023> (Accessed March 2024).

Office for Standards in Education (Ofsted) (2023b) *Schools Commentary: the emerging picture from 2022/2023 inspections*. Updated 7th November 2023. Available at: <https://www.gov.uk/government/publications/school-inspections-statistical-commentaries-2022-to-2023/schools-commentary-the-emerging-picture-from-202223-inspections#:~:text=Graded%20inspections,Figure%201%3A%20Outcomes&text=In%202022%2F23%2C%2078%25,70%25%20in%202021%2F22.> Accessed March 2024.

Office for Standards in Education (Ofsted) (2019) *School Inspection Update Jan 2019*.

Available at:

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/772056/School inspection update - January 2019 Special Edition 180119.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/772056/School_inspection_update_-_January_2019_Special_Edition_180119.pdf) (Accessed June 2023).

Office for Standards in Education (Ofsted) (2021) *Review of Sexual Abuse in Schools and Colleges: Executive Summary & Recommendations* Available at:

<https://www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges> (Accessed April 2024).

Office for Standards in Education (Ofsted) (2022) *Ofsted Strategy 2022 – 27*. Available at:

[https://assets.publishing.service.gov.uk/media/626677b6e90e0716945b1722/Ofsted Strategy 2022 2027.pdf](https://assets.publishing.service.gov.uk/media/626677b6e90e0716945b1722/Ofsted_Strategy_2022_2027.pdf) (Accessed April 2024).

Office for Standards in Education (Ofsted) (2023) *Education Inspection Framework. Updated 14th July 2023*. Available at: <https://www.gov.uk/government/publications/education-inspection-framework/education-inspection-framework-for-september-2023> (Accessed June 2024).

Office for Standards in Education (Ofsted) (2024a) *School Inspection Handbook*. Available at: <https://www.gov.uk/government/publications/school-inspection-handbook-eif/school-inspection-handbook-for-september-2023> (accessed Nov. 2024).

Office for Standards in Education (Ofsted) (2024b) *Main Findings: State Funded Schools Inspections and Outcomes as at Dec 31st 2023*. Available at [https://www.gov.uk/government/statistics/state-funded-schools inspections-and-outcomes-as-at-31-december-2023/main-findings-state-funded-schools-inspections-and-outcomes-as-at-31-december-2023](https://www.gov.uk/government/statistics/state-funded-schools-inspections-and-outcomes-as-at-31-december-2023/main-findings-state-funded-schools-inspections-and-outcomes-as-at-31-december-2023) (Accessed June 2024).

Ofqual (2022) *Infographics for GCSE results, 2022*. Available at:

<https://www.gov.uk/government/publications/infographic-gcse-results-2022/infographics-for-gcse-results-2022-accessible> (Accessed September 2023).

Ofqual blog (2018). *GCSE grades 9 to 1: a Guide for Parents*. Available at:
<https://ofqual.blog.gov.uk/2018/03/02/gcse-9-to-1-grades-a-brief-guide-for-parents/>
(Accessed May 2024).

ONS (Office for National Statistics) (2019) *Long-term Trends in UK Employment, 1861 – 2018*
Available at:
<https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/compendium/economicreview/april2019/longtermtrendsinukemployment1861to2018> (Accessed April 2023)

Palinkas, L., Horwitz, S., Green, C., Wisdom, J., Duan, N. & Hoagwood, K. (2015) 'Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research', *Administration and Policy in Mental Health and Mental Health Services Research* Vol.42(5) pp.533-44.

Parliamentary Business (2006) *Brief History of Special Educational Needs: Select Committee on Education and Skills: Third Report*. Available at
<https://publications.parliament.uk/pa/cm200506/cmselect/cmeduski/478/47805.htm>
(Accessed April 2023).

Partridge, J. (1965). *Life in a Secondary Modern School*. London: Pelican Books.

Patton, M. (1980). *Qualitative evaluation methods*. Thousand Oaks, CA: Sage Publications.

Perkins, D. (2008). *Making learning whole: How seven principles of teaching can transform education*. San Francisco, CA: Jossey-Bass.

