

UNIVERSITY OF WINCHESTER

How are outcomes for learners judged in the further education and skills sector?

Robert David Rees

ORCID Number: 0000-0002-5897-6677

Doctor of Education

August 2018

This thesis has been completed as a requirement of a postgraduate research degree of the
University of Winchester.

Declaration and Copyright Statement

Declaration:

No portion of the work referred to in this Thesis has been submitted in support of an application for another degree or qualification of this or any other university of other institute of learning.

Copyright Statement:

Copyright © 2018 Robert David Rees

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the author. Details may be obtained from the RKE Centre, University of Winchester.

This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instruction may not be made without the permission (in writing) of the author).

No profit may be made from selling, copying or licensing the author's work without further agreement.

Dedications and Acknowledgements

The fact that I am writing this acknowledgement at the beginning of my doctoral thesis is testament to the fact that education changes lives.

My own educational journey would have been doomed from the start if it hadn't been for three inspirational teachers. I left school with one GCSE in music, which was only achievable because of Ms Linet Purcell (née Willans). Ms Purcell was enthusiastic, an encourager and supportive and would let me spend all my spare time practising in the music rooms. She clearly loved teaching, so much so, that it was she, without knowing, who inspired me to teach.

Beth Thomas and Ann-Marie Marks were my A-level teachers. As I was someone with one GCSE presenting myself to the local college of further education, they had every right not to accept me on to their courses. If that was me today, I am sure I would be on a foundation level study programme, resitting GCSE English and mathematics. Because of them, their patience, their understanding, the extra time they spent with me and the enthusiasm they each had for their subject, I achieved two A levels and was the first in my family to go to university. All three teachers took a risk on me and without them, I would not have had the brilliantly successful career that I have enjoyed so much, or the business I now have.

Completing my study would not have been possible without the support of my partner, Peter. His support and encouragement have been essential, especially in the times when I could no longer think clearly or on the days where I would delete more words than I wrote. And, as my business partner, he agreed that my EdD should count as one of our clients, so instead of this being an evening and weekend study, I was, thanks to Peter, doing this as part of my job.

Finally, I owe a great deal to my supervisors, Dr Bridget Egan and Dr Kerry Ball at the University of Winchester. They have provided me in equal measure with encouragement, intellectual challenge and knowledge of research methods that I will always aspire to match. This powerful combination has sustained me throughout all phases of this research and has challenged me to read and think far more widely. They have even talked about future research – such aspiration!

UNIVERSITY OF WINCHESTER

ABSTRACT

How are outcomes for learners judged in the further education and skills sector?

Robert David Rees

ORCID Number: 0000-0002-5897-6677

Doctor of Education

August 2018

Abstract

My study analyses how inspectors use success rate, value added, attendance, progression and destination data to evaluate student performance and to make judgements on outcomes for learners in further education colleges. Ofsted (2017) acknowledges that there is little empirical evidence about the validity of inspection judgements. My previous experience suggests that inspectors approach and interpret data inconsistently. To date, hardly any research into this aspect of inspection has been published. I analysed 165 further education college inspection reports from inspections that took place between September 2012 and July 2015, along with the publicly available success rate data the inspectors would have used to evaluate student performance. For aspects where data is not collected centrally, this was obtained directly from colleges within the scope of this study. Data are analysed using approaches that reflect the ways in which Ofsted uses numerical information. My findings indicate that there is (often considerable) disparity between the success rate data and the judgements made by inspectors, and that the use of the national success rate average as a foundation from which student performance is assessed is inconsistent. I further suggest that, because most of the data (other than success rates) is generated by the individual colleges in the absence of any national guidance, inspectors are not able to compare one college with any other objectively. Moreover, not all inspection reports make judgements to the same criteria, with too many reports missing key judgements. This evidence indicates that there is a lack of transparency about which student groups are included or excluded from the data used by inspectors to inform their judgements. I therefore conclude that the consistency with which inspectors use data and information to evaluate student performance and make judgements on outcomes for learners, 'requires improvement', and suggest how this situation might be improved.

List of Contents

Title	Page
Declaration and Copyright Statement	2
Dedications and Acknowledgements	3
Abstract	4
List of Contents	5
List of Tables	8
Chapter 1: Introduction, Background and Context	
1.0 Introduction	11
1.1 Introduction, background and context	14
1.2 The development of further education and an introduction to the college sector in England	15
1.3 The further education college sector in England and its provision within a system	24
1.4 Further education policy: learning, confusion or amnesia	26
1.5 Ofsted: the regulator and inspector	28
1.6 Outcomes as a default for the standard of teaching and learning	30
Chapter 2: Literature Review	
2.0 Introduction	33
2.1 Inspection and emotion	35
2.2 Performing for inspection	41
2.3 Accountability and inspection	44
2.4 Quality improvement and inspection	48
2.5 The reliability of inspection: judgement and improvement	52
2.6 Pan-European models of inspection	56
2.7 Conclusion	67
Chapter 3: Research Methodology	
3.0 Introduction	70
3.1 Researcher's philosophical position	71
3.2 The rationale for choosing the timeframe and sample	72
3.3 Hypothesis testing	74
3.4 Gaining access to data and information	76
3.5 Inspection reports	78

3.6 How are the reports analysed for student success?	79
3.7 Success rates	85
3.7.1 What are success rates	85
3.7.2 Where has the success rate data come from	89
3.7.3 How are success rates attributed to inspection	89
3.8 The national average	90
3.9 Accessing non-publicly available data	91
3.10 Attendance	92
3.11 How are the reports analysed for attendance	93
3.12 Value added	94
3.13 How are the reports analysed for value added and the progress made by students?	95
3.14 Progression data	96
3.15 How are the reports analysed for progression	97
3.16 Destination data	97
3.17 How are the reports used to analyse destinations	98
3.18 Limitations of the methods used in this research	99
3.19 Sampling	100
3.20 Measurements	101
3.21 External validity	101
3.22 Conclusion	102

Chapter 4: Presentation of Data

4.0 Introduction	104
4.1 The sample	105
4.2 Success rates	106
4.2 Inspection cycles	108
4.4 Success rates, textual analysis and national average for students aged 16-18	109
4.4.1 Analysis	109
4.5 Success rates, textual analysis and national average for students aged 19+	112
4.5.1 Analysis	113
4.6 Success rates, textual analysis and national average for WBL students (apprentices)	116
4.6.1 Analysis	117
4.7 Student attendance: college data, textual analysis and its use in inspection	120
4.8 Value added: college data, textual analysis and its use in inspection	124
4.9 Progression: college data, type of data and textual analysis	129
4.10 Destination: college data, type of data and textual analysis	131
4.11 Conclusion	134

Chapter 5: Discussion

5.0 Introduction	135
5.1 How consistently are data and information used by inspectors to evaluate and make judgements on student success?	136
5.2 How is attendance evaluated by inspectors? Is the evaluation consistent across inspections?	139
5.3 To what extent is value added used by inspectors to evaluate and make judgements on student progress?	143
5.4 How are progression and destination evaluated by inspectors and is this evaluation consistent across inspections?	147
5.5 How transparent are inspectors about the type and validity of data and information used when evaluating, judging and reporting on outcomes for learners?	151
5.6 Implications of the findings	154

Chapter 6: Conclusion and Recommendations

6.0 Introduction	156
6.1 Conclusion	156
6.2 Recommendations	158
6.3 Contribution to research	166
6.4 Limitations of the research	167
6.5 Identification of future research areas	168
6.5.1 How are data and information used by inspectors to evaluate student performance and make judgements on outcomes for learners?	168
6.5.2 Short inspection: what relationship is there between student performance and the grade awarded to a college?	170
6.5.3 How important are English and mathematics achievement rates when evaluating student performance in further education colleges?	171
6.5.4 What is the relationship between student performance and the timing of an inspection in the further education sector?	172
6.5.5 Do students who attend better, achieve better?	174
6.5.6 Outcomes for learners – an example	175
6.6 Personal learning	176

Bibliography	178
---------------------	-----

Appendices

Appendix 1: Textual Analysis Codes	193
Appendix 2: Freedom of Inspection Request, letter	198
Appendix 3: Freedom of Inspection Request, data table	199
MPhil/PhD THESES OPEN ACCESS / EMBARGO AGREEMENT FORM	201

List of Tables

Chapter 1

Table 1.1	The development of further education in England	p21
-----------	---	-----

Chapter 2

Table 2.1	European inspection regimes similar to that in England	p57
-----------	--	-----

Table 2.2	European inspection regimes that differ to the one in England	p58
-----------	---	-----

Table 2.3	European countries with no process of external evaluation of school standards (inspection)	p59
-----------	--	-----

Chapter 3

Table 3.1	An illustration of how success rates are calculated, as per SFA Guidance	p74
-----------	--	-----

Table 3.2	The data range of colleges inspected against the documentary analysis (Code 6): The success rates for 16-18 year olds is above average/high	p75
-----------	---	-----

Table 3.3	The data range of colleges inspected against the documentary analysis (Code 33): The success rates for 16-18 year olds is in line with the national average or just above/below it	p75
-----------	--	-----

Table 3.4	The data range of colleges inspected against the documentary analysis (Code 60): The success rates for 16-18 year olds is below the national average, too low or poor	p75
-----------	---	-----

Table 3.5	The coding framework for the analysis of inspection reports regarding the success rates for students	p81
-----------	--	-----

Table 3.6	Examples of how inspection report texts regarding student success rates have been allocated a code for analysis	p82
-----------	---	-----

Table 3.7	Examples where a textual analysis code cannot attributed with regards to student success rates	p83
-----------	--	-----

Table 3.8	The codes used in the analysis of success rates judgements for students aged 16-18 and 19+	p84
-----------	--	-----

Table 3.9	The codes used in the analysis of success rate judgements for apprenticehips	p84
-----------	--	-----

Table 3.10	An illustration of how success rates are calculated, as per SFA guidance	p85
------------	--	-----

Table 3.11	Qualification levels and descriptions	p87
------------	---------------------------------------	-----

Table 3.12	Changes to the success rate methodology 2013/14	p80
Table 3.13	The relationship between QSR report and both the calendar and academic years of inspection	p90
Table 3.14	Description of attendance judgements found in inspection reports	p93
Table 3.15	Description of value added judgements found in inspection reports	p95
Table 3.16	Description of progression judgements found in inspection reports	p97
Table 3.17	Description of destination judgements found in inspection reports	p99

Chapter 4

Table 4.1	Total number of inspections in scope, by academic year and inspection grade for OfL	p106
Table 4.2	Total number of responses received from the FoI, broken down by geographic area	p106
Table 4.3	The data range of colleges inspected against the documentary analysis (Code 6): The success rates for 16-18 year olds is above average/high	p110
Table 4.4	The data range of colleges inspected against the documentary analysis (Code 33): The success rates for 16-18 year olds is in line with the national average or just above/below it	p111
Table 4.5	The data range of colleges inspected against the documentary analysis (Code 60): The success rates for 16-18 year olds is below the national average, too low or poor	p111
Table 4.6	The data range of colleges inspected against the documentary analysis (Code 87): The success rates for 19+ year olds is above average/high	p114
Table 4.7	The data range of colleges inspected against the documentary analysis (Code 114): The success rates for 19+ year olds is in line with the national average or just above/below it	p114
Table 4.8	The data range of colleges inspected against the documentary analysis (Code 141): The success rates for 19+ year olds is below the national average, too low or poor	p115
Table 4.9	The data range of colleges inspected against the documentary analysis (Code 165). The success rate for WBL is above the national average/High	p117

Table 4.10	The data range of colleges inspected against the documentary analysis (Code 174). The success rate for WBL is in line with the national average or is just above/below	p118
Table 4.11	The data range of colleges inspected against the documentary analysis (Code 183). The success rate for WBL is below the national average too low or poor	p119
Table 4.12	Total number of responses received from the FoI regarding	p121
Table 4.13	Presentation of attendance data by academic year, documentary analysis and the type of attendance data	p122
Table 4.14	FoI responses on what type of value added tool was used by the college at the point of inspection (V1Y)	p126
Table 4.15	FoI responses on whether the inspection team explored value added data and information	p126
Table 4.16	FoI responses on how much weight was placed on value added as a tool for measuring student outcomes	p127
Table 4.17	Documentary analysis of judgements made regarding Level 3 Value Added/Progress Measure	p127
Table 4.18	Analysis of Level 3 value added judgements by OfL grade	p128
Table 4.19	College responses to the type of data used by the inspection team to assess the standards of student progression	p129
Table 4.20	Presentation of progression data by academic year and documentary analysis	p130
Table 4.21	College responses to the type of data used by the inspection team to assess the standards of student destination	p132
Table 4.22	Presentation of destination data by academic year and documentary analysis	p133
Chapter 5		
Table: 5.1	Ratio of inspectors:students regarding direct observation of student progress	p146

Chapter 1: Introduction, Background and Context

1.0 Introduction

I have worked in the further education college sector for most of my working life – as a teacher, middle manager and senior manager. Whilst in these roles I have seen at first-hand how young people and adults, many of whom failed to reach their potential in school, achieve qualifications and develop skills that will prepare them well for their future life; be the first in their family to go to university and, as a result, have a successful career. Indeed, this is me, this is my life story.

The further education system in England is complex. The further education college sector is ever changing and colleges themselves, often catering for thousands or in some cases tens of thousands of students every year, are the invisible institutions of the educational landscape or, to use a well referred to quote, further education is the Cinderella sector (Gravatt, 2004). However, a college becomes visible when it is inspected. Whether the result is good or otherwise, it is likely to be reported in the local press: if it is good, the college will use it for self-promotion and publicity and if the result is not so good, it is reported as a news story. If the result is remarkable, its coverage might also extend to the national media and industry trade papers. It cannot be overestimated how considerable the impact of an inspection is, on a college's reputation, on its ability to generate income and indeed on its potential to influence whether it closes, merges or expands.

Ofsted despite being in existence for 25 years, acknowledges that there is a perception that assessments made by inspectors across all sectors are too often unreliable and it is surprising that it also acknowledges that there is little empirical evidence about the validity of inspection judgements (Ofsted 2017). This lack of evidence did not stop the then Ofsted Chief Inspector stating in a newspaper article that Further Education (FE) was 'inadequate at best' (McInerney, 2016a). My study, therefore, aims to assess the consistency and accuracy with which inspectors use data to evaluate student performance and make judgements on outcomes for learners.

I have acted as a college nominee for Ofsted inspections in several large colleges in London and in the South West. I have led on, supported or advised others on approximately 16 inspections, including three full college inspections. On each occasion, the data requested by

the lead inspector has varied from that outlined in the handbook to a position where much more data was produced than officially required. Furthermore, the way that the data was used and interpreted was variable, especially the data produced by the college, which unlike success rates are not subject to any external benchmarking. And it is these experiences that have given rise to this study, to identify whether my experiences were coincidental, isolated examples of inconsistency or in fact representative of the college sector as a whole.

As a result, I have in this study analysed 165 college inspections which took place between September 2012 and July 2015. This covers one complete inspection cycle. I have sought and catalogued success rate data that would have been used by inspectors at these inspections and where only locally produced data existed, I requested this information directly from colleges.

Chapter 1: the remainder of this chapter is dedicated to placing the further education system and college sector into a national context. It explores, in some detail, what further education is and what it does and it outlines how it is impacted by continuous and repetitive policy development. Inspection and regulation and their relationship to the further education college sector are also explored and tracked in this chapter, outlining the four inspection regimes and four inspection bodies to which colleges, since incorporation, have been beholden. This chapter concludes by introducing some key themes explored in later chapters, starting with the notion that outcomes for learners act as a default for the standard of teaching and learning.

Chapter 2: this chapter concerns itself with reviewing current literature on inspection. What is acknowledged early on in this chapter is that research on the further education system generally is limited, and on inspection within the further education college sector, it is even more so. Therefore, much of the literature reviewed is based on research from the schools sector, both nationally and internationally. The most common themes associated with an inspection are explored such as: emotion, performing (for an inspection), accountability and quality improvement.

Chapter 3: begins with a philosophical position and a detailed account of a set of experiences which led to the existence of this study, along with results of the hypothesis testing. The remainder of this chapter is dedicated to the research methods used to outline the scope of my study and the timeframe which it covers. It also identifies and explores the

instruments used or produced by/for inspectors, such as success rate data, locally produced data and the inspection reports and how they have been analysed for use in this study. This chapter concludes by examining the research limitations.

Chapter 4: this data analysis chapter identifies how the data and information used by inspection teams relate to the judgements made by them because of both textual and statistical analysis. Each aspect under the theme of outcomes for learners is analysed and where there are discrepancies or apparent inconsistencies, they are explored in detail. Where there is nationally available data, such as success rates, inspectors' evaluations and judgements are analysed against the national average. Where the data has been provided directly from colleges, the same principles of analysis apply, although it is acknowledged that there are limitations with this dataset, not just for use in this study but also for inspectors.