Peters, R.S. (1983) 'Philosophy of Education' in Hirst, P. (ed) *Education theory and its Foundation Disciplines*. (pps 30 – 61). London: Routledge & Kegan Paul.

Plowden, B. (1967) *Children & their Primary Schools (The Plowden Report)* London: HMSO.

Potter, D (2004) *The Role of Federations in Turning Round the Weakest Schools*. Academies Division, DfE.

Poulos, C. (2021) *Essentials of Autoethnography*. New York: American Psychological Association.

Pring, R. & Roberts, M. (2016) *A Generation of Radical Educational Change*. London: Routledge.

Reynolds, D. (1976). *The delinquent school*. In Hammersley, M. & Woods, P.(Eds.), *The Process of Schooling* (pp. 217 – 229). London: Routledge and Kegan Paul.

Reynolds, D. (2007) *Schools Learning from their Best: the Within-School Variation Project* Nottingham: National College for School Leadership. Available at: <https://dera.ioe.ac.uk/7381/1/download%3Fid=17377&filename=schools-learning-from-their-best.pdf> (Accessed April 2023).

Reynolds, D. (2023). *Education Without Failure: Some Practical Proposals for Educational Change*. Speech to the Comparative and International Education Society Conference, Florida, 22nd March 2023. Available at: https://www.youtube.com/watch?v=iFRcPe3ep8A&ab_channel=ComparativeandInternationalEducationSociety (Accessed November 2023).

Reynolds, D., Chapman, C., Kelly, A., Muijs, D. & Sammons, P. (2012) 'Educational effectiveness: the development of the discipline, the critiques, the defence, and the present debate', *Effective Education* Vol 3(2), pp.1–19.

Reynolds, D. & Hopkins, D. (1993) 'World Class school improvement: an Analysis of the Implications of Recent International School Effectiveness and School Improvement Research for Improvement Practice', in A. Hargreaves, A. Lieberman, Fullan, M, and Hopkins, D. (Eds.) *International Handbook of Education Change*. London: Springer Publishing, pp.1275-1285.

Reynolds, D., Potter, D. and Chapman, C. (2003) 'A Review of Research and Practice', in B. Davies and J. West-Burnham (Eds.) *Handbook of Education Leadership and Management*. London: Pearson-Longman, pp.542-550.

Reynolds, D., Sammons P., Frayne, B., Van Damme, J., Townsend, T., Teddlie, C. and Stringfield, S. (2014) 'Educational effectiveness research (EER): a state-of-the-art review', *School Effectiveness and School Improvement* Vol 25(2), pp.197 – 230.

Reynolds D, Stringfield S & Schaffer G (2006) *The High-Reliability Schools Project*, in Harris A & Chrispeels J (Eds) *Improving School s and Educational Systems*. London, Routledge.

Richards, C. & Norris, F. (2024) *The Alternative Big Listen. Main Findings on School Inspection*. Available at: <https://iHH1.wsiHH.com/blobby/go/d5727bbc-b57f-49db-9752-91d0efb69788/downloads/ABL%20Main%20Findings%20110724PDF.pdf?ver=1725085134965> (Accessed September 2024).

Richardson, L. (2000) *Writing: A method of inquiry*, in Denzin, N. and Lincoln, Y. (Eds) *The Sage Handbook of Qualitative Research*, 2nd edition. Thousand Oaks CA and London: Sage Publications. pp.923–48.

Rosenshine, B. (2012) 'Principles of Instruction', *American Educator*, Spring 2012.

Rosenthal, L. (2004) 'Do school inspections improve school quality? Ofsted inspections and Examination Results', *Economics of Education Review*, Vol 23(2), April 2004, pp.143-151.

Rowley, J. & Slack, F. (2004) 'Conducting a Literature Review', *Management Research*, Volume 27(6), pp.31-39.

Rudduck, J. (1991). *Innovation and Change: Developing Involvement and Understanding*. Buckingham: Open University Press.

Rutter, M., Maughan, B., Mortimore, P. and Ouston, J. (1979) *Fifteen Thousand Hours: Secondary Schools and their Effects on Children*. London: Open Books.