Chapter 5: this discussion chapter examines the results based on the data in Chapter 4 and uses as its structure the research questions outlined in Chapter 2. In addition, the content of this chapter explores these questions more broadly and suggests that there are too many examples of where the validated data used by inspectors to evaluate student performance and make judgements is inconsistent. Moreover, the notion that inspectors use un-validated, locally produced data, which limits comparability and transparency, is also explored – as are the purpose and usefulness of the inspection report itself.

Chapter 6: begins with the conclusion, which summarises the analysis in Chapter 3 and the discussion in Chapter 4 and from which six recommendations are made. These recommendations, if enacted, aim to improve the consistency, transparency and comparability of further education college inspections. The final sections of this chapter identify the contribution of my study to research, the limitations of this study, and five future studies.

1.1 Introduction, Background and Context

The Office for Standards in Education (Ofsted) have been responsible for the regulation and inspection of the further education and skills system in its entirety since 2007. In that time, there have been several changes to the inspection regime itself, as identified through the four handbooks for inspecting colleges, including a series of shifts in the re-alignment of teaching and learning as the de facto limiting grade for college inspections Ofsted 2007, 2009, 2012 & 2015). In addition, there have been numerous changes in policy affecting qualifications, programmes of study and the mandatory requirement to study English and mathematics.

Therefore, this introductory chapter places into context the landscape within which colleges of further education operate. It also aims to give some background about the development of the sector over time, how it is impacted by ever-changing policy and how its relationship with Ofsted, now as its sole inspection body, has been cultivated.

To this end, this chapter has been structured as follows:

- the development of further education and an introduction to the college sector in England
- the further education college sector in England and its provision within a system
- further education policy: learning, confusion or amnesia
- Ofsted: the regulator and inspector
- outcomes as a default for the standard of teaching and learning
- the fear of inspection: playing the game
- the reliability of inspection: judgement and improvement

1.2 The development of further education and an introduction to the college sector in England

The earliest government document which describes the constituent parts of the further education sector as it would be recognised today are found in the final report of the Department Committee's review on technical education (RCS, 1906). The three constituent elements included in the report are the provision of:

- juvenile education for young people aged 14-16 and those aged 16-18;
- adult education; and,
- technical education.

It should be noted at this stage that whilst the areas cited above are indeed broadly reflective of the further education sector in both the twentieth and twenty first centuries, there are some stark differences.

At the time of this review, whilst there is a suggestion that this review concerned itself with all young people aged 14-19, it would not be until the Fisher Act of 1918 (some 12 years later) that the school leaving age would be increased to 14. Similarly, the provision made in the Act regarding adult education and technical education more specifically was highly reliant on volunteerism - volunteerism in that there was no compulsion on students to attend (education more generally encompassing technical education) and no compulsion on local authorities to make available such provision. This theme can also be seen when looking at the impact (if any) of the 1917 review entitled 'The Report of the Consultative Committee on the Education of the Adolescent'. Whilst this review, undertaken by the Board of Education, also considered education and training in the 'post-school' stage (BoE, 1926) it would be nine years before it was published and it left unchanged any compulsion for people to attend, or for local authorities to provide such provision.

The 1944 Education Act for the first time wrote into law the government's anticipated relationship between academic, general and technical education. However, whilst it is clear about academic education in the form of grammar schools and equally as clear about universal education in secondary schools, there is far less clarity about both technical education and schools. As had happened previously, there was no compulsion on local authorities to establish technical schools within their areas and indeed Gillard (2011)

identifies that many authorities did not want to set up and maintain expensive technical schools. As a result, it would be for this reason that by 1958 technical schools provided education for just 4% of the secondary school age population (ibid).

With regards to further education the 1944 Education Act, like the 1906 report cited above, fails to give an overarching strategy or framework within which to operate; instead it lists areas of provision or cohorts that are to be targeted (Doel, 2018). Moreover, the list is a collection of functions that do not fit into either university or school-based education. They are:

- provision for people with special educational needs;
- access courses for adults;
- programmes for the unemployed; and,
- higher technical qualifications such as HNDs and HNCs.

However, what the 1944 Education Act did do, was to mandate all local authorities, by law, to provide 'adequate' vocational and non-vocational provision for students aged 16-18 and those aged 19+ studying on either a full-time or part-time basis. The implementation of the 1944 Education Act, specifically the new offer for young people post-16 and adults, would not be revisited in any form of major legislation or policy review until the late 1980s, when as part of the Education Reform Act (1988) the mechanisms by which colleges were funded began to move away from local authorities to a scheme of delegation under the newly formed Polytechnics and Colleges Funding Council. This indeed paved the way, for what is often referred to as 'Year 0' or 'the birth' of the further education sector (Doel, 2018), whereby colleges would be released from the control of local authorities and be established and accountable as incorporated institutions under the Further and Higher Education Act (1992).

The Act (1992) brings together, for the first time, an overarching framework within which the further education college sector operates. In summary, it states:

- that both the funding and regulatory functions covering the college sector will be the responsibility of the Further Education Funding Council (FEFC); and,

- that the 'new' further education sector delivers full-time education for 16-18 year-olds, part-time education and full-time education for those over 18 and provision for persons with learning difficulties

It should now be noted that, as outlined above, there has been little legislative and/or policy shift throughout most of the twentieth century. However, contemporaneously and immediately following the implementation of the Further and Higher Education Act (1992) and the formalisation of the further education sector, it has been subjected to numerous policy shifts, and reviews whose recommendations have been ignored or partially implemented and a series short-lasting reforms. Norris and Adam (2017, p 5) summarise the changes to the further education sector as follows:

- 28 major pieces of legislation related to vocational, FE and skills training;
- six different ministerial departments with responsibility for education;
- 48 secretaries of state with relevant responsibilities; and,
- no organisation surviving for longer than a decade.

For the purpose of illustration, the following pieces of legislation or government policy are some that have impacted directly on the further education sector:

Item	Summary
Education and Training for the 21 st Century – a white paper (1991)	<p>A white paper introduced by the Secretary of State for Education and Science (Ken Clark) proposed (in summary):</p> <ul style="list-style-type: none"> • a three-way national qualifications framework of academic, vocational and technical (A-Levels, GNVQs and NVQs); • to promote parity of esteem between vocational and academic qualification; • to provide opportunities for young people to reach higher levels of attainment; and, • to give colleges more freedom to expand provision and to respond to the needs and interests of their 'customers'. <p>The policy implications of this white paper were seen to be motivated by the aim of 'restricting access to A Levels and to developing an alternative for those who wanted to study full-time post-16' (Hodgson and Spours 1997, p11). This white paper would lead directly to the Further and Higher Education Act (1992).</p>
Dearing Review (1996)	<p>A review of qualifications for 16-19 year-olds commissioned by the Secretary of State for Education and Employment. Specifically, 'to consider and advise the Secretaries of State for Education and Employment and for Wales on the ways to strengthen, consolidate and improve the framework of 16-19 qualifications' (Dearing 1996, p 1).</p>

	<p>Gillard (2011) points out that the review was all but pointless, in that the Secretary of State for Education and Employment in her letter to Dearing in 1995 wrote 'our key priorities remain to ensure that the rigour and standards of GCE A Levels are maintained. As a result, and in order to balance consolidation (of what already existed) and to reform Young (1997, p 38) she states that the report's recommendations are at best 'a weak variant of a framework approach' but that it 'provides some of the conditions for moving to a unified system in the future'.</p>
Kennedy Review (1997)	<p>Learning Works: widening participation in further education was commissioned by the FEFC at the transition stage between the out-going conservative government and the incoming new labour government. The review itself was a celebration of further education – of its successes, its reach and its impact (if not widely recognised) on society.</p> <p>The recommendations made in the review focused on raising awareness of the further education sector and the contributions it could make to widening participation. And, that the sector should commit itself to working in partnership with employers, local authorities and others with the aim of widening participation.</p>
Learning and Skills Act (2000)	<p>The Learning and Skills Act set about de-coupling the funding and regulatory (inspection) roles held by the FEFC. This led to:</p> <ul style="list-style-type: none"> • the Learning and Skills Council (LSC) being established as the funder of further education, including capital development costs; • the creation of the Adult Learning Inspectorate (ALI); and, • Ofsted assuming responsibility for the inspection of provision for students aged 16-18. <p>From this point until 2007, college inspections would be carried out jointly by both ALI and Ofsted inspectors.</p>
Tomlinson Review (2004)	<p>The Tomlinson Review – <i>14-19 Curriculum and Qualifications Reform</i> was established by the Secretary of State for Education and Skills (Charles Clarke). It was an indirect response to an earlier green paper <i>14-19: extending opportunities, raising standards</i> (2002) and its subsequent white paper <i>14-19: opportunity and excellence</i> (2004)</p> <p>Its recommendations were radical, in essence proposing a new education framework similar to that of the Baccalaureate – one overarching diploma framework comprising core-learning and main learning. Within the framework a student could choose an academic, vocational or technical pathway.</p> <p>The report itself cites the key strengths of this much needed reform as including the need raise participation, to get the basics rights and strengthen vocational routes, in addition to making the system more transparent and easier to understand (Tomlinson 2004, p 5)</p>
14-19 Education and Skills - a white paper (2005)	<p>This white paper is the then government's response to the Tomlinson Review (2004). It does not take on the central tenet of the review to create an overarching diploma. Instead, it commits the government to strengthening GCSEs and A Levels whilst creating a diploma specifically targeted at young people aged 14-19 in a range of vocational subjects and ensuring that these 14 'lines of learning' would first be rolled out incrementally, starting with the first four in 2008 and that all 14 would be made a national entitlement by</p>

	<p>2015. Mansell (2010, p 16) noted that the government had predicted that some 50,000 young people would take up the option of the diploma in each year of its implementation. At the beginning of the second year, where 100,000 young people were expected to have taken up the diploma, only 36,000 had.</p> <p>However, the government in 2009 continued to promote the diploma as a viable alternative to A-Levels, vocational and technical qualifications and set out in the white paper '<i>Your Child, Your School, Our Future</i>' (2009) formal arrangements for local consortia to deliver the diploma as an option for all 14-19 year olds.</p> <p>In the government's budget of 2011 it was announced that the remaining lines of learning still to be developed would not be developed and that funding for all of the other 14-19 diplomas would cease.</p>
Education and Inspection Act (2006)	<p>The Education and Inspection Act (2006) brought the education inspection regimes (with the exception of higher education) under one organisation, Ofsted. As a result of the Act, which abolished the Adult Learning Inspectorate, all inspections and their associated frameworks would fall under the purview of HMI of Education, Children's Services and Skills, including inspections of further education colleges and training providers.</p>
Leitch Review (2006)	<p><i>The Leitch Review of Skills: prosperity for all in the global economy – world class skills</i> was commissioned by the Chancellor of the Exchequer in 2004 to 'identify the optimal skills mix for 2020 to maximise the economic growth, productivity and social justice...and consider the policy framework required to support it' (Leitch 2006, p3).</p> <p>Its main conclusions were (in summary) that by 2020:</p> <ul style="list-style-type: none"> • the majority of the adult population should achieve basic, functional skills in literacy and numeracy; • the majority of adults should be fully qualified to Level 2 (equivalent to 5 good GCSEs A* - C); • that the balance of intermediate skills should shift from Level 2 to Level 3; and, • more than 40% of adults should be qualified to Level 4 and above. <p>The two most notable changes to come out of this review were that:</p> <ul style="list-style-type: none"> • There would be greater competition among providers regarding the education and training of adults; and, • The leaving age for participating in education and/or training would be raised from 16 to 18 years old (which would be included as part of the Education and Skills Act, 2008)
Further Education and Training Act (2007)	<p>The Further Education and Training Act (2007) is a technical act detailing the changes to:</p> <ul style="list-style-type: none"> • the Learning and Skills Council (moving from local to regional offices); • foundation degree awarding powers for further education colleges; and, • such miscellaneous items such as the qualifications of principals for example.

Education and Skills Act (2008)	<p>The Education and Skills Act (2008) placed onto the statute book that all young people (from 2014) would stay on in education and/or training up to the age of 18.</p> <p>The Act itself outlines:</p> <ul style="list-style-type: none"> • the government’s duty with regards to young people; • the role of the local authority in implementing the duty; • the role of employers in implementing the duty; • the role of parents in implementing the duty; and, • attendance (failures to comply and enforcement)
Apprenticeships, Skills, Children and Learning Act (2009)	<p>The Apprenticeships, Skills, Children and Learning Act introduces a wide range of measures affecting young people aged 16-19. The Act puts into place a statutory framework that states that all ‘suitably qualified’ people aged 16-18 have the right to become an apprentice. In addition, the Act also introduces the right for employees to request time away from their duties to undertake training.</p> <p>With regards to the framework supporting the further education sector, this Act abolished the Learning and Skills Council, replacing it with two agencies:</p> <ul style="list-style-type: none"> • Young People’s Learning Agency (for those aged 16-18); and, • The Skills Funding Agency (for those aged 18+ and covering college capital development funds). <p>Importantly, the Act also states that, on the demise of the LSC, responsibility for funding education and training for 16-18 year-olds transfers (back to the local authority) – although its intent for this to happen was short lived and was never enacted.</p>
Wolf Review (2011)	<p>The Secretary of State for Education and the Minister of State for Further Education, Skills and Lifelong Learning commissioned Professor Alison Wolf to undertake a review of vocational education in England. In the report's foreword, the Secretary of State (Michael Gove) summarises that too many people aged 14-16 are studying for qualifications that have little or no value and that about one third of those aged 16-19 are studying for qualifications that offer little chance of progression to higher levels of education or to meaningful employment (Wolf 2011, p32)</p> <p>Furthermore, the report states that that the most valuable vocational skills, those of English and mathematics, are not taken seriously enough post-16 for those who have not achieved the minimum standard, set at Grade C GCSE.</p> <p>Many of the key recommendations cited within the Wolf Report were positively welcomed in the government’s response (DFE, 2011). In summary, those with a considerable impact on the further education were:</p> <ul style="list-style-type: none"> • that all young people who have not achieved a GCSE in English and/or mathematics by the age of 16 will be required to achieve it/them by the age of 19; • that all young people aged 16-18 will study within the Study Programme framework. That means that they will study for either A

	<p>Levels or a vocational qualification, alongside English and mathematics GCSE (where appropriate) and undertake mandatory work experience;</p> <ul style="list-style-type: none"> • a review of the provision for students with learning difficulties and/or disabilities; and, • an evaluation and review of the apprenticeship frameworks to identify adequacy for those apprentices aged 16-19.
Education Act (2011)	<p>With regards to further education, the Education Act (2011) is another technical act. The Act outlines the abolition of the Young People’s Learning Agency (YPLA), transferring its functions and duties to the Secretary of State, via the Department of Education’s Education Funding Agency.</p>
The Budget: Autumn Statement (2011)	<p>During the Autumn Statement (2011) the Chancellor of the Exchequer made two key announcements that affected the further education directly.</p> <p>First, it was announced (as outlined above) that there would be no funding for the development of the remaining 14-19 diplomas as set out in the 14-19 Education and Skills - a white paper (2005) and affirmed in the white paper <i>'Your Child, Your School, Our Future'</i> (2009). In fact, it was also announced that all diploma lines would cease to be funded, marking the end of the 14-19 diploma.</p> <p>Second, it was announced that a new £180 million a year programme would be established to support the most disadvantaged 16 – 18 year-olds in further education. This new programme replaced the Educational Maintenance Allowance (EMA).</p>
Technical and Further Education Act (2017)	<p>The Technical and Further Education Act (2017) scribes into law a number of provisions affecting the further education sector. They are summarised as:</p> <ul style="list-style-type: none"> • the introduction of an insolvency regime for further education and sixth form colleges; • the renaming of the Institute of Apprenticeships as the Institute for Apprenticeships and Technical Education; and, • it being compulsory for schools and other providers to give their students access to colleges and training providers as part of a comprehensive and impartial careers service and that such careers guidance to be reported on as part of a school’s inspection.

Tables 1.1 The genealogy of further education in England

It should be noted, however, that whilst the above is a comprehensive list of legislation and policy development related to the further education sector since incorporation, it does not include all consultation documents, white and green papers, Ofsted handbooks for the inspection of further education and skills, independent reviews and position statements. Where appropriate, they have been referred to elsewhere in this study.

The further education sector is one that can best be described as poorly defined, often being referred to as the 'Cinderella Sector' in that it is stuck in a 'calibrated hierarchy of worthwhile achievement' that emphasises academic success through higher education and is therefore overlooked when compared to its school and university siblings (Kennedy 1997, p1; Panachamia 2012, p1 & Gravatt, 2004). Despite catering for 2.9 million people annually, including 773,000 students aged 16-18, and receiving an annual allocation from the government of £7.5 billion (Association of Colleges, 2015) the further education sector whilst important is not well understood (Parry, 2009; Coffield et al, 2008; Hodgson & Spours, 2008).

Before the raising of the participation age (DfE, 2014), further education could best be described as 'post-compulsory education and training beyond the age of 16, but excluding higher education' (Hodgson & Spours 2011, p1). However, in reality the system is more complex – a system that is a heterogeneous amalgam of institutions, provision and funding bodies and until relatively recently, regulatory regimes.

Within that context, it is worth noting that the delivery of further education takes place in a range of venues and sectors. These include:

- Colleges of further education;
- Sixth form colleges;
- Specialist colleges such as land based, art and design, dance and drama and for students with learning difficulties/disabilities;
- Adult and community education colleges;
- Independent training providers; and,
- Others, including prisons, further education in higher education and the military.

This complexity is rooted in the fact that the further education system in relative terms is a new system within the educational landscape in England. The first university in England (University of Oxford) was founded in 1167 and the first school (Lincoln School) was founded in 1090. It would not be until the Education Act (1944) that the further education system was formally established, although not in a state that could be recognised today, The further education sector as it is known and understood today came about as a result of the implementation of the Further and Higher Education Act (1992). This Act would free colleges

and other providers from local authority control, and lead them towards becoming independent, self-determining organisations.

Prior to the Education Act (1944) which mandated Local Education Authorities by law to provide 'adequate' provision of both vocational and non-vocational types for students aged 16-18 and those aged 19+ studying on both full-time and part-time modes, education and training for young people and adults outside of the school system were piecemeal. They were inconsistent and generally poor in offer and quality (Hodgson et al, 2015). This reputation and approach to the establishment and development of such a new system continue to evolve against what is experienced as a cyclical background of poor outcomes, qualifications that are not fit for purpose. Moreover, this takes place within a political landscape that is confused as to what it wants from the system (City & Guilds, 2015; Wolf, 2011; Wilshaw, 2016; Foster, 2005).

Hodgson & Spours (2011) make clear that in recent years the age-related ministerial bifurcation has divided the policy narrative to a point that there is no overarching national conversation for further education or of lifelong learning. Indeed, it could be argued that until recently there has never been a national plan for further education or even any requirement for anyone to attend it (Unwin & Bailey, 2014). This is despite the number of reviews (since colleges become independent of local authority control), governmental white papers and holistic changes, the most recent of which has been the introduction of the Study Programme and the mandate to study English and mathematics until the achievement of a GCSE graded A*-C (DFE, 2014). This is in addition to the re-organisation of the further education and sixth form college sectors through area reviews.

1.3 The further education college sector and its provision within a system

As outlined above, the further education system is one that is all encompassing. Within the system and specifically with reference to this thesis, there are a range of general further education colleges. These colleges vary in size and location from major towns and cities to rural communities. However, what they deliver, to a large extent, is quite similar; only the size of the provision (in student numbers and associated income) will act as the greatest variable. The provision can be described using the following four categories:

- Vocational, occupational and technical education and training
- Apprenticeships
- General education i.e. functional skills and GCSEs
- A-Levels

The further education college sector in its entirety and as individual institutions, offers a curriculum and provision similar to that delivered by the further education system, i.e. to all ages, modes of delivery, mixes of provision (including levels) on-site and off-site. It is this complexity, with the addition of a set of success rate measures, different from those used in schools but with the same regulatory inspection judgements, which Unwin & Bailey (2014) state contributes to the 'lack of status and visibility enjoyed' by the school and university sectors in public life and indeed within the research community.

Colleges of further education, of which there are 216 in England (AoC, 2015) have a low political profile, in many ways due to the fact that they deliver, or are known for delivering, vocational education and training (Hodgson et al, 2015). People will know about apprenticeships, even if their thinking is out of date, based on a very traditional system of on the job, technical training. And, of course, everyone will know about GCSEs and A-Levels. However, vocational education and training, in the main are known only to those who have directly experienced it. The place of vocational education and training within the British system has undergone several reinventions and numerous policy shifts since the 1944 Education Act. As a result, confidence in such an ever-changing, wide-ranging and, for some time, overlooked sector has been limited (Foster, 2005; Coffield et al, 2008; Wolf, 2009).

Often seen as the pathway associated with those who are not able to progress along 'traditional' or 'academic' routes, participation on full-time programmes in further education

colleges, post-16, up until the rise in the participation age, was on the decline. The numbers progressing to university or to higher level education had stalled, which is contrasted with the increase in both participation and progression from GCSEs and A-levels (QCA, 2008).

It is commonly agreed that learning and educational progress benefits both individuals and society with regard to the economy and social inclusion (Kennedy, 1997; DfES, 2005; Coffield et al, 2008; DfES, 2004). It is important, therefore, for students aged 16 and adults who make educational decisions for future employment or study based on aspirations, to have a broad and appropriate curriculum from which to choose. And, as students are individuals, provision should be tailored to the specific aims and goals of each student.

However, for some time, educational opportunity post-14, and indeed throughout both the primary and secondary stages, has been dominated by an outcomes or product based system which seeks to judge a student's 'academic' ability (Lumby & Foskett, 2005). General Certificates of Secondary Education (GCSEs) were introduced as part of the 1988 Education Reform Act (ERA) and replaced the long-standing 'O' Level qualifications. Originally simulating post-16 learning, GCSEs, and the associated grading system, became a selection tool and finally a barrier to progression (Raffe & Spours, 2007; Young & Leney, 1997).

In addition, due to the centrality of GCSEs and the long standing 'A' Levels within the school curriculum, these 'benchmark' school leavers' qualifications (which are not only used for measuring student success, but also the school's success) have marginalised alternative or vocational provision (Hodgson & Spours, 2008; Hodgson & Spours, 2008a; Lumby & Foskett, 2005). As a direct result, many disengage. Over 40% of students in state funded secondary schools, leave school without the baseline target of 5 GCSE grades A*-C including English and mathematics (including equivalents) in 2015.

It is widely accepted that the perception of vocational education and training when compared to academic education is of a lower order. Whilst there are a number of arguments that can be attributed to such a belief based on history, culture and society, the fact is, currently, a student's first encounter with such assertions is towards the end of compulsory schooling and during the selection process for further education. In brief, students with high GCSE grades will more than likely progress through the academic route ('A' Levels) whilst students with more precarious grades will progress onto a vocational

route, a deficit choice, if indeed they stay on in education (Lumby & Foskett, 2005; Hodgson & Spours, 2003; Hodgson & Spours, 2008a).

Furthermore, to support such educational opportunities for a certain cohort of young people, vocational options for students aged 14-16 are limited and are an alternative for those deemed to be poor at academia (Lumby & Foskett, 2007; Wolf, 2011). Hence, the status of vocational education is low and is often seen by stakeholders, such as employers and admission tutors in higher education, as inferior to their 'academic' cousins (Young, 1996, cited in Lumby & Foskett, 2005). This can be attributed, in part, to the comparatively short term and piecemeal approach to constructing a cohesive and rational strategy for vocational education and training (DfES, 2004). This is in complete contrast to A Levels which have been a staple part of the post-16 offer for over 65 years.

Lumby & Foskett (2005) attribute the impression of vocational education and training as being lower in value and status to a range of historical processes and systems which have 'prioritised one form of learning over another, which has been used to establish or confirm social structures'.

1.4 Further education policy: learning, confusion or amnesia?

The policy arena that the further education system and subsequently the further education college sector find themselves in is confused, quick to change and lacking in sustainability (Higham & Yeomans, 2007; Lumby & Foskett, 2005; Raffe & Spours, 2007a). Leading up to and between incorporation (when colleges opted out of local authority control) and 2016, there have been eighteen Secretaries of State and the same number of Ministers (of State or Under Secretaries) who have had responsibility for the further education system – its direction, funding and regulation. In addition, this governmental leadership of further education in England has been directed through five departments and five funding bodies. Furthermore, successive governments have commissioned numerous reviews and produced a series of white papers, all of which have sought to determine a radical overhaul and long term strategy for the sector (DCSF, 2008; DCSF & DIUS, 2008; DfE, 2016; DfEE, 1999; DfES, 2004; DfES, 2005; DfES/LSC, 2007; Foster, 2005; Kennedy, 1997; Leitch, 2005 & Wolf, 2011). However, in reality, this has led to what Lumby and Foskett (2005) refer to as policy 'busyness' and 'continuous innovations' and what Avis (2009) calls 'policy hysteria', the reflection of which is continual shift in governmental priorities; even when such innovations

are implemented they are often short lived (Foster, 2005; Keep, 2011 & Nash & Jones, 2015).

Despite the white noise of policy change, it could be argued that the closest the further education sector has got to an overarching policy narrative emerged from two seminal papers commissioned by the then Labour government and which still permeate the thinking of the college sector today. These commissioned reports were 'Learning works: widening participation in further education' (Kennedy, 1997) and 'Skill in the UK: the long term challenge' (Leitch, 2005). As their titles suggest, the further education college sector is often seen and celebrated for delivering education and training, often for those who have not succeeded within the school sector or for adults who are returning to education in order to change their lives and they imply that the provision within the system is designed to respond directly to the needs of employers and ultimately the economy. Indeed, Orr (2009, p481) asserts this bifurcation, identified by the Labour government as delivering 'social justice through widening participation in education' and 'enhancing national economic competitiveness through workforce skills'.

However, even these two pillars of the further education system and college sector more specifically have come under some tension in recent years. For young people, the mixed economy of providers from whom they can choose to continue their educational career post-16 is now within a landscape of academies, free schools and private training providers, with schools, often with small sixth forms offering some vocational provision. Moreover, it is essential that, as the options for study post-16 are embedded within a complex system of education and training providers, they are underpinned with appropriate and well informed impartial advice and guidance; whereas Freshminds (2014) report that they and a number of national bodies, including the Education Select Committee, have concerns about the quality of school-delivered advice and guidance.

At the beginning of the coalition government's term of office, it announced that the financial scheme supporting students to stay on, in what was then post-compulsory education, would change. The Education Maintenance Allowance would be withdrawn and replaced with a £180m bursary scheme. The Association of Colleges (AoC), as part of their 16-18 recruitment survey of all colleges, identified that for the first time in 20 years the number of 16-18 year olds starting at a further education college had decreased overall by 0.01% and that 49% of colleges had seen a decrease, in some cases up to 15% of their intake. The two

main reasons cited were the withdrawal of the EMA and rising transport costs (one of the uses of the EMA) (Mamon, 2011 & BBC, 2011).

For adult students (those aged 19 and above), the Conservative and Liberal Democratic coalition government from 2010 and subsequently, from 2015 the Conservative government, have overseen a real terms decrease in the adult education budget (between 2009 and 2014) of some 35% with an additional 24% reduction for the 15/16 academic year. As a result, the number of adult learners returning to colleges has reduced considerably and the number of courses offered to this age group has declined accordingly, despite the introduction of a loan facility, currently only available for students studying at level 3.

This reduction of funding per 16-18 year old and across the adult education budget (and in some cases individual colleges' ability to manage their budgets) has led to the latest government policy of area reviews (which are designed to identify different models of collaboration between colleges including merger). This is likely to result in the biggest structural change to affect the further education college sector since incorporation, as one of the expected outcomes is to have 'fewer, larger and more financially resilient organisations' (BIS & DFE, 2016, p3).

1.5 Ofsted: the regulator and inspector

Like the perceived indiscriminate development of the further education system and the college sector cited above, there have been a number of bodies and agencies since incorporation that have underpinned its regulation, funding and development. Whilst the remaining commentary in this section will focus on the regulator and inspection organisations, it is important to note that since the Further and Higher Education Act 1992 gained royal assent, there have been four funding agencies, ten development/improvement agencies and seven organisations that have identified basic and/or professional standards for staff working within the further education system (Fletcher, 2015). Many of the functions housed within the various organisations are not new. Instead, their make-up, regardless of purpose, has been as the result of re-structure, marrying and divorcing activities, amalgamating work streams and staff to best deliver policy initiatives, legislation or indeed efficiency.

The Office for Standards in Education (Ofsted), has been responsible for the regulation and inspection of further education and skills, for young people aged 14-19 since 2001 and for all of its provision (with the exception of higher education) since 2007. Prior to 2007 and following incorporation, regulation and inspection were fractured along the lines of age and provision (Fletcher, 2015; Fletcher, Gravatt & Sherlock, 2015).

Although inspection had played a major role in the landscape across the further education system, where it had (at the hands of Her Majesty's Inspectorate) been subjected to irregular, rhetorically developmental inspections with few consequential inspections, its culture and its structure of high stakes accountability experienced today, came into play as a result of a one-stop-shop approach to inspection arising from the Education and Inspection Act 2006 (Baxter & Clarke, 2013; Courtney, 2012; Fletcher, Gravatt & Sherlock, 2015). The implementation of the Act would see the merging of the Adult Learning Inspectorate into Ofsted. Although the two agencies had worked together under a common inspection framework from 2001, they had approached inspection with different emphasise – one focusing on judgement, the other on development (Fletcher, 2015; Baxter, 2013)

The history of regulatory and inspection bodies as they affected the sector is set out below:

- Joint inspections by the Further Education Funding Council and the Adult Learning Inspectorate from 1993 until 2000;
- Work-based Learning inspections by the Training Standards Council from 1997 to 2001 (before being subsumed in to the Adult Learning Inspectorate);
- Joint inspections between 2000 and 2007 by Ofsted and the Adult Learning Inspectorate; and,
- Ofsted inspections from 2007 onwards

Since 2007, under the common inspection framework, there have been four handbooks which outline the framework by which colleges are inspected. Published and implemented in 2007, 2009, 2012 and 2015, each document is expected to make clear to all interested parties, including the general public, the remit by which Ofsted inspects the various activities within the system and how it or more specifically its inspectors make judgements (Ofsted, 2007; Ofsted, 2009; Ofsted, 2012, Ofsted, 2015). Moreover, true to the foundation principle of the new inspection regime, coming out of John Major's Citizens' Charter, all handbooks and other Ofsted documentation are explicit on how inspection reports are to be written so

that they can be used by parents and other consumers of the further education system (Major, 1991; Baxter, 2013).

However, whilst the underlying principle remains, or at least remains the aim of the regulator, it is assumed that the comparability between one Ofsted inspection and the following is the same – that is, that the measurement of progress (or decline) is based on the same set of parameters or indicators. Otherwise, it would have to be mean that all people reading and using these reports to make decisions, be they college or school leaders, other inspectors or, arguably most importantly, parents and consumers, are fully informed of the changes to scope or emphasis, policy levers and/or overriding priorities (Park, 2013; Coffield, 2012).

1.6 Outcomes as a default for the standard of teaching and learning?

The current government's White Paper Education Excellence Everywhere, published in March 2016, clearly states that

'Ofsted will consult on removing the separate grade judgements on the quality of teaching, learning and assessment to help clarify that the focus of inspection is on outcomes...' (DfE, 2016 p22)

For further education colleges, this would be seen as a major policy shift to the inspection framework that is currently operating, the original version of which was introduced to the sector in September 2012, or, more specifically, this would be a policy reversal to the pre-2012 guidance for inspections as identified in the handbook for inspecting colleges (Ofsted, 2012).

Outcomes for learners have been central throughout the evolution of the performance management regime and therefore of Ofsted inspections since its inception in the early 90s and subsequently as it applied to the further education sector (Fletcher, Gravatt & Sherlock, 2015). In its broadest forms, outcomes can be classified as examination results and therefore, the judgement is about the number of students within an organisation, studying for a particular qualification (or set of them i.e. GCSE) or at an appropriate level who achieve their qualification (Park, 2013). In practice, and as identified by Ofsted, outcomes for learners involve a wider set of criteria which inspectors are asked to consider to inform their

judgement for this category. Within the 2012 handbook, the specific criteria which inspectors 'must' evaluate [are] the extent to which:

- All learners achieve and make progress relative to their starting points;
- Achievement gaps are narrowing between different groups;
- Learners develop personal, social and employability skills; and,
- Learners progress to courses leading to higher-level qualifications and into jobs that meet local and national needs (Ofsted, 2012 p43).

Outcomes as above, of all the other key headline judgements, are based on a considerable amount of objective information, and therefore, it is argued, for the area of examination results, Ofsted's role is one of validation. As a result, the rest of the inspection is used to identify what the experiences of students are as they pertain to teaching, learning and assessment and how both together are being planned and improved as a result of leadership and management (Waldgrave, 2014; Boocock, 2014). It should be noted at this stage that due to the overriding priority of outcomes, and specifically examination results, Waldgrave (2014) asserts that, whatever the outcomes, they have the ability to underpin a pre-determined judgment, (explicitly or implicitly) and therefore the inspection as a whole becomes an evidence finding, confirmatory process.

It is interesting to note then, that inspections of further education colleges and other providers in the further education system, from September 2012, sought to place teaching, learning and assessment at the heart of the inspection processes, relegating outcomes to a secondary position. This can easily be seen from the inspection reports themselves and the relationship between the grading between:

- Outcomes, Leadership and Management and Overall Effectiveness (pre-2012), and,
- Teaching & Learning, Leadership and Management and Overall Effectiveness (post-2012)

For inspections pre-September 2012, in the majority of college inspections, the grade awarded for outcomes was also the grade awarded for leadership and management and, by default, the grade awarded for overall effectiveness. The grade for teaching and learning could be higher or lower than the other key judgements without compromising the top grade. For inspections post September 2012, the grade awarded for teaching and learning in

the overwhelming majority of inspections up until the introduction of the current guidance for inspecting colleges (Ofsted, 2015) would match the grade awarded for leadership and management and again, these combined grades act as a default for the overall effectiveness judgement. The grade awarded to outcomes could be and sometimes is lower than grades awarded for the other key aspects.