Ryan, G. (2018) *Introduction to positivism, interpretivism and critical theory*. Open University Online Study Guide. Available at: <https://oro.open.ac.uk/49591/17/49591ORO.pdf> (Accessed May 2024).

Sahadewo, G.A. (2023) 'School quality and labour market earnings', *Journal of Applied Economics*. Vol 26,(1), pp. vol. 26(1), pages 216.

Sallis, E. (1993) *Total Quality Management in Education*. London: Kogan Page.

Sammons, P. (2007). *School Effectiveness and Equity: Making Connections. A review of school effectiveness and improvement research*. Reading, UK: CfBT Education Trust.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key Characteristics of Effective Schools: A Review of School Effectiveness Research*. London: Institute of Education, University of London.

Sandbrook, D. (2012) *Seasons in the Sun: The Battle for Britain, 1974 – 79*. London, Allen Lane.

Saunders M, Lewis P, Thornhill A, Bristow A (2019) *Research Methods for Business Students: Chapter 4: Understanding research philosophy and approaches to theory development*. London: Pearson Education.

Savin Baden, M. & Major, C. (2013) *Qualitative Research: The Essential Guide to Theory and Practice*. London: Routledge.

Scanlon, M. (1999) *The Impact of Ofsted inspections*. NUT/NFER, Slough.

Scheerens, J. and Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.

Schluter, D. (2000). *The Ecology of Adaptive Radiation*. Oxford: University Press.

Scott, J. & Clery, E. (2013) *British Social Attitudes 30: Gender* in Park, A., Bryson, C., Clery, E., Curtice, J. and Phillips, M. (eds.), *British Social Attitudes: the 30th Report*, London: National Centre for Social Research. Available at:

https://www.bsa.natcen.ac.uk/media/38457/bsa30_gender_roles_final.pdf (Accessed April 2023).

Soiferman, L.K. (2020) *Compare and Contrast Inductive and Deductive Research Approaches* in *Inductive and Deductive Research Approaches*. Winnipeg: University of Manitoba.

Spouse, J. (2001) *Intervention for Learning: Supporting School Improvement*. University of Gloucestershire; PhD thesis.

Spry, T. (2001). 'Performing autoethnography: An embodied methodological praxis', *Qualitative Inquiry*, Vol.7(6), pp.706-732.

Stenhouse, L (Ed) (1980) *Curriculum Research and Development in Action* London: Heinemann Educational Books.

Stoll, L. and Myers, K. (Eds) (1988) *No quick fixes: perspectives on schools in difficulty*. London: Falmer Press.

Stringfield S & Teddlie C (1988) 'A Time to Summarize: The Louisiana School Effectiveness Study', *Educational leadership: Journal of the Department of Supervision and Curriculum Development*. Vol 46 pp. 43-49.

Stringfield, S. & Teddlie, C. (2011) *School effectiveness research, 1932 to today, including a call for future research*. London: Routledge

Stubbs, B. (2016) *National figures for pupils gaining 5+ A*-C GCSE grades*. Available at: <http://www.bstubbs.co.uk/5a-c.htm> (Accessed April 2024)

Stubbs, B. (2023) *Student Performance Analysis: National Percentage Figures for UK A-Level Grades*. Available at: <http://www.bstubbs.co.uk/a-lev.htm> Accessed March 2024.

Sutcliffe-Braithwaite, F. (2018) *Class, Politics, and the Decline of Deference in England, 1968 – 2000*. Oxford: OUP.

Sutton Trust Website 'About Us'. Available at <https://www.suttontrust.com/about-us/> (Accessed March 2023).

Tennison, J. (2020) *How does Ofqual's Grading Algorithm Work?* Available at: <https://rpubs.com/JeniT/ofqual-algorithm> Accessed March 2024.

Thatcher, M. (1971) Secretary of State's Speech to the North of England Education Conference, Buxton. January 6th, 1971.