Chapter 2: Literature Review

2.0 Introduction

Over the 2015/16 academic year, just under 25,000 inspections were carried out by Ofsted, of which 85 took place within general further education colleges (GFEs), representing a sample of 38.1% of this sector (Ofsted 2016, p6-7). By contrast, whilst the most number of inspections (in terms of physical inspections) were in maintained secondary schools (666), maintained primary schools (2,468), maintained schools with early years (2,423) and early years registered providers (20,761), their representative samples were much smaller than that of the GFE sector at 19.7%, 14.7%, 14.5% and 30.8% respectively (ibid.). It is therefore obvious from these numbers, that inspection has the ability to reach, affect and impact on the lives of very many teachers, academic leaders and indeed students, especially when these numbers are multiplied across the life-cycle of the common inspection process.

Whilst this thesis overall and the analysis of outcomes for learners are singularly focused on general further education colleges, there is very little research and consequently literature covering this sector. As a result, the majority of the literature reviewed, with a few rare exceptions, uses studies and research from the primary and secondary phases of education. However, as Ofsted is responsible, and has been since before the timeline covered in this thesis for the inspection for both schools and colleges, many of the previous studies affecting school inspections are relevant to those carried out in GFE colleges and indeed other educational settings within their remit.

Therefore, in order to place the literature reviewed into context, as it applies to secondary and primary schools, it is important to map out the history of inspection from its inception to the current day. Baxter (2013), identifies that inspection has been part of the schooling system since 1838 and up until 1992 with the creation of the Office for Standards in Education, when inspection was undertaken by Her Majesty's Inspectors of Schools under the direction of the Chief Inspector. Ofsted, at its inception, became the regulator of education with the overarching remit to control inspection and analyse inspection data, both to inform ministers as to the health of the system and to publish information on performance in individual schools (Lee & Fitz, 1997). Unlike the HMI regime, all schools in England would be subject to inspection every four years and the emphasis would transfer from that of development to quality assurance.

Since its inception, there have been many publications and guidance documents from all inspectorates, Ofsted, ALI and the Training Standards Council. However, Coffield (2017) shows that regardless of textual changes and changes of emphasis or priority areas, many of the underlying principles of inspection remain the same. Cited within both acts of parliament affecting school inspections and the accompanying policy guidance such as 'The Common Inspection Framework' and the numerous handbooks for the inspection of schools and other education providers, some of the key changes have been:

- The number of judgements on a school has reduced considerably from twenty seven to four, plus one overall grade for effectiveness. The current four are: leadership and management; teaching, learning and assessment; personal behaviour development and welfare, and outcomes for students (Ofsted, 2016a);
- The reduction in the number of grades awarded from seven (ranging from excellent to very poor) to the current four: outstanding (grade 1), good (grade 2), requires improvement (grade 3) and inadequate (grade 4) – the last two replacing the earlier judgements of satisfactory and unsatisfactory;
- The cycle of inspections, which initially stipulated that each school was to be inspected every four years. This changed to every six years as part of the School Inspection Act 1996 (amended in 2000) with the introduction of both full and short inspections with less effective schools being inspected more regularly. The Education Act 2005 and Education Regulations of the same year reduced the inspection cycle to three years (Ofsted, 2004), only for it to be increased to five under the Education Regulations 2009. Five year cycles remain for schools judged to be good at their last inspection, whereas schools judged to be outstanding can go for longer periods (following an interim assessment) and those requiring improvement or who are inadequate will be subject to an inspection within three years or very regular monitoring visits respectively; and,
- The notice period given to schools has decreased by some margin. At the outset, a school would have been given a number of months' notice of an impending inspection; this has been reduced to between six and ten weeks under the New Framework (Ofsted, 2000). However, as part of a considerable shake-up of inspections in 2005, the notice period to schools became 48 hours and the duration of inspection reduced from the full school week to between two and three days (Elliot, 2012 & Ofsted, 2004). Currently, a school will be given approximately one day's notice and in some cases, no notice (Roberts & Abreu, 2016).

This chapter continues with a critical review of literature, most of which concerns itself with aspects of inspection such as its influence on individual and organisational behaviours, and its relationship to both accountability and quality improvement

2.1 Inspection and emotion

As a non-ministerial department, Ofsted's primary purpose is to implement, through inspection and regulation, government legislation. Although working through the Department for Education, Ofsted is accountable for its work directly to parliament. However, Ofsted is as much an influencer of government policy as it is a regulator of it (Baxter, 2013; Raffe, 2008; Fletcher, 2015). It has the power to 'form and inform' policy that influences the way in which standards in education [in England] are understood, even when the parameters change from one inspection to the next (Baxter, 2013; Raffe, 2008; Ofsted, 2014c; Jones & Tymms, 2014).

The paradox of Ofsted as an inspector and a policy influencer plays itself out to such an extent that regardless of the Chief Inspector's encouragement of head teachers to be 'mavericks' or to 'take their own path', most heads, according to McInerney (2016), diligently, even slavishly, follow Ofsted pronouncements, where teachers see inspection as a hoop to jump through or where they engage at an arbitrary level, often suffering from 'inspection fatigue' (Chapman, 2002 p268). Inspection, or at least the structure of school and college inspection that is in place today, is, as Fletcher (2015) puts it, a 'high stakes' model, making judgements from within a small time window, so creativity and entrepreneurship in school and college leadership are seen as too risky, further demonstrating the central reliance on student outcomes and not necessarily their experience of how they achieve them (Forrest, 2015; Courtney, 2012).

As a result, to many schools and colleges inspection becomes a game that needs to be won as the ramifications for not playing well or losing are considerable, not only on the life of the school but for individuals, specifically head teachers. So understanding the game that you are playing and what the goal looks like or even where it is, is a must for successes. Muijs & Chapman (2009, cited in Courtney, 2012 p1) state that, within the context of inspection it is a 'well-known phenomenon that organisations will concentrate their efforts on those things that they are judged on'. So if this is the case, what is the game and how is it played?

Ofsted realised some time ago that schools and colleges prepared for inspection. For further education colleges, under the first framework, which saw Ofsted responsible for the inspection of all of its provision (excluding higher education), colleges were given three weeks' notice of their intended inspection (Ofsted, 2007). This timeframe would allow colleges, as best they could, to prepare students, staff, documents and facilities to present the best picture possible. Ofsted, following a period of consultation, reduced this notice period to two working days (Ofsted, 2009 p9) with the specific aim of getting as close as possible to the everyday experience of students and by doing so, reducing the provider's ability to present a false case and, arguably, removing the 'tick box' culture of inspection (Fletcher, 2015; Baxter & Clarke, 2013).

However, whilst the reduction in notice period from three weeks to two working day (in reality, four days due to notice being given on most occasions on a Thursday), this has not led to a complete reduction in what Fletcher calls (2015 p18) 'gaming behaviour' where due to an impending inspection, whether it has been called or whether it is anticipated, the normal operation of a college is 'distorted' in the pursuit of giving off a good impression. In part, it is argued that this type of behaviour is understandable given the possible consequences of being judged to 'require improvement' or worse, being judged to be 'inadequate' - under such a target driven accountability structure and where one voice, in the case of Ofsted, has the power to dominate or even terminate discussion (Park, 2013). Therefore, in light of such consequences, the behaviour and priorities of leaders and teachers in both the preparation for and during inspection can be influenced by what they think Ofsted inspectors are looking for (Courtney, 2012). Ehern (2014), specifically notes that teachers during inspection will move away from their normal practice and will 'teach to the test' or to inspection criteria i.e. what they believe Ofsted want to see. Moreover, Ehern (ibid.) and Fletcher (2015) note that the curriculum and other activities in schools and colleges are amended, often narrowed, in the anticipation of an inspection, again to what head teachers and college leaders believe inspectors will judge to be at least good.

But if inspection is there to identify what the standard of education is and that standard is being identified consistently, it is worth exploring why gaming behaviour is necessary. Moreover, if the judgements are predicated on a 'false' presentation rather than the actual standard, consideration should be given to the exacting relationship between inspection and improvement.

Inspection is now a considerable and consistent part of the accountability and performance culture of the education system in England, the effects and side effects of which can come at a cost (de Wolf & Janssens, 2007; Ehern & Swanborn, 2012). One such side effect, which dominates the literature analysing the emotional impact on teachers and academic leaders, and which is recognised to be considerable, is stress (Perryman, 2005; 2006; 2007 & 2009; Ouston, Fidler & Early, 1997; Clapham, 2015). This is especially the case when the outcomes of an inspection can have serious ramifications for the school, individuals and the wider community. During an inspection, a school or college can become an 'organisation for the gaze' (Ball 1997, p332) with the potential for additional pressure and anxiety to be placed on school and college leaders, permeating the entire organisation and specifically those who will have a direct contribution to the overall judgement, at the forefront of whom are teachers.

One early study following the creation of the Office for Standards in Education (Ofsted) in 1992, and its subsequent responsibility for inspection policy and implementation, identified that stress, above and beyond normal stress factors, was heightened as a result of inspection (Brimblecombe, Shaw & Ormston, 1995). Specifically, they state, as a result of evidence from a national survey of secondary schools and a small number of interviews, that the two compounding factors which are perceived to increase stress in teachers as it relates to inspection are:

- 1) the fear of the unknown, that is to say the lack information as to what exactly inspectors are seeking to judge beyond the text cited in the common inspection framework; and,
- 2) the tensions teachers felt as a result of being pulled in many directions as to what was important.

The increase in stress, they suggest, has the ability not only to change individual and organisational behaviour, but it can place individuals in a state of fear. This fear, they assert, can lead to an individual questioning their own professionalism leading to a decrease in professional self-esteem, which can play out during inspection. This has the potential to negatively impact on inspection evidence and therefore the validity of any judgements made (ibid.; de Wolf & Janssens, 2007; Gray & Gardner, 1999). Sandbrook (1996) also concludes from research undertaken in primary schools that teachers were most anxious about losing control in the classroom during an observation by an inspector. This was because the

teachers were either unclear about what exactly was being observed or they changed their individual behaviour to a manner that they thought was expected of them.

Loss of professional identity, loss of self and loss of pedagogic values are key themes evidenced in a case study by Jeffrey & Woods (2005). Supporting the assertions made by both Brimblecombe, Shaw & Ormston (1995) and Sandbrook (1996), their findings from a qualitative case study from a small primary school, within an area of high social and economic deprivation, signify that the impact of inspection often leads to negative emotional reactions. Such reactions put forward by the teachers in this school (and in another four schools associated with the case study) include professional uncertainty, inadequacy and what Jeffrey & Woods (2015 p10) describe as 'grief and bereavement for professionalism, autonomy and identity'.

Moreover, such loss of professional identity is put forward by Pennickx & Vanhoof (2015) as a point in time, when a teacher, due to their experience of inspection, loses their self-confidence and 'starts to ignore their own professional ideals' (ibid., p489). Their study and subsequent conclusions are based upon a review of published literature from 1995 to 2012 concerning itself with the emotional consequences of inspection. It should be noted that Pennickx & Vanhoof (2015, p495) state as part of their conclusion that research in the area of inspection and emotional consequences needs further insight. However, they do state that of the twenty eight papers reviewed in their study, all provided evidence of 'negative emotions' from school staff prior to, during and/or following an inspection, although to varying degrees.

The reported effects and consequences of inspection and emotion so far have been taken from literature that, in the main, comes from research undertaken within the English education system. However, the following three studies give a comprehensive insight to the relationship between inspection and emotion as experienced in other European countries. Two papers by Ehern, Gustafsson, Altrichter et al (2015) and de Wolf & Janssens (2007) present theoretical studies of control mechanisms in education and public performance indicators, namely inspections, and seek to catalogue the key characteristics and to profile the side effects of each system. In addition to their reviewed literature, Ehern, Gustafsson, Altrichter et al (ibid.) sought primary evidence via an on-line survey from primary and secondary school principals (73 respondents). It should be noted that whilst the literature reviews in these two studies are purported to be pan-European, the majority were written on

the English inspection system. However, whether their evidence is considered as a whole or in part (just mainland European countries), stress continues to be identified as a considerably emotional experience in teachers and academic managers throughout the inspection process. Moreover, de Wolf & Janssens (2007, p383) go so far as to suggest that stress can go simply beyond an undesirable state to a situation that 'can [also] hinder the assessment and jeopardise the validity and reliability of the assessment'. This also has the ability to be compounded as proposed by Jeffrey & Woods (2005, p2); inspectors make no allowances for emotional responses, either to the inspection process itself or to teaching and learning, leading to an assumption that teaching quality should be defined in terms of 'technical competencies' (Ofsted, 1993; Troman, 1996).

The third study, using primary qualitative data from Belgium, is that by Penninckx, Vanhoof, De Maeyer & Van Petegem (2015). They seek to interpret what perceived effects and side-effects there are following an inspection and how these and the perceived inspection quality affect a school's policy making capacities. Unlike the majority of studies used in this literature review, the outcome of this study of 2,202 teachers puts forward the notion that factors leading to greater stress levels include the context within the school, rather than it being solely related to the outsider coming in. It reports that teachers within schools that have a higher level of autonomy or 'stronger policy-making capacities' (ibid., p343) will experience a greater degree of stress or anxiety throughout the inspection process – a relationship which has previously been explored by Scanlon (1999) and Sandbrook (1996) and in a pan-European study by Hall, (2017).

The literature reviewed, as it applies to the topic of emotion and inspection, whether it be via a theoretical analysis of evidence or through primary qualitative research, is in the main based on a case study or as the result of a one off survey. Two further studies, by Ousten, Fidler & Early (1997) and Perryman (2002; 2006; 2007 & 2009), give a longer term perspective to the implications of emotion and inspection.

Ousten, Fidler & Earley (1997) was an early study looking into the after effects of an Ofsted inspection. The evidence for this study came from two cohorts of secondary schools. The first cohort were inspected in the autumn term of 1993, the second from the autumn term of 1994. Each survey was targeted at all inspected schools within the timeframes outlined above, of which 284 responded from the first cohort and 399 from the second. A follow-up survey was sent to the original respondents two years later i.e. 1995 and 1996 respectively.

Most responses were completed by the head teacher and have therefore been categorised as the senior manager's view of the inspection process and subsequent effects. As part of the survey, schools were asked to identify long-term positive or negative outcomes of inspection. The positive outcome identified was the 'confirmation that it was a good school' (ibid., p100). However, two negative outcomes were recorded, one of which was 'the stress and demoralisation it may create for staff at all levels' and 'the negative impact that it may have on the community' with the latter point featuring as a consequence of this new type of inspection – external accountability (Perryman, 2007).

The second long-term study of inspection and its impact on teachers and managers, is in fact a set of studies spanning a period of seven years from one school. Throughout the life of this case study research project, the school has been on a journey, initially being classified as being in special measures. Each individual study marks a specific point in the inspection cycle (Perryman, 2002; 2005; 2006 & 2007), culminating with two studies on the performance of staff for inspection and the continuation of improvement post-Ofsted (Perryman, 2009 and 2010). The early context of the secondary school within this set of studies, is that it was a Fresh Start school (Perryman, 2002 p46), which, according to the DFEE (1998, p14), is a school that has been judged by Ofsted to be failing. The school, following an inspection a year after re-opening after under the scheme, would be inspected and placed in special measures.

Two studies of particular note are Perryman (2002 & 2007) both of which are concerned with the experience of staff working under the intense regime of special measures. In these case studies, the impact of being both a Fresh Start school and subsequently being placed in special measures would result in very quick changes in senior management (on more than one occasion), constant monitoring and inspection by Ofsted and, as Perryman (2009, p628) concluded, a staff team perfecting their ability to play the inspection game. The impact of such scrutiny is considered during the two years in between inspections when Perryman (2007) interviewed, over the period of two years, up to and including the school's latest inspection, a group of long-term teachers and a small group of middle and senior managers.

The study takes places over two years and in three time frames:

- Year A: The Autumn term, approximately one year after coming out of special measures;
- Year Bi: The following spring, the six weeks leading up to and including inspection; and,
- Year Bii: The summer term, where people involved earlier in the study were re-interviewed about how they felt about the inspection and how it had affected them.

The emotional impact in this situation was considerable, due to the specific context of the school, a context in which Ofsted would inspect the school with regularity and at one stage, six times in one year. It is noted in this study that due to the perceived culture of constant surveillance, judgement and public accountability, teachers and managers felt that they worked under a disciplinary system embedded with fear and in an organisation which existed purely for the purposes of 'passing an inspection' (Perryman, 2007 p189). For this to succeed in the way that it did, it is acknowledged, was the result of emotional dissonance - whereby individuals are forced to play a game or to project differing emotions and actions to those that they actually feel and believe, the result of which manifests itself as a 'technical workforce to be managed and controlled rather than a profession to be respected' (Tomlinson, 2001 p36).