Timpspn, J. (2024) Televised interview with James O'Brien. Available at: https://www.youtube.com/watch?v=pOlF0pLKB1Y&ab_channel=LBC (Accessed Oct.2024)

TIMSS (Trends in International Maths and Science Study) (1995) *TIMSS 1995: Highlights of Results* Available at: <https://timssandpirls.bc.edu/timss1995i/HiLightB.html> (Accessed April 2023).

Toye, R. (2013). 'From 'Consensus' to 'Common Ground': The Rhetoric of the Postwar Settlement and its Collapse', *Journal of Contemporary History*, Vol.8(1), pp.3–23.

UK Parliament (2024) *Teacher Recruitment, Retention and Training*. Available at: <https://publications.parliament.uk/pa/cm5804/cmselect/cmeduc/119/report.html> (Accessed June 2024)

University of Gloucestershire (2020) *Research Ethics Handbook* Feb 2020 Last updated: 12 February 2021. Available at: <https://www.glos.ac.uk/information/knowledge-base/research-ethics-handbook-feb-2020/> (Accessed August 2022)

UTC (University Technical Colleges) Website (2023). Available at:
[https://www.utcolleges.org/#:~:text=University%20Technical%20Colleges%20\(UTCs\)%20are,employers%20in%20the%2021st%20CenturyUTr](https://www.utcolleges.org/#:~:text=University%20Technical%20Colleges%20(UTCs)%20are,employers%20in%20the%2021st%20CenturyUTr) (Accessed November 2023).

Walker, G. (1963) 'Englewood and the northern dilemma', *The Nation*, 197, pp.7-10.

Warnock, M. (1978) *The Warnock Report: Special Educational Needs. Report of the Committee of Enquiry into the Education of Handicapped Children and Young People*. London: Her Majesty's Stationery Office.

Webb, J. (2022) *The Gift of a Radio*. London: Penguin Random House.

Weber, G. (1971) *Inner-City Children Can be Taught to Read: Four Successful Schools*. Washington D.C.: Council for Basic Education.

Welsh Government (2022) *School improvement guidance: framework for evaluation, improvement and accountability*. Available at: <https://hwb.gov.wales/evaluation-improvement-and-accountability/school-improvement-guidance-framework-for-evaluation-improvement-and-accountability> (Accessed Jan 2024).

Wilson, H. (1963) Speech to the Labour Party Conference, Scarborough, October 1963. (Generally known as the "White Heat of Technology" speech).

Wilson, J. (2010) *Essentials of Business Research: A Guide to Doing Your Research Project*. Thousand Oaks, CA: SAGE Publications.

World Bank/Government of Samoa (2005) *School Improvement Manual* Available at: http://wbgfiles.worldbank.org/documents/hdn/ed/saber/supporting_doc/EAP/Samoa/SAA/MESC_2006_School_Improvement_Plan_manual.pdf (Accessed January 2023).

Wyatt, I., & Hecker, D. (2006) *Monthly Labour review* March 2006. Washington DC: Bureau of Labour Statistics.

YouGov (2021) *Pass or Fail? The current state of the English school system according to parents*. Published June 2021. Available at: <https://yougov.co.uk/society/articles/36233-current-state-english-school-system> (Accessed March 2024).

Young, M.D. (1958) *The Rise of the Meritocracy*. London: Thames & Hudson.

APPENDICES

I – V: SEMI-STRUCTURED INTERVIEWS

The main research evidence came from semi-structured interviews with 32 people. Information about these interviews and the selection of interviewees appears in the Methodology chapter; the additional information provided as Appendices comprises:

Appendix I: Information provided to potential interviewees within a letter of invitation to participate

Appendix II: Informed consent form.

Appendix III: The list of questions around which discussions were structured.

Appendix IV: The list of interviewees, their identify concealed by the attribution of letter codes, together with a one-line c.v. to show their major role(s).

Appendix V: Sample post-interview letter/email of thanks to interviewees

6: **.APPENDIX VI** provides the Coding Tree, described in Section 3.2.4.

APPENDIX I: INFORMATION TO POTENTIAL INTERVIEWEES; INVITATION TO PARTICIPATE

UNIVERSITY OF GLOUCESTERSHIRE

(date)

Dear

Thank you for your preliminary agreement to take part in an interview to contribute to my doctoral research at the University of Gloucestershire. This letter explains:

- more about the research topic
- arrangements for the interviews
- guarantees of confidentiality/anonymity and your right to withdraw
- and seeks your informed consent to take part.

The title of the thesis is *Improvement in What?* and is designed to further our understanding of three aspects of school improvement in England:

- the combination of external (socio-econo-political) and internal (educational) factors that led education to be dominated by school improvement during the period from 1970 to 2020, and which promoted the rapid evolution of the school improvement paradigm;
- what key groups of educators (researchers, policy makers, school leaders) have understood by 'improvement', and the actions they have taken in response;
- the impact of fifty years of school improvement, and implications for the future.

Interviews will take about one hour, and you can choose where and how (face-to-face or by video conference). I intend to record interviews (but see the consent form below).

Taking part is voluntary. If you decide to participate, you are free to withdraw from the study, without stating a reason, at any time, including up to a week after the date of the interview. Any information you give will be anonymous and be treated confidentially. If respondents are quoted in the thesis, attribution will be to "a leading researcher/school leader said ..", not a named person. Data will be kept securely and will be destroyed after the study. The University of Gloucestershire School of Education and Humanities research ethics panel has reviewed and approved this study. Here you can view the [University's Research Participants' Privacy Notice](#).

Would you please place a Y or N in each of the boxes on the attached consent form and return it by email to me? If you any queries, please contact me. Thank you in advance for your help.

(David Potter)

APPENDIX II: INFORMED CONSENT FORM

Attached to 1.1, completed by participants.

IMPROVEMENT IN WHAT?

Informed consent

Please tick boxes

1. I confirm that I have read and understood the information provided for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time up to a week after the interview, without giving any reason. ☐
3. I agree to take part in the above study. ☐
4. I give permission to be recorded during the interview. ☐
5. I give permission for quotations from the interview to be used anonymously in any report or publication from this work. ☐

_____	_____	_____
Name of Participant	Date	Signature

_____	_____	_____
Name of Researcher	Date	Signature

APPENDIX III: INTERVIEW QUESTIONS

Issued in advance to participants; these were the questions around which discussions were structured.

1. I am interested in differing views of what 'school improvement' is: what do you understand by the term? What is to be improved?
2. Where does your understanding come from, and has it changed?
3. From your experience, how do you believe that improvement happens in schools?
4. Can you point to any key moments/critical events in your own involvement with school improvement? What you might call 'flashbulb moments', moments when to quote Alan Bennett, "history rattled over the points"?
5. How successful do you consider school improvement to have been?
6. What of the future - more of the same, or are there limits on improvement? Changing focus?

APPENDIX IV: INTERVIEWEES: CODES & ONE-LINE CVs

Interviewees are identified in the text by coded initials and sometimes by their key role. To show the richness of experience of the interviewees, I have included a one- line CV for each.

AA	Headteacher
BB	Headteacher multiple times, system leader
BC	School senior leader
CC	Headteacher multiple times, MAT CEO, national government officer, researcher, school improvement consultant, system leader
CD	School senior leader
DD	Headteacher multiple times, MAT CEO, system leader
DE	School senior leader
EE	Headteacher multiple times, MAT CEO, system leader, national government official, school improvement consultant
EF	School senior leader
FF	Headteacher multiple times, MAT CEO, system leader
FG	Researcher, senior local government officer
GG	Headteacher, local government officer, researcher
GH	Headteacher, local government officer, HMI, school improvement consultant
HH	Headteacher multiple times, MAT CEO, system leader
HJ	Headteacher, Ofsted inspector, school improvement consultant
JJ	Local government officer, Ofsted inspector, school improvement consultant
JK	School senior leader
KK	Senior national government officer, MAT chair, researcher
LL	School senior leader, national government officer, MAT chair, school improvement consultant, system leader
MM	Researcher, senior national government officer
NN	Headteacher, local government officer, school improvement consultant, Ofsted inspector
PP	Headteacher, MAT CEO, system leader
QQ	Researcher, senior national government officer
RR	Headteacher multiple times, MAT CEO, system leader, researcher
SS	Headteacher, Ofsted inspector, national government officer, MAT chair, system leader
TT	Headteacher
UU	Headteacher, multiple times
VV	Headteacher, system leader
WW	Senior local government officer, Ofsted inspector, school improvement consultant
XX	Headteacher, multiple times, system leader, school improvement consultant
YY	Headteacher, MAT CEO, system leader
ZZ	School senior leader, Ofsted inspector