2.2 Performing for Inspection

Such technical approaches to teaching, for the purpose of being successful at inspection, are what Perryman (2006, p150) refers to as 'performativity...about performing the normal within a particular discourse'. In this study, the discourse identified is not only of the inspection as a single event or series of events, but also of the role of Ofsted as the 'omnipotent presence' (Clapham, 2015 p622) both of which can fundamentally affect teaching practice.

The case study school used by Perryman (2006) is the same used in Perryman (2002 & 2005). The timing of this research bridges the gap between the last year that the school was in special measures (Year A) and its first year being released from them (Year B). The research method used to gain evidence was semi-structured interviews of those teachers who had been at the school at the point when it went into special measures. Interviewees in

Year A consisted of 13 individuals most of whom had managerial responsibilities and only three of whom were main-scale teachers. In Year B, this cohort had reduced to 10 as three had left to take roles elsewhere and the three main-scale teachers had been promoted to middle-management positions. Furthering the analysis in earlier studies on inspection and emotion, specifically anxiety and stress, Perryman (2006) puts forward the notion that in order to survive, teachers and managers need to comply within a framework of prescriptive norms or put more specifically 'perform the game' (ibid., p159). The game, it is asserted in this setting and in the wider context of inspection, concerns itself with the loss of professional autonomy and the assertion of teacher as professional, being replaced by a system of conformity, continuous observation and fabrication (Perryman, 2009; de Wolf & Janssens, 2007; Pennickx & Vanhoof, 2015). It is further asserted by Perryman (2006, p156) that 'under the Ofsted regime...there is no room for freedom of expression – unless a school 'behaves' and 'jumps through hoops', it will not be released from the regime'.

Similarly to stress and anxiety, both gaming and fabrication can also be seen as a side effect or 'unintended strategic behaviour' to the process of inspection (de Wolf & Janssens, 2007). In Clapham's study (2015), the relationship between inspection, performativity (the scientific descriptions of people and cultures with their customs, habits and mutual differences) and fabrication is explored. This study has been created against the backdrop of comments made by the then Her Majesty's Chief Inspector of Education, Children's Services and Skills, who commented publicly that 'putting on a show was deeply irritating...and that there was no Ofsted preferred teaching style' (Wilshaw, 2013). The focus of this study is the perspectives and experiences of two teachers: one at the beginning of their career and the other an experienced or, as Clapham (2015 p619) puts it, a veteran teacher. Whilst the researcher recognises that the sample used in this study is small and that there is acknowledgement of the implications of their own subjectivity and reflexivity (due to their past experience and prior knowledge), the validated evidence from both participants and eleven other contributors can be assumed to be the experience of other teachers, within this school and others. Positively, both participants in this study wanted to be inspected. They wanted to be inspected as they believed that this external assessment validated their self-assessment; gave them the opportunity to 'show off' what they were already doing and to heighten motivation, at least for the new teacher whose aim was to be outstanding (Clapham, 2015 p618). Neither wanted inspection to be a game, passed or achieved, based on fabricated evidence and thereby presenting to the inspection team a false picture of their own practice and the school.

However, as Ball (1998, p198) concluded, 'to be successful in inspection, teachers are inwardly focused upon the survival of their institution within the education market' a market that is based on metrics such as the achievement of A*-C at GCSE and a successful inspection outcome. As a result, whilst both participants in this study believed that they had professional autonomy, they were indeed not working in a fabricated system, but as Clapham (2015) describes, a post-fabricated regime - post-fabrication that takes the form of school's internal quality assurance processes, including Mock Ofsted inspection (Mock-steads); the internal pressure for teachers to become Ofsted successful practitioners; the constant inspection readiness culture permeating the entire organisation and the role of the 'virtual inspector' influencing key strategic decisions (Clapham, 2015 p621). It is these every day influencers, driven by a high stakes inspection regime, which have the ability to, and in this case are proven to, affect individuals' approach to teaching and learning.

Perryman (2009) refers to the aspects that Clapham (2015) calls post-fabrication as normalisation. Through the experiences of teachers and managers in the case study school cited elsewhere in this chapter, it is suggested that due to the pressure of inspection and its centrality in all aspects of a school's being (whether it is currently being inspected or not), fabrication with regards to documentation, teaching delivery and decision making is so inherent that it now forms part of the normal working discourse and is in itself nothing out of the ordinary. Plowright (2007, p384) states that inspection is 'a game that is understood by all parties' and that 'schools ignore playing the game at their peril'.

2.3 Accountability and Inspection

The relationship between school accountability and school inspections is explored in a short report by Ehern, Perryman & Spours (2014). This report takes into consideration a range of (mainly) contemporary, international literature and considers to what extent all aspects measured by the inspection process (student achievement, teaching and leadership) influence schools policy. It is made clear in this study that there are several considerable factors that have a bearing on the way that schools are impacted by inspection: first, that the context of the school's setting not being taken into consideration, especially for schools serving areas of considerable disadvantage, contributes to the variability of inspection outcomes, a theme supported by Courtney (2012), and second, that inspection and any relationships to student achievement are unclear, stating that schools, through their own data-sets and the threat of a negative inspection judgement can cajole a school to improve outcomes. However, Ehern, Perryman & Spours (ibid.) and Elmore & Fuhrman (2001) note that such improvements in students' achievement rates can be as a result of both a narrowing of the curriculum or more time and energy being spent teaching to the test or on specific groups whose results will be measured and judged as part of the inspection accountability framework (Shaw, Newton, Aitkin & Darnell, 2003). Finally, this study makes clear that the Ofsted accountability framework inspects schools as individual entities and that all performance measures and Ofsted reports are published by single school. This it argues 'legitimises a one size fits all strategy for success to national standards' with the consequences of 'encouraging risk averse behaviour in schools and window dressing of successful rituals' (Ehern, Perryman & Spours, 2014 p6).

In their study, West, Mattei & Roberts (2011) proffer a range of accountability measures that have, since the publication of exam results (a part of the 1980 Education Act) and of the two key policies which came out of the 1988 Education Act, namely local management of schools and the introduction of a national curriculum and testing programme, continued to play a central role at the heart of the English school system (DfES, 1980 & DfES, 1988). The publication of such information, it is argued, would improve transparency and would move accountability out to the key stakeholders and users of the service. This study identifies, via a range an analysis of policy documents, research literature and empirical studies, an accountability framework that recognises the complexity within which schools operate (West, Mattei & Roberts, 2011 p42). However, what is concluded is that the two forms of accountability which underpinned the public accountability arena installed in the early 1980s,

continue. They are 'market' and 'hierarchical' forms of accountability, both of which, it is stated, 'may have major tangible sanctions associated with them' (ibid., p57). Market and hierarchical forms of accountability are inextricably linked. Whilst the latter concerns itself with who or what a school is accountable to (governing bodies, academy chains, Ofsted, the Department for Education) coupled with the publication of data and information, the former concentrates on its impact. This position considers the role of parents and other interested parties, as consumers, who have the technical ability to move their children (and the money that follows them) to another school if they believe that they would receive a better education elsewhere, and it shows that this decision is based on available data and inspection reports, all of which can be disputed regarding accuracy and consistency (Richards, 2001; Waldegrave & Simons, 2014; Sinkinson, 2004; Baxter & Clarke, 2013), brief and over-standardised reports (Fielding, 2001), and achievement rates based on a narrow, un-contextualised set of tests (Parks, 2013; Courtney, 2012; Ready, 2013).

Furthermore, Courtney (2013) clearly identifies that accountability in any school-based situation sits with the head teacher and as a result, explores their experiences of inspection following the implementation of the 2012 framework. Using a mixed-method approach, survey material and interviews with a small group of practising head teachers, he aims to identify the consequences (positive and negative) of the new framework. Much of the analysis presents a picture of a regime that encourages the narrowing effect of inspection. For example, where judgements and commentary were expected to be made under the previous regime but are no longer a category in their own right, in areas such as care guidance and support, community engagement, and pupil health, respondents stated that less emphasis would be dedicated to them. Instead, participants state that their experience of inspection under this new regime focused on value added and the progress of students, with a considerable increase in the use and presentation of data and on information on pupil achievement, especially in English and mathematics. However, whilst value added, as cited earlier, remains part of the framework, removing 'contextual value added' from its data sources gives further rise to the notion of inspection as 'one size fits all approach' (Courtney, 2012; Park, 2013; Ready, 2013).

Crehan (2016) draws attention to alternative synonyms for 'accountability' being 'answerability' and 'responsibility' on the one hand or 'liability' and 'culpability' on the other. She points out that the highest scoring countries in terms of PISA scores approach accountability in the former way, and contrasts this with the blame culture in England and the

USA, which uses the latter approach. She observes further that in these high-performing countries support for underperforming schools, school managements and teachers is systematic and positive in approach, because accountability is seen to be collective rather than individual (Crehan 2016, p220). She further found that in these education systems inspectorate teams were drawn from high-performing educational professionals (usually school leaders, that school performance was monitored not through high-stakes testing so much as through the use of school-level data and less regular national assessments. Inspections, when they took place, did so over a long enough period for the inspectors to form a balanced judgement (ibid, p257). She therefore proposes that one measure to be taken to improve educational attainment in England would be to combine school accountability with school support (rather than sanctions) (ibid, p257). While Crehan's discussion focuses on the years of mainstream schooling, and does not address questions about post-16 education, much of her discussion of accountability could equally be applied in the post-16 sector. Moreover, an approach to inspection that viewed accountability more positively, combined monitoring with support, and was carried out by successful practitioners, would be likely to be more welcomed by college managements, and more trusted by staff.

As stated earlier on this chapter, only a small number of published sources consider the impact of inspection and government policy as they affect the further education sector. One such study (Lumby, 2003) evaluates accountability in further education as it applies to government policy. Much emphasis is placed in this study on the differing groups of 16-19 year old students that attend colleges of further education and sixth form colleges, and it is acknowledged that those studying at the former institutions were more likely to be from socially and educationally disadvantaged backgrounds, studying predominantly vocational qualifications. As a result, it is this sector that, following incorporation as part of the 1992 Further and Higher Education Act, has undergone measurable changes in its accountability structures, not least as a direct response to numerous policy developments (Raffe & Spours, 2007; Coffield, Edward, Finlay et al; 2008; Hodgson & Spours, 2008). However, when compared to the literature on the school sector, which Lumby (2003) concludes is weighted to the presentation of 'staff as victims of an accountability culture...with its inevitable changes in focus and practice to the detriment of teaching and learning (ibid., p9), this study indicates that internal factors are likely to have as much effect on accountability, and therefore on the experience of their staff and students, as external factors.

Finally, Park (2013), in a detailed analysis of the relationship between inspection and school accountability, sets out both the case for the current Ofsted regime being 'profoundly toxic' and 'failing to achieve the stated goal of improving education' and makes two recommendations to support the aim of creating a 'multi-perspective' inspection process (ibid., p13). This hard-hitting study situates the inspection process at the heart of an inspection system the outcomes of which are based on a small and narrow set of externally assessed qualifications or core subjects, such as SATs and GCSEs, not taking account of an individual's academic ability or to the detriment to a wider curriculum – judging all the same way, taking little or no account of the school's context. Or as Birdwell, Grist & Margo (2011 p21) identified in their study, that the majority of schools that they visited spent considerable time and effort in improving the achievement rates for those studying GCSE, but very little on the remaining 50%. In addition, Park (2013) also asserts that the source of this toxicity is firmly rooted in Ofsted itself and its inspections as 'it gives one voice the power to dominate others, terminating the discussion' (ibid., 43) which in turn perpetuates the feeling, real or perceived, that there is an inherent lack of trust between school leaders, policy makers and the wider public.

2.4 Quality Improvement and Inspection

'Improvement through inspection' was the Ofsted strap-line (Institute of Education/Ofsted, 2004 p17). It very much saw its role as not only judging educational providers within its remit, but as having a part to play beyond the formal face to face inspection. Perryman (2010) identifies early in their study that 'school improvement is increasingly defined in benchmarking terms' (ibid., p184). This point is made based on the commentary by Marley & Rassool (1999, p80) that, 'the standardisation of the National Curriculum, the Ofsted inspection criteria and the publication of national league tables have all contributed to the emergence of a nationally uniform systems approach to school improvement guided by the development...of technical expertise grounded in market based realities'. As a result, quality improvement should be considered alongside school effectiveness. This case study school is the same as Perryman's (2002, 2003, 2005, 2006, 2007, 2009), a school that had been placed in special measures and which had subsequently been inspected numerous times before these measures were removed. The participants in this study were senior and middle managers. The collection of evidence was through interviews, observation and participation in an inspection week, similar to the timeframe outlined in Perryman (2009). Much of the direct feedback from the participants in this study suggests that Ofsted had been the driving force behind the school's improvement. A number suggested that it not only provided the stimulant for change, but it focused managers' attention on the areas that really mattered. However, Perryman (ibid., p189) comments that '...Ofsted did appear to be the driving force behind improvements in teaching and learning , but only if the Ofsted criteria were used to judge success'. The impact of this technical approach to improvement and/or success, it is suggested, might be the inability of the school to sustain its improvement whilst not being under direct threat of inspection and, through its conformity, failing to gain the organisational knowledge, skills and experiences to flourish independently within its own paradigm (Jackson & Wallis, 2006).

Published research into any causal relationship between inspection and school improvement is limited (Ehern & Visscher, 2008; Jackson & Wallis, 2006). This is the opening premise by Jones & Tymms (2014, p315), who in their study seek to identify what, through both the analysis of legislation and policy documents, Ofsted's inspection role is in school improvement. The predominant research method used in their study is a policy scientific approach, using multiple data sources such as documents and interviews for its analysis and validation. The parameters within which this study is based are: a focus on maintained

schools in England catering for students aged 4-16 years old, and the school inspection framework which was in place September 2009 – August 2011. What is made clear in this study is the legislation in the Education and Inspection Act 2006 (DFES, 2006) whereby Ofsted 'is to perform its functions for the general purpose of encouraging the improvement activities within the Chief Inspector's remit'. In addition, the framework for school inspections (Ofsted 2009a, p4) outlines the mechanisms by which Ofsted will use school inspection to encourage improvement. Jones & Tymms (2014, p321) note the four stages by which Ofsted promotes school improvement. They are:

- 1) Setting expectations:
- 2) Increasing a school's confidence by confirming the school's own view of itself;
- 3) Recommending priorities for school action planning (and in some cases follow-up);
and,
- 4) Promoting improvement through conversations between the inspection and senior managers.

Furthermore, Jones & Tymms (2014) cite two additional but linked influencers that can galvanise schools to improve: the publication of the inspection report, and parents. The publication of the report, which is available on a publicly available web-site, and which gives overall judgements (numerical and text) with supportive narrative judgements, is a notable external driver for a school to receive a 'good' inspection result. Parents act as the second influencer, using the published reports as a basis on which to choose which school to send their child to (and with funding following the student that has wider ramifications, which are not discussed here) and, as a collective body, they directly pressure the schools, possibly via the head teacher or the governing body, to improve.

Altrichter & Kemethofer (2015) in their study, identify and analyse whether accountability pressure through school inspection promotes results. Their study takes into consideration data and information from approximately 2,300 primary and secondary school principals from a number (seven) of European countries via an on-line survey. It is acknowledged that, due to the size of the cohort and its geographical spread, only the principals of the schools have been surveyed and therefore it is only their subjective responses that are being used. However, the relevance of seeking the views of this one target group is deemed adequate on account of the key relationship between the principal and 'accountability pressure' (ibid., p42). Accountability pressure, which forms the central tenet of this study, is defined as

'pressure on individual schools...to act in conformity with the standards of an accountability system and to take action to improve school quality and effectiveness' (ibid., p37). Their findings identified varying degrees of accountability pressure, with the experience of responding principals from England being categorised as being under the most pressure. In addition, it is reported that inspection systems which have the highest form of accountability pressure, again such as England's, were more likely to be engaged in an improvement process and aware of stakeholder reactions. However, such improvement, within systems identified as having the highest accountability pressure, it is stated, can have a number of unintended consequences, such as narrowing the curriculum or discouraging new teaching strategies. Schools which sit within a system of higher accountability pressure are more likely to engage with development activities. However, there is no evidence provided in this study that suggests that the outcomes of such activity is a better inspection result.

Forrest (2015) offers one of only a small number of studies that give an insight into the interplay among Ofsted, their inspections and the further education system. Their study focuses on the relationship between inspection and improvement in three colleges of further education. Whilst the study concludes that 'there is a stark contrast in the findings between the narrative expressed by...commentators on Ofsted's approach and the experience within the three case-study colleges' (ibid., p308) suggesting that the inspection itself was positive and collaborative, it is acknowledged that some participants viewed Ofsted as 'an anonymous, politically driven body' and that there were 'tensions' when a college carried out activity that went wider than the parameters within the Common Inspection Framework (CIF).