- 32 interviewees, of whom one third (11) are female
- 90% have been a senior school leader; two thirds (22) have significant system leadership experience

APPENDIX V: SAMPLE POST-INTERVIEW LETTER/EMAIL OF THANKS TO INTERVIEWEES

Dear

Many thanks once again for your time on ...day. Our discussion will be of great value in my doctoral thesis, especially our discussion of

Let me repeat your right to withdraw, either in full or any comment you made, if you think it would be unhelpful or damaging.

With best wishes

Etc.

APPENDIX VI: CODING TREE

The process of data analysis (Sec 3.2.4) is shown here in diagrammatic form: marginal notes/memos on the interview transcripts were drawn together initially into codes (column 2); these were amalgamated into categories/sub-themes (column 3) and then into broad themes (column 4): an iterative process.

No.	CODES	CATEGORIES & SUB-THEMES	THEMES
1	Watershed 1980		
2	Expansion of research	)	
3	Rapid legislative change	) Dominance of Improvem't	
4	Specialist SE/SI research/researchers	)	
5	Ofsted impact	)	
6	Opening schools to scrutiny	)	
7	HT accountability	) School accountability	DOMINATION OF EDUCATION BY THE THEME OF IMPROVEMENT POST-1980
8	Agency of schools	)	
9	Dominance of test/exam results	)	
10	Continuous change	) Speed of change	
11	Relentless pace	)	
12	Safeguarding new dominant	)	
13	Mental health pupil/community	) Rise of welfare agenda	
14	Pastoral support	)	
15	Expansion of HT responsibilities	)	
16	Social outcomes: behaviour	)	
17	Ditto: attitudes	) Link between academic &	
18	Ditto: attendance	) PD development	
9	Dominance of test/exam results	)	
19	Refined measures of academic outcome	)	
20	Improvements in teaching	)	EVOLUTION OF THE MEANING OF 'IMPROVEMENT'
21	Improvements in curriculum	) Improvements in inputs	
22	Improvements in leadership	)	
23	Artificial improvement	Gaming	
38	Tension good school/improvement		

No.	CODES	CATEGORIES & SUB-THEMES	THEMES
24	Turn-round	)	
25	Un/constrained schools	) Dimensions of improvement	
26	Improvement trajectory	)	
27	Collective purpose/Individual responsibility	)	
28	Need for change	)	
29	School culture	) HT leadership	
30	Focus	)	
31	Team building	)	CONDITIONS FOR SCHOOL IMPROVEMENT
32	SE principles	)	
33	Centrality of teaching	) Principles/systems	
20	Improvements in teaching	)	
34	Consistency	)	
35	LA support	)	
36	MAT support	) External support	
37	Govmt/Ofsted support	)	
38	Tension good school/improvement	)	SIGNIFICANCE OF BASIC IMPROVEMENT OVERLOOKED
39	Schools as humane places	)	
10	Continuous change	) Speed of change	
11	Relentless pace	)	
40	Decline of trad industries	)	
41	International competition	) Economic factors	
42	Knowledge economy	)	
43	Neo-liberalism	)	EXPLANATION OF RAPIDITY OF EDUCATIONAL CHANGE;
44	Market replaces welfare state	) Political factors	CLASSIFICATION OF FACTORS
3	Rapid legislative change	)	
45	Decline of deference	)	
46	Affluence and choice	) Social factors	
47	Permissive society	)	
48	Ruskin	)	
49	ROSLA	) Educational factors	
50	Comprehensivisation	)	
51	School failures	)	