Concrete evidence that suggests a positive causal relationship between inspection and school improvement is a theme that has been explored from soon after Ofsted's inception. See Gibbon, 1996. In an early study on this subject, Chapman (2002), taking the experiences of teachers and managers from ten schools, identifies a number of factors that would suggest that Ofsted inspections, especially of schools in challenging circumstances, add little value. Where they do add value, the study reports, is when inspection is used as a 'lever' to affect change – a point made by the group of head teachers and senior managers – the group who held the most positive view of the inspection process (ibid., p261). However, this group also recognised that the process was restrictive in other ways, such as requiring short-term planning. The experiences of middle managers and teachers were less positive than those of senior managers, with regard to inspection and its perceived impact on improvement. Being

continually ready for an inspection was recognised as a strength of leadership by the teams closest to and within the classroom, in order to reduce a school's 'vulnerability to variability', but beyond this point, middle managers and teachers stated that 'the inspection process had only marginal influence on their teaching...practice' (ibid., p266) and that key areas of weakness had already been identified by the case study schools ahead of their inspections, and as a result, there 'were no surprises (ibid., p266). Clapham (2015) concludes his study by proposing a framework for inspection based on 'context specificity changes at all levels and post inspection relationships' (ibid., p270); this, it argues, will remove the school's focus on 'short-term strategies' and 'quick fixes', to be replaced by tackling serious weaknesses, underpinned by long-term capacity building and quality improvement.

The inspectorates in England have undergone several re-interventions. Under the Learning and Skills Act (DFEE, 2000), the Adult Learning Inspectorate (ALI), was given responsibility for the inspection of all provision covering students over the age of 19, and over the age of 16 for students on work-based learning programmes. This remit excludes higher education provision. Jackson & Wallis (2006), in their study, consider to what extent inspection, specifically within work-based learning settings, is a tool used for improvement. Unlike many of the case studies within this literature review, it is not the principal or teacher at the centre of this piece of research, but the college nominee (normally a senior member of the college staff, who acts as the conduit between the inspection team and their organisation). In this two phased study, 13 nominees are interviewed in the first phase, concentrating on inspection as a mechanism for change and a further 12 in the second. This latter cohort was used to examine the impact of inspection, particularly re-inspection, as a driver to sustainable improvement (ibid., p254). For the individuals and the organisations which they represented, the four day 'snap-shot' inspection was similar to that reported elsewhere, in that it was judgement led. However, for the group of organisations which were subject to re-inspection, under the [then] recently introduced model of re-inspection, the process was far more developmental. Moreover, the timeframe from start to finish of this 'developmental' process was approximately 12 months from the original inspection, which Jackson & Wallis (2006) identified as leading to much more of a partnership model where the nominees were able to garner information and knowledge from the inspectors who, they realised, carried with them considerable insight from their roles as inspectors. Indeed, it would be this two-way process of inspection that the second cohort of participants would identify as having the greatest impact on long-term quality improvement. As Fullan (2001, p74) states:

'educational change takes time and significant rather than superficial change takes even longer'.

2.5 The reliability of inspection: judgement and improvement

Every judgement made as a result of an Ofsted inspection is predicated on there being evidence to support such judgements (Ofsted, 2012 p39). For each of the key themes, inspectors take into consideration a range of primary and secondary evidence and first-hand accounts of leaders, teachers, students and others who are involved in the business of the college e.g. employers. Depending on the timing of the inspection, some of the information required by the inspection team will be in the public domain, for example, success rates as published in the Qualification Success Rate (QSR) report, recently renamed Qualification Achievement Rates (QAR) published by the Skills Funding Agency (SFA) and these will be available to the inspection team prior to their visit. The onus is therefore on the inspection team to identify, triangulate, collate and interpret from a considerable base of evidence and to arrive at an agreed (with themselves) judgement for each aspect being inspected.

From the literature surveyed it is clear that there are two types of inspectors: the team that judge (through inspection) and the team that support, with the specific aim of improvement (normally following a less than good outcome at their most recent inspection).

The judgements that inspectors are asked to make at any college inspection are values based, but there is no explicit explanation or guidance on what these values are or the aims for which they are to be achieved (Richards, 2012). Therefore, like any public body or organisation with a similar regulatory function, its credibility, which for Ofsted has to be its independence and workforce, is paramount (Baxter, 2013). However, inspection is not merely a tick box exercise, or at least that is not the perception that is produced and propagated. Baxter & Clarke (2013) discuss the notion of the 'invisible tick box', balancing the notion between professional judgement and consistency. In a highly charged public arena, to support its credibility Ofsted needs to be both robust and authentic in its approach. This should be less contentious when it comes to the judgement of the broad range of outcomes (success rates, attendance, value added scores and destination), assuming that the parameters of success (and failure) are standard across the sector. For if success rates in their widest forms in one college are judged to be good, surely it stands to reason that

another college with the same or very similar results and data-supported indicators would be judged the same.

For the judgement of teaching and learning, inspectors make direct observations of the process, through targeted lesson observation or shorter learning walks, often supported by an internal manager (to discern the accuracy of the college's own assessment as to the standards of teaching and learning). It is this key process that is used to identify what is happening in the school (Waldgrave, 2014) and therefore, it is this area, under the 2012 inspection cycle for colleges, which has dominated the perceived reduction in standards across the college sector. It should be noted, however, that this policy change towards the centrality of teaching and learning as a de facto limiting grade, coincided with the volcanic and mandatory introduction of English and mathematics to the post-16 curricula for all students who had failed (or whom the school system had failed) to achieve a GCSE grade A*-C. This policy came into being with very little notice and with no work-force development or recruitment plan for specialist English and mathematics teachers (Wolf, 2011; Burke, 2016; DfE, 2014).

The relationship between inspection and improvement on the other hand, in which Ofsted clearly sees it has a role, (Ofsted, 2015 pp6&7), includes:

- setting standards
- reporting on performance
- raising expectations of performance
- providing challenge and the impetus to act where improvement is needed

There is little evidence in school based research which directly links inspection, or even the support from inspectors, to improvement, specifically to student achievement (Ehren, 2014; Chapman, 2002; Baxter, 2013; Jones & Tymms, 2014). However, there is research which suggests that an inspector who is linked to a particular school for support and challenge activities, and who has a more regular, insightful and in-depth understanding of the context in which it operates is more likely to be seen as a partner of improvement (Forrest, 2015); however there are questions about the capability of inspectors to offer ideas of excellence evidenced by their own success in school or college career (Baxter, 2014) for example, contextual value added, an indicator previously used by Ofsted to measure a school's performance compared to other schools with similar catchment and socio-economic profiles.

Ofsted, as part of their Raising the Bar (Ofsted, 2002), removed the consideration of value added within the context of inspection, arguing that the standard of education should be the same for all and to deny that would be to have low expectations, especially for those from disadvantaged communities (Courtney, 2012). Forrest (2015), in his research on inspection and improvement in three further education colleges, stated that inspectors, who were in a college under the guise of supporting it to improve, even though the relationship was still tactical, were more communicative and open to dialogue regarding 'context specificity' (Chapmann 2002, p270).

Ofsted (2017) published a very small study to assess the reliability of their new short inspections. Ofsted claims that short inspections are a 'more proportionate approach to reduce the burden of inspections on good schools' (ibid., p3). They are also clear that this study was 'designed to be the first step towards collecting a body of evidence on the reliability of inspection practice' and acknowledge that there is little empirical evidence which objectively analyses the reliability and validity of inspection judgements. However, it is interesting to note that, although this is indeed a first step, 26 inspections were considered as part of the study, of which the results of only 24 completed inspections were analysed (ibid.). The inspections were carried out in the Autumn of 2015 and the summer of 2016. Furthermore, the scope of this small study is narrow, considering only short inspections, when indeed Ofsted (2016, p6-7) states that in the academic year 2015/16, it carried out over 25,000 inspections, of which 2,468 were in maintained primary schools (Ofsted 2016, p6-7) and its findings, therefore, are in no way representative of the education system as a whole or the primary sector.

Whilst the title of this study by Ofsted suggests a whole sector analysis of the reliability of inspections, the actual analysis, as stated above, concentrated on short notice inspections in primary schools. The methodology used to assess inspection reliability was to deploy two inspectors, comprising a lead inspector and a methodology inspector, into the same school on the same day and for each to inspect the school independently. Throughout the day, each inspector was asked to reflect and note down their reflections, for comparisons to be made. The findings of this study were that of the 24 inspections, in the 22 (91.7%) which were completed, 'inspectors agreed on their final decisions' (Ofsted 2016, p6). And whilst it did acknowledge that the same evidence was interpreted differently in these 22 schools by the two inspectors, the difference was not so great as to change the end result. With regard to the two schools whose outcomes did not match, the reasons put forward in this study

were: different interpretations of the school's self-evaluation document and initial discussions with the senior team (1 school), and the methodology of the study (1 school).

The premise of this study, and indeed short notice inspections generally, is that if a school was judged to be good at the previous inspection, there is an assumption that the school is still good (Ofsted 2016, p10). However, there is no indication in this study as to the foundation on which this assumption is made. Indeed, of the 24 schools in this study, 11 (46%) converted to a full inspection. The conversions in these schools were on the basis that inspectors '[had] insufficient evidence to satisfy themselves that the school remains good' (Ofsted 2016, p11). Of those 11 schools who converted, three were judged to require improvement and three were judged to be inadequate.

Ofsted (2016, p39) acknowledges that this sample size is too small to secure external validity and that the study itself lacked sufficient independent observation across the sample, giving rise to the accuracy of inter-observer agreement. Thus, it is clear that the foundation on which this study was carried out and the potential lack of independent evaluation mean that it cannot provide clear evidence of either the validity or the reliability of inspection judgements.

2.6 Pan-European models of inspection

As identified elsewhere in this study, the purpose of inspection is clear. It is to inform users and potential users about the standard of a provider; to inform the Secretary of State about the standards of education and training in England across the sector and to inform providers what their strengths are and which areas require improvement. Ofsted inspectors make their decisions and judgements at each individual inspection. The interval between inspections depends on the outcome of the previous inspection and can range from one year (for colleges that have been judged to be inadequate) to four years (in a normal cycle for colleges deemed to be good). However, there are examples of where inspections have not taken place for eight or, in a small number of cases, ten years.

It is important at this stage of the study to place the English system of inspection into context, together with those inspection regimes or systems of external evaluation from across Europe. Again, as with most of the literature reviewed in this chapter, the following sections use research from the primary and secondary school sectors (or equivalent) as very little research has taken place in Europe regarding inspection as it applies to the further education sector.

The inspection regimes across Europe vary considerably, but for the purposes of this study they have been categorised into three distinct groupings:

- those that have a similar regime to that in England;
- those that have a regime which is different to that in England; and,
- those that do not have an external assessment (inspection) regime.

European countries that have a similar regime to England's

As it can be seen in Table 2.1 below, there are a number of countries that have a similar inspection regime to that of England, in that the inspection regime is an external assessment, making judgements on school standards (or the standards of the school). However, it is important to note that whilst external or independent assessment is a common attribute across this group of countries, other characteristics differ. For example, whilst in England there is very little engagement between the inspection team and the school being inspected, if any, the relationship in Lithuania is continuous, with visits before and

after the inspection itself. It is suggested that this is to create a culture of continued self-evaluation and improvement. Similarly, in Belgium (German) and Wales there is a greater emphasis on inspection being used as a partnership tool (between the inspectorate and the school) to improve the standard of education, training and outcomes.

Country	Key attributes and/or purpose(s) of inspection
Austria	<ul style="list-style-type: none"> • carried out by a team of external inspectors • responsibility at a federal level • inspection regime covers all school types • national tests are used to identify the quality of school standards
Belgium (German)	<ul style="list-style-type: none"> • carried out by a team of external inspectors who form part of the university sector (lecturers) • focus is on the improvement of outcomes, teaching and standards
Belgium (Flemish)	<ul style="list-style-type: none"> • inspectorate is a non-ministerial government agency responsible to parliament • undertaken by a team of external inspectors from the agency
Czech Republic	<ul style="list-style-type: none"> • inspectorate is a non-ministerial government agency responsible to parliament • undertaken by a team of external inspectors from the agency • executes administrative control of state funds
Republic of Ireland	<ul style="list-style-type: none"> • forms part of the Education and Skills department • inspectors are a range of current and former teachers and managers • focus is to improve the quality of education
Lithuania	<ul style="list-style-type: none"> • carries out external inspections via an external inspection team • supports schools with inspection by visiting before and after an inspection to help create a culture of self-development and improvement planning
Netherlands	<ul style="list-style-type: none"> • as in Lithuania but also assesses the standard of financial management and other regulatory functions
Poland	<ul style="list-style-type: none"> • as in the Netherlands
Romania	<ul style="list-style-type: none"> • carried out by a team of external inspectors • remit for inspection is whole school improvement
Sweden	<ul style="list-style-type: none"> • carried out by a team of external inspectors • focus is on outcomes for learners
Wales	<ul style="list-style-type: none"> • carried out by a team of external inspectors • remit for inspection is quality improvement

Table 2.1 European inspection regimes similar to that in England

European countries that have a different regime to England's

There are also a number of European countries that, whilst having an external evaluation or inspection process, have a range of functions and a framework that differ somewhat from the English inspection regime. Included in Table 2.2 is an analysis of inspection regimes and their key attributes from this second group of European countries. These inspections differ in a range of ways from the English inspection regimes and those cited in the above tables and these differences can be summarised as follows:

- external evaluation focuses on compliance with legislation and/or use of resources, rather than on the standard of teaching and learning, outcomes or the student experience; and,
- different inspection expectations and regimes across different types of schools in the same country.

Country	Key attributes and/or purpose(s) of inspection
Belgium (French)	<ul style="list-style-type: none"> • inspection is focused on a school's compliance with the decree (law) of 2007. The previous regime focused on the standard of teaching
Cyprus	<ul style="list-style-type: none"> • inspections are carried out only on a section of the education system, namely the lower secondary stage
Denmark	<ul style="list-style-type: none"> • external evaluations of schools are based on a screening programme and a monitoring of a school's self-evaluation • direct intervention (inspection) only takes place when a school has been poorly performing for a period of time
Estonia	<ul style="list-style-type: none"> • there are a variety of inspection processes depending on the school type and who has responsibility. The types of inspections include: <ul style="list-style-type: none"> • thematic inspections; • inspections following a complaint; • inspections assessing compliance with legal requirements; and, • inspections assessing the use of resources.
France	<ul style="list-style-type: none"> • inspections in the primary sector focus on the standard of teaching • inspections in the secondary phase of education, but not as a systematic process • schools agree performance targets every three years and inspections are carried out on those that fail to reach their agreed targets
Germany	<ul style="list-style-type: none"> • regular external evaluations take place, but are controlled through the 'local education authority' • the purpose of inspection is to monitor quality and encourage improvement
Hungary	<ul style="list-style-type: none"> • there are two forms of inspection: sub-regional inspections assess standards and national inspections monitor compliance with the law
Italy	<ul style="list-style-type: none"> • a new system of inspection was introduced in 2015 • its emphasis is on self-evaluation and improvement
Malta	<ul style="list-style-type: none"> • the standard of a school is judged through self-assessment and compliance with the national curriculum
Slovakia	<ul style="list-style-type: none"> • inspections in the primary sector focus on the standard of teaching • inspections in the secondary phase of education, but not as a systematic process
Slovenia	<ul style="list-style-type: none"> • school evaluation is focused on compliance with regulation and the law
Macedonia	<ul style="list-style-type: none"> • inspections in the primary sector focus on the standard of teaching • inspections in the secondary phase of education, but not as a systematic process

Table 2.2 European inspection regimes that differ to the one in England

European countries that have no process of external evaluation of school standards

Finally, a small group of European countries, cited in Table 2.3 have no or at least very little external assessment of school standards. The assessment regimes in these countries are unified in their approach, in that they all depend on self-assessment.

Country	Key attributes and/or purpose(s) of inspection
Croatia	<ul style="list-style-type: none">• reliance on self-assessment and national tests to indicate standards and areas for improvement
Bulgaria	<ul style="list-style-type: none">• no framework for the assessment for school standards
Finland	<ul style="list-style-type: none">• reliance on self-assessment which is locally designed
Greece	<ul style="list-style-type: none">• reliance on self-assessment
Luxembourg	<ul style="list-style-type: none">• reliance on self-assessment and quality improvement plans
Norway	<ul style="list-style-type: none">• reliance on school improvement

Table 2.3 European countries with no process of external evaluation of school standards (inspection)

Like much of the literature reviewed in this chapter regarding inspections in England, the research on external school evaluations across other European countries shares the same considerable limitations. These limitations include:

- the sample size being too small, especially when the study purports to review or consider the inspection system of an entire country;
- the narrow choice of research participants, namely principals or inspectors, who will have one view of the effects, conditions and the use of external school evaluations;
- the research methods used to extrapolate from mainly qualitative data; and,
- the focus on subjective measurements within the inspection process, such as feedback, the link between inspection and school improvement, and inspection and emotion.

Furthermore, and as with the literature reviewed on the impact of inspections in England, there is very little research across Europe that applies to the post-compulsory or the further education sector (or equivalent).

In their exploratory study on school self-evaluation and school inspections in Europe, Janssens and van Amelsvoort (2008) seek to identify the relationship between inspection and inspection regime and its influence on the ways that schools evaluate themselves. In this case, they have used as their sample eight education inspectorates in seven countries,

namely, England, Scotland, the Netherlands, Northern Ireland, Belgium (Flemish inspectorate) , Denmark and Germany (Hesse and Lower Saxony inspectorates).

At the outset of their study, Janssens and van Amelsvoort (2008, p16) state that across Europe there is 'no one definition of school self-evaluation or quality assurance'. However, they do acknowledge that inspection is 'defined as the process of periodic, targeted scrutiny carried out to provide independent verification, and to report on whether the quality of schools is meeting national and local standards' – a definition taken from an earlier study on supervising the quality of inspection by Janssens (2007, p47). It is against this definition of inspection and the definition of school self-evaluations or quality assurance as being 'a process, directly or indirectly aimed at school improvement' (Janssens and van Amelsvoort 2008, p16) that this study is based.

Janssens and van Amelsvoort (2008) in their study, draw a number of conclusions about the relationship between the school self-evaluation and inspection. They conclude that where school self-evaluation is incorporated into the school inspection system, there is considerable influence by the inspectorate on its design and content. That is to say that there are examples of where the self-evaluation document and even the style of writing closely follow that of the inspection framework, such as in England, Scotland, the Netherlands and Northern Ireland. In these cases, Janssens and van Amelsvoort (2008, p20) assert that the school self-evaluation orientates itself towards both improvement and accountability. Conversely, in those countries where school self-evaluation does not play such a compelling role in inspection, or indeed any role at all, such as in Germany, Denmark and Belgium, its focus become more 'improvement-orientated' (ibid).

However, the basis of the findings incorporated in their study comes from an analysis of n=25 school self-evaluation reports, and whilst the study does not spell out the breakdown of reports by country (to ensure that there is an appropriate distribution across the countries or inspectorates within the study), there would be an average of n=3-5 reports per country or per inspectorate. And, although Janssens and van Amelsvoort (2008, p19) seek to provide assurance regarding the validity of their study by stating that their findings have been shared with a group of 'key inspectors', who were not only research participants but also an expert panel, it is therefore not possible to assume the accuracy of the study's findings with any great confidence.

In another pan-European study, Gustafsson, Ehern, Conyngham et al (2015) attempt to identify how school inspections impact on the improvement of schools. They place their study within a context that states that whilst inspection or external evaluations are used as a tool for 'controlling' and 'promoting' the quality of schools, there is little or limited research within this area (Gustafsson, Ehern, Conyngham et al 2015, p47). The premise of their study is to test an established conceptual model designed by Ehern, Altrichter, McNamara and O'Hara (2013) which was developed by analysing inspection policies and frameworks underpinning school inspections in six European countries. The six countries used for this study were the Netherlands, England, Sweden, Ireland, Austria (the state of Styria) and the Czech Republic. In the case of Ireland, there is no reference as to whether the regime used is that of Northern Ireland or the Republic of Ireland. Gustafsson, Ehern, Conyngham et al (2015, p50) state that the inspection regimes chosen for this study share similar attributes – 'the embodied observation of classrooms and schools'.

This study, like the one above by Janssens and van Amelsvoort (2008), uses at its foundation the legislation, policy documents and inspection framework documentation to define the 'common mechanisms to link school inspections and their intended outcomes' (Gustafsson, Ehern, Conyngham et al, 2015 p48). These mechanisms include aspects of an inspection such as: frequency; the regular cycles of inspection; the setting of standards and thresholds; and, feedback. This documentary analysis or 'programme theory' is used to identify the range of intended outcomes by inspectorate, including the level of central control and 'high stakes' inspection in England and the Netherlands (ibid). And, in the same vein as Janssens and van Amelsvoort's (2008) study, Gustafsson, Ehern, Conyngham et al (2015) seek data directly from actors involved in the inspection process. In this study, as in the one above, information is sought from school principals. They consider that school principals are the best 'informants of change in school' as they are believed to be the key players in both preparing the school for inspection and implementing any subsequent action plan (ibid).

To gather information directly from school principals from both primary and secondary schools, Gustafsson, Ehern, Conyngham et al (2015) designed a questionnaire of some thirty-four questions. The questions were grouped under the following headings:

- setting expectations;
- stakeholder sensitivity to reports;
- accepting feedback;

- improving self-evaluation;
- improvement actions for capacity building (teacher participation in decision making, teacher co-operation and transformational leadership); and,
- improvement actions for school effectiveness (opportunity to learn, assessment of students, assessment of school and structured teaching).

The number of respondents by question is not given within the study and the targeted sample size by country varies considerably. For example, in England the targeted sample was 211 primary schools and 211 secondary schools that were 'close to the threshold for monitoring inspections'. (ibid, p51). So, whilst it is suggested that the response rates from the England sample are 29.38% (primary) and 20.10% (secondary), it is known that in any one year Ofsted carries out over 2,400 primary inspections and approximately 650 inspections in secondary schools, most of which schools have been judged to be good and therefore close to threshold for monitoring inspections (Ofsted, 2016). Therefore, it can be argued that the original sample size with regards to the sample from England is very small and the response rate even smaller. For the remaining countries, the target sample is inconsistent with examples of all schools being targeted in Ireland and Austria (Styria) and those in the Netherlands being spread across three threshold groups – no risk, risk and high risk. Furthermore, the response rate is wildly different in each country and ranges from the lowest at 4.18% to the highest at 77.2%. It should be noted that the authors cite the variation in response rate as a limitation of their study, but no recognition is given to the sample size.

There are further, quite serious limitations highlighted by the authors of this study. Like the study by Janssens and van Amelsvoort (2008), their study relies solely on the views of principals, giving a single perspective of inspection. Another limitation is the short duration of the study and finally, and arguably most importantly, the generalisation of the findings. Their findings, which as outlined above are based on documentary analysis (policy analysis), a survey of one very small group of participants (in an inspection) over a short time frame, are arguably inane, in that they state that 'the impact of school inspections on the quality of education is driven by the setting of expectations, standards and norms, with self-evaluation and encouragement of capacity building and better teaching and learning' Gustafsson, Ehern, Conyngham et al (2015, p55).

The following three studies reviewed are from Sweden, where the Swedish School Inspectorate, by comparison to Ofsted, is a relatively new organisation, only starting to operate in the autumn of 2008. Each study considers the effects of inspection. One, mainly descriptive case study, compares the experiences of inspection in two schools, whilst the other two studies examine inspection within the political context. In these latter two studies, one considers the political context leading up to the establishment of the Swedish School Inspectorate and the other considers its influence post establishment.

Segerholm and Hult (2018) in their case study of two schools, explore how governing, which in the context of their study refers to leadership and leaders, learns through an inspection process. Specifically, is to identify what is learnt, how it is learnt, who learns it and with what motives. As a result, it is stated that the design of the study made it possible to 'study the inspection process as a whole' and therefore 'directly relate what was learned, reactions and emotions' within each school's context (Segerholm and Hult 2018, p128). The research methods used to collect data are broad. As in the studies outlined above, the starting point for analysis is the documents shared with the inspection team prior to the inspection. This is followed by direct observation of the inspection in each school once it is under way and it concludes with a range of interviews – with head teachers, teachers and inspectors. However, it is unclear under what conditions observations were carried out; how many teachers were interviewed (although in a footnote it suggests four or five); or, how many inspectors were interviewed.

Segerholm and Hult (2018, p128) state that these two schools were chosen out of a larger group of eleven to 'illustrate variations...in both local contexts and conditions and different types of school'. The context of the two schools was that one was a large urban secondary school with 500 enrolled students and the other a small public rural school with 50 enrolled students. Much of the content of their study is descriptive of the process undertaken by the researchers. Their summary of their findings is similar in nature to that in the study by Gustafsson, Ehern, Conyngham et al (2015) in that they offer very little, if any, new knowledge. For example, they conclude that the better the feedback (from inspectors) to teachers and head teachers the better it is accepted. Or, in the case of the large urban school which did not receive a favourable inspection outcome, there were feelings of 'disappointment' and feelings of '[low] self-esteem' (Segerholm and Hult 2018, p137). This point is also made in a recent study by Behnke and Steins (2017) whose qualitative study analysing principals' reactions to inspection feedback in Germany concludes that receiving a

positive inspection result was cited as the main reason for accepting the feedback from inspectors. However, their study, although limited by the small response rate (n=20) makes a further, important observation in that the feedback accepted by principals where the inspection has resulted in a positive outcome, gives no indication as to whether it is valuable or realistic (Behnke and Steins 2017, p101).

Finally, in their study, Segerholm and Hult (2018) present one overall arguably obvious conclusion, which is that 'inspection cannot be a matter of applying the same techniques in an instrumental way' because there are human beings involved at every stage of the process (Segerholm and Hult 2018, p 138). And that it [inspection] is always based on human practice and values (Scwandt, 2002).

The inspection regime in Sweden is similar to that in England. The Swedish School Inspectorate design and implement an inspection framework that is similar to that of Ofsted. Schools are externally evaluated by the inspectorate, as part of a cycle; they are judged on the standard of teaching, learning and assessment, achievement rates and how well they are led and managed; and every report following a school inspection is publicly available.

However, the inspection regime in Sweden as it is today under the scope of the Swedish School Inspectorate, when compared to the English inspectorate Ofsted, is a relatively new one, having come into existence only in 2008. The creation of the Swedish School Inspectorate comes at the end of a period which saw the 're-installing' and 'reinforcing' of school inspections, starting in 2001 (Ronnberg 2014, p385).

In their studies, both Ronnberg (2014) and Carlbaum (2016) analyse the political discourse and context surrounding the development of school inspections. In her study, Ronnberg (2014) analyses the political motives for increased central control and the external evaluation of school across two periods namely 2001 to 2003 and 2006 to 2008. These two periods cover two election cycles and therefore two different governments. The first period was under a social democratic government and the second period was under a non-socialist government.

The specific premise of this study is to consider if there is any difference in the approach to inspection under two different governments and if so, what justification there might be for such a difference. Documentary and descriptive analyses are the two research methods used

in this study. Ronnberg (2014, p385) states that the documents analysed include '34 bills, parliamentary minutes and motions, commission report and auditory notes' in addition to 'official texts relating to the two decision-making processes...to the final debate and decision in parliament'. As a result, their study clearly maps through official documents the to-ing and fro-ing between both governments, specifically on emphasis and prominence. For example, Ronnberg (2014, p390) states that during the second period of government 'educational results and performance' became a central focus of inspections, whereas under the first administration, inspection had been focused on educational quality and equality. However, the conclusion of this study is that, despite changes of emphasis or prominence on different aspects of inspection, both political parties came 'across as unified in their need for increased national control over education and school inspection' (ibid 2014, p392) - a picture of inspection and its development that is arguably reflective of the inspection regime in England under both Conservative, Labour and Coalition governments (Clark 2013, Baxter 2013 and Waldegrave and Simons, 2014).

In her study, Carlbaum (2016), like Ronnberg (2014), traces, over time, developments across the educational landscape and how inspection, which had been abandoned as a result of wider political reforms (such as marketisation, decentralisation and managerialism) was reinstated in 2003. Inspection in Sweden is now, as has been mentioned, a heavily regulated process undertaken on behalf of the state by the Swedish Schools Inspectorate. The reinstatement of inspections was the result of 'increased debate and criticism of poor performance, increased differences among schools and [a] lack of relevant information for students and parents to make an informed school choice' (Carlbaum 2016, p137). However, the scope of their study draws on the concept of 'the politics of blame' and specifically the relationship between accountability and blame when gaps in student performance are identified within the context of an inspection (ibid). Carlbaum (2016) puts forward the idea that in Sweden, the public debate around the crisis in educational standards in the early part of the twenty first century centred on the 'attitudes' or 'low expectations' of students. In particular, it is noted that the Swedish Schools Inspectorate showed that such excuses for 'low expectations' and 'low results' are particularly prevalent among students from ethnic or lower economically social backgrounds (Carlbaum 2016, p133 & SSI, 2011).

The research method and material used in this study are the use of secondary data for primary research. Decision reports (summary reports of school standards) drawn from individual schools within four municipalities, inspected in 2011 or 2012, have been analysed

(Carlbaum 2016, p135). It should be noted that there is no reference in this study to show either how many inspections took place within the time period mentioned or how many reports were analysed. In addition, as the scope of the study focuses, at least in part, on 'performance gaps' and the context regarding 'education achievement' and inspection, there is no mention of any exam data being collected and analysed to validate inspectors' judgements (Carlbaum 2016, p137).

The initial finding put forward by Carlbaum (2016, p138) is that in all decision reports 'large performance gaps' existed between schools. This measurement is derived from the pass rates in all subjects and the average grade of eligibility for upper secondary school programmes against the national average. Carlbaum (ibid) states that where contextual factors are considered to have affected achievement rates, they are in schools which are reportedly underperforming and which, despite efforts to 'raise standard of achievement, have remained inadequate'. Furthermore, Carlbaum (2016, p139) states that in differing aspects of inspection, the politics of blame is being levelled at a range of organisations, processes and individuals, such as local authorities (material context), principals and teachers (professional context) and schools (situational context). In the case of local government, the Swedish School Inspectorate, it is suggested, attributes inadequate student performance, in part, to the distribution of resources. However, the distribution of resources is not the only aspect identified as a weakness in this area; attendance zones and differing staffing ratios are also cited. Therefore, Carlbaum's (ibid) assertion that by emphasising resources in its reports the Swedish Schools Inspectorate is blaming local government, an attribution of blame which is arguably unfounded.

With regards to both the professional and situational contexts, whilst Carlbaum's study (2016) clearly identifies evidence which supports assertions of blame being attributed to principals, teachers and school, it also provides evidence which contradicts its findings. For example, it is acknowledged that the situational context is 'complex' and the summative conclusion is that linguistic and socio-economic explanations from teachers and managers are regarded by the inspectorate as an 'excuse'. However, evidence presented in their study, quoted from one of the decision reports, indicates that 80% of the ninth grade population achieved the standard in Swedish as a second language. As a result, the inspectorate puts forward the case that linguistics and language are not necessarily an automatic reason for poor educational results, and that it might indeed be 'low expectations at the school' that result in low achievement rates. (2016, p142).

2.7 Conclusion

The literature reviewed in this chapter presents a comprehensive understanding of the nature of inspection in England and its impact, whether intended or otherwise, on individuals, organisations and the education sector as a whole.

For individuals, much of the research in this chapter indicates that inspection is wrought with emotional turmoil – due to the fact that teachers feel that they are being tested or part of a game in which they are unclear about the rules. Despite these facts, evidence in the more long-term studies such as Perryman (2006) indicates that teachers will perform for inspection, and indeed for any internal process that mirrors inspection, even if it meant fundamentally changing their 'normal' teaching practice. However, it is worth noting that the emotional effects on staff, identified in a number of studies as stress and/or anxiety, are not limited to the time that inspectors are in school – it is when inspection or to be more precise the threat of inspection is used by managers as a leadership tool in the day to day running of a school.

For organisations, there are examples in the literature reviewed in this chapter that where an inspection goes well (where they are judged to be at least good) schools and colleges are more likely to be content with the process of inspection. This also applies to the feedback given to school leaders in England and in other European countries, in that it is welcomed when the outcome is a positive one. However, there is evidence to suggest that when an inspection concludes and the judgement is 'requires improvement' or 'inadequate' or equivalent, a school's engagement with the inspection process has the ability to become less productive. Moreover, there are studies in this chapter which indicate that the road to recovery becomes more complex as a result. For example, when inspection is used as an accountability tool there is often a change in the school leadership, processes and organisational culture which are not always positive. In addition, similarly to the points above, where Ofsted or the threat of Ofsted permeates a school or college it has the ability to be a distraction and to take away time and resources from areas that are perceived as not being important.

For the education sector, inspection of schools in England since the inception of Ofsted, and more recently of colleges acts as the omnipresent force. As a result, the literature in this chapter illustrates that inspection has a wide-ranging impact across the education sector. In

the studies reviewed above, there is empirical evidence that inspection has a role within accountability processes and inspection has the ability to affect leadership, management and teaching practice. Furthermore, inspection can affect quality improvement processes and a school's approach to improving standards.

Research in the field of Ofsted and inspections in England and inspections (or similar) in European countries is limited. There are no studies in England or across Europe, that have been identified, that: considers one entire sector of the education system; the data and information used by inspectors to inform their judgements (either qualitative or quantitative); or, tested objectively to what extent inspectors' judgements are founded on evidence.

A number of studies reviewed in this chapter use one or two schools to undertake their research. These case studies are often based on a school's exposure to one or two inspections and mostly where their result has been a negative one. Furthermore, even the studies which look at the impact of inspection over time, they too map the journey and experiences of inspection in one school. In both cases this often leads to the findings being limited to either generalisations or highly specific experiences.

Even in larger studies, particularly those which seek to investigate the role and impact of inspection in different European countries as well as specific studies which consider the same question, the response rate from participants (often head teachers or principals) is regularly low. In these studies, the findings, like those coming out from the case studies are generalised and sometimes obvious as acknowledged by studies' authors.

Finally, there are two areas which merit particular note. The first is that the overwhelming majority of the literature reviewed in this chapter is from the school sector. Very little research exists relating to the further education or post-compulsory sector and its relationship with and to inspection. Secondly, hardly any research exists which considers the theme of outcomes for learners or its consistent parts of student achievement, attendance, value added or progression/destination in any part of the education system in England or in Europe.

As a result, it is my study, in the form of this thesis, which determines how consistently data and information are used by inspectors (in England) to make judgements on outcomes for

learners in the further education sector. To achieve this, the following key questions have been identified:

- 1) How consistently are data and information used by inspectors to evaluate and make judgements on the performance of students (success rates)?
- 2) How is attendance evaluated by inspectors? Is the evaluation consistent across inspections?
- 3) To what extent is value added data used by inspectors to evaluate and make judgements on student progress?
- 4) How are progression and destination evaluated by inspectors and is this evaluation consistent across inspections?
- 5) How transparent are inspectors about the type and validity of data and information used when evaluating, judging and reporting on outcomes for learners?

Chapter 3: Research Methodology

3.0 Introduction

My study analyses how consistently inspectors use data and information when evaluating student performance and making judgements on outcomes for learners in further education colleges - although the premise of this study might be applied to other educational sectors, such as the primary and secondary sectors, and higher education.

Inspection reports, and the judgements on educational and academic outcomes made by inspectors, are predicated on consistency, so that those who use inspection reports to make decisions about where to study, for example, can do so with confidence – confidence in knowing that if they are comparing one college to another, they are doing so on a like for like basis. However, it can be argued that, whether it be the 25,000 annual inspections or the 165 further education college inspections used in this study, consistency in the evaluation of student performance and in the judgements made by inspectors is hard to achieve. It is important, therefore, that any level of inconsistency becomes known.

It is clear from the literature reviewed in the Chapter 2 that any form of research relating to the further education sector is limited and relatively small scale when compared to other areas of education such as schools and universities. Moreover, it is made even clearer from the literature reviewed in the previous chapter that most research to date about inspection is founded on emotion: how people, specifically head teachers and teachers, feel about inspection and how the outcomes of inspection impact on individuals and the wider community. In the majority of studies the ontological stance has been constructivist and their epistemological position is that of interpretivism. As a result, the qualitative research methods deployed in these studies and used to support the evidence, findings and recommendations within the literature reviewed have included: longitudinal case studies of one school or two schools, questionnaire surveys of teachers and head teachers, comparative case studies including observation and researcher participation across two or three schools, and the comparative experiences of individual teachers.

My study, by comparison, is large in scale. It textually analyses 165 inspection reports from further education colleges that were inspected over a three year period. It is a study that

also uses a considerable amount of secondary data, which has been taken from a series of large datasets or a smaller set of secondary data directly from colleges.

It should be noted at this stage that the secondary data used in my study is descriptive statistics and it is used in exactly the same way as it was in colleges and by inspection teams. The measure of central tendency used in the creation of both the validated data (success rates and national averages) and in the un-validated, local data (attendance, progression and destination) is the mean average. Using such aggregated data, Smith (2008, p91) suggests, is 'probably the most straightforward way of using secondary data' and one of its strengths is 'their value in providing a context to in-depth work as well as enabling researchers to combine data from different sources'. The sources of and types of data are explicitly referenced throughout this chapter.

I have opened this methodology chapter with a philosophical position, the basis of which is my experience of leading a number of inspections, as part of my role as a senior manager. It is this experience which has given me, and continues to give me, insight to what data and information are used and how they are used by inspectors prior to and during an inspection. And it is this first-hand experience which has informed the research strategy used to answer the research questions raised at the end of Chapter 2.

As a quantitative study, beyond the philosophical position and the rationale for choosing the timeframe and sample, the rest of this chapter, which concerns itself with both textual and statistical data collection, their analysis and their sources, is preceded by the process and result of the hypothesis test.

3.1 Researcher's Philosophical Position

My study aims to determine, objectively, how inspectors use data and information to evaluate student performance and to make judgements on outcomes for learners. For this purpose I adopt a positivistic and objective stance that is 'concerned with facts rather than impressions' (Saunders, Lewis & Thornhill 2016, p128), and that these facts are consistent with the notion of observable social reality as identified in this study through the inspection reports (ibid.; Remenyi, Williams, Money, et al 1998). The data used is the same data as inspectors used when making the judgements on outcomes for learners in the cycle of inspections investigated.

My study is thus based on a highly structured approach to the collection and analysis of both textual and statistical data, which forms the foundation of the quantifiable analysis in Chapter 4 and which underpins the discussion and recommendations in Chapters 5 and 6 respectively. This approach, which seeks to remove any influence of emotion through the collection and analysis of secondary data, is in sharp contrast to the subjective, qualitative methods used in much of the research published in this area, as identified in Chapter 1.

3.2 The rationale for choosing the timeframe and sample

Ofsted has been inspecting colleges of further education with complete autonomy, that is as an individual inspectorate and not in partnership with similar organisations, since 2007, so to all intents and purposes there was a new system of inspection, as identified in the handbook for the inspection of colleges (Ofsted, 2007). As a result, I decided early on that there were too many variables in the first wave of inspections (between 2007 and 2012) that might render the foundation of any analysis and its consequent conclusions, invalid. Such variables included the merging of staff teams from other inspectorates whose professional behaviours were noted as being 'softer' than those in Ofsted, and the lack of historical data and information on how that data was interpreted and reported by inspectors from more than one agency.

Therefore, it was important that I identify a period of time in the inspection cycle which:

- Was covered by the guidance of one common inspection framework (and handbook for inspecting colleges);
- Allowed Ofsted sufficient time to assimilate its inspection team under one culture, set of expectations and standardised processes;
- Used one system to measure the success of students, particularly those aged 16-18; and,
- Also provided the data for colleges in the two years prior to the timeframe chosen to be used, as a way of identifying success rate trends prior to inspection.

Taking all of the above into consideration, I concluded that the most appropriate timeframe which could be used for the purposes identified in this thesis was September 2012 to July 2015. This timeframe is supported by the facts that:

- It was covered by one Common Inspection Framework and Handbook for Inspection Colleges (Ofsted, 2012);
- Ofsted had been the sole inspector and responsible for its own team of inspectors for five years;
- The success rate methodology for 16-18 years olds had remained the same for the two years before the implementation of the 2012 framework and, despite the methodology change in 13/14, it was comparable for the duration of this study; and,
- For the most part, adult success rates were measured consistently up until the 13/14 reporting year, affecting a small proportion (n=18) of possible inspections, not able to be analysed as part of my study.

I identified that during this time frame, using the selection tool on Ofsted's web-site and selecting the provider type of 'general further education and tertiary', 185 full inspections took place. However, the following inspections were discounted from the overall sample for the following reasons:

- 9 inspections took place at a college designated as a land based college;
- 2 inspections were in colleges which were titled as an adult college;
- 1 was designated as a specialist art and design college; and,
- 1 was designated as a specialist college for students with learning difficulties and or/disabilities.

A further 12 inspections were discounted from my study due to a lack of continuous success rate data. Where this was the case it was due, in the main, to the college(s)' having merged or been taken over by another college. As a result, the number of colleges which fall within the scope of my study is 165. These 165 inspections were carried out in 150 further education colleges. 15 colleges were inspected during the timeframe of this study, due to their grading (either 'requires improvement' or 'inadequate') at this first inspection. As at September 2016, there were 216 general further education colleges in England (AoC, 2015)

3.3 Hypothesis Testing

The purpose of this thesis is to identify how consistently data and information have been used by inspectors to evaluate student performance and to make judgements on outcomes for learners across the further education college sector. Payne & Payne (2004, p112) state that 'a hypothesis is a reasoned but provisional supposition about the relationship between two or more social phenomena'. The scope of this study is predicated on such a reasoned and provisional supposition based on my time as a college nominee and my experience of working directly with inspectors who approached the use of data and information very differently. This opinion became my working hypothesis. However, to ensure that the premise of my study was not founded on a set of coincidental experiences or on a set of outpouring assumptions based on my own and others' experience of inspection, I analysed a small sample of reports. My initial analysis focused on evaluative statements and judgements made by inspectors regarding success rates of students aged 16-18. I chose the first academic year of my study for which there were 62 inspection reports. This not only gave me the opportunity to analyse the success rates across two QSR years, but it also gave me a sufficiently sized sample to decide if there was merit in expanding the analysis to the remaining two academic years and across the provision.

NB. Whilst there is a full explanation of what success rates are and how they analysed on page 86, below in Table 3.1 is a simple illustration preceded by a brief explanation of how success rates are calculated.

The SFA (2014, p10) states that 'success rates are calculated as the number of learning aims achieved divided by the number of starts.' They can also be calculated by multiplying the achievement rate by the retention rate. This calculation is illustrated in Table 3.1.

Number of people starting a course	Number of people retained	% of those retained	Number of people achieving their qualification	% of those achieving (from those retained)	Success rate % (Ret x Ach)
10	8	80	8	100	80

Table 3.1 An illustration of how success rates are calculated, as per SFA guidance.

The results of the hypothesis test can be seen in Tables 3.2, 3.3 and 3.4 below. They are also included in the data tables in Chapter 4, as part of the full analysis.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13	4	78.9 – 86.8	7.9	79.7	-0.8 to +7.1
11/12	2013	12/13	11	77.1 – 90.3	13.2	82	-4.9 to +8.3

Table 3.2 The data range of colleges inspected against the documentary analysis (Code 6):
The success rates for 16-18 year olds is above average/high

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average) %	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13	4	76.1 – 83.5	7.4	79.7	-3.6 to +3.8
11/12	2013	12/13	17	77.2 – 84	6.8	82	-4.8 to +2.0

Table 3.3 The data range of colleges inspected against the documentary analysis (Code 33):
The success rates for 16-18 year olds is in line with the national average or just above/below it

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13	1	70	-	79.7	-9.7
11/12	2013	12/13	12	71.2 – 87	15.8	82	-10.8 to +5.0

Table 3.4 The data range of colleges inspected against the documentary analysis (Code 60):
The success rates for 16-18 year olds is below the national average, too low or poor

The above tables reference a total of 39 colleges (n=39 number in the sample) where the performance of students aged 16-18 was analysed. This data suggested that there were 23 inspections where no clear judgement or evaluative statement on student success for this age group was made. It also indicated that there was considerable variation in the differences in success rates in each of the years and across the evaluative statements. For example, in Table 3.2 of the 11 inspection reports in 2013, where the performance of students aged 16-18 was identified as being high or above the national average, the data range, between the lowest and highest success rate, was 13.2%. A slightly higher data range was found across the 12 inspections that also took place in 2013 (Table 3.4), where the performance of students aged 16-18 was judged to be low or below the national average.

In addition, there was also evidence to suggest that the relationship between the success rate and the national average was variable, with some large ranges. For example, of the 11 inspections in 2013 used in the previous paragraph (Table 3.2), the success rate data range, when compared to the national average, varies from -4.9% to +8.3% which, at the lowest end at least, is at odds with the judgements made by inspectors. And, again of the 12 college inspections in 2013 (Table 3.4), the relationship between success rates and the

national average ranges from -10.8% to +5.0, which at the top end bears no resemblance to the judgement that success rates are poor or below the national average.

The result of these initial analysis and findings was that there was some, and in a number of cases, considerable variation across success rates, the national average and the judgement made by inspectors. As a result, a full analysis spanning three academic years, or the lifespan of the handbook (Ofsted, 2012), was undertaken. And, in order to analyse the remaining aspects under the theme 'outcomes for learners' which uses as its evidence source validated and un-validated objective data, information was sought directly from colleges via a freedom of information (FoI) request.

3.4 Gaining access to data and information

Smith (2008, p8) acknowledges that secondary data for the purpose of analysis has been used in social research for over three hundred years, although its recognition and potential, according to Heaton (2008, p33), have been on the rise since the mid 1990s.

My study is based on the collection, analysis and presentation of secondary data for primary research which Hakim (1982, p1) describes as the 'further analysis of an existing data set which presents interpretations, conclusions or knowledge additional to, or different from, those presented in the first report on [an] inquiry and its main results'. However, my study is not a mere re-analysis of the data and information which inspectors might have used when evaluating performance and making judgements either to confirm or to argue for alternative judgements or grading, as a kind of blind test or second marker. The secondary data used in my study is used to assess with what consistency inspectors use student success rates, for example, and what relationship that success rate data has to the judgement made by inspectors, as cited in the inspection report. It should be noted that the data on which the judgements are based is very rarely to be found in an inspection report.

Therefore, underpinning the analysis within this thesis are three sets of data, namely:

- Ofsted inspection reports;
- Student success rates; and,
- Additional data on attendance, value added, progression and destination.

Two of the above sets of data are available publicly and the other has been sourced directly from colleges who were inspected between September 2012 and July 2015.

Ofsted Inspection Reports: following each individual college inspection (and all other inspections), a written report is published on the Ofsted web-site. This report summarises the key findings of what the inspectors observed or identified from their time at the college, structured under each of the inspection themes. In addition to the written elements of the report, each section is preceded by a judgement, in both numerical and text formats. Prior to publication, each organisation inspected would have seen the draft version of the report and would be allowed only to check and suggest alterations based on factual accuracy. Once published, these reports are made available to the public.

Student Success Rates: the data for student success rates in further education colleges is publicly available and published on the government's web-site. Following the collection and analysis of results from all further education colleges for the latest academic year, a list of all colleges, and indeed independent training providers, is published and made available to the public.

Additional Data: as identified above, the only data which the theme of outcomes for learners is assessed against, and which is available to the public is on student success rates. However, much of this category, which is the basis of this thesis, is not collected centrally or published in any meaningful way locally. Therefore, it was essential for the research to access this information via a freedom of information request. The subject of the data and information requested concerned itself with student attendance, value added, and student progression and destination. The identified person in each college from whom this data could be retrieved varied considerably, so FoI emails were sent to the person identified on each of the colleges' web-sites. Typically, responses were received from the principal, vice principal, director of management information systems or the FoI officer (or people in similar type positions).

3.5 Inspection Reports

Following a college's inspection (or indeed that of any other provider under Ofsted's remit), a report is written by the lead inspector, which takes into consideration all of the evidence collected by the inspection team. At this stage, a college would have already been informed verbally of the grades awarded for each aspect being inspected, along with the main findings, including strengths and recommendations for improvement (Ofsted 2012, p36). This would have taken place on the final day of inspection.

All inspection reports are published on Ofsted's web-site, most usually within 25 days from the end of the inspection - although there have been a small number of cases which have fallen outside this time scale. See Camden, 2017, Robertson, 2017 and Offord & Burke, 2017. The aim of these reports underpins the first main purpose of an inspection which is to: 'provide users with information about the quality of provision in England; to help inform them about the providers they use or about the providers they may use in the future' (Ofsted 2012, p5). As such, a report provides a summary of the inspection, using the evaluation schedule for the inspection of further education and skills as its framework (Ofsted 2012, p39).

The reports analysed as part of my study maintained a general formula in the way that they were designed, ensuring that key data could be captured. For the purpose of this study these related to:

- The date of the inspection;
- The college name (although the name of the college has been anonymised, it has been used to identify geographical location);
- The grades awarded for each aspect inspected;
- The evaluative text judging each aspect under the theme 'outcomes for learners'; and,
- The name of the lead inspector.

The date of the inspection was not only used to locate the academic year in which the inspection took place, but it was also used to identify the validated success rate data on which the evaluation of success rates would have been founded. For most of the time throughout the timeframe used in my study, these data reports were entitled Qualification

Success Rate reports (QSRs) and subsequently become Qualification Achievement Rate reports (QARs). Whilst the nomenclature may have changed, the methodology has not. Therefore, to ensure continuity of language, the success rate data reports are referred to in the rest of my study as QSRs.

3.6 How are the reports analysed for student success?

At the hypothesis testing stage, it was essential to identify a quantifiable coding structure which would be used to analyse the text within the inspection reports summarising the inspection teams' evaluations and judgements on student performance under the theme 'outcomes for learners'. This quantifiable numeric data, referred to in my study as textual analysis, acts as the central statistical tool, which, when compared to the student performance data, for example, on success rates (used by inspectors to make such evaluative statements and judgements) forms the statistical and analytical framework of my study. The detail on how these codes were constructed is outlined later in this chapter and a full list of these codes can be found at Appendix 1. At the hypothesis testing stage a sample of 30 inspection reports from across two regions was analysed, pertaining to the evaluation and judgements made on the success of students aged 16-18. This analysis indicated that judgements were made using three broad statements. These statements, or concepts, were identified according to the frequency of the same, similar and/or related terminology and phrasing used by inspectors, in what (Carley 1993, p82) refers to as a 'sociolinguistic community' (this process of textual analysis has been applied to the other aspects of my study, namely attendance, value added, progression and destination). For success rates, these were:

- That students aged 16-18 achieved well, that their success rates were above the national average or that they were high;
- That students aged 16-18 achieved in line or broadly in line with the national average, or that their success rates were slightly above or slightly below; and,
- That success rates for students aged 16-18 were poor, below or 'significantly' below the national average or that too many students did not complete their course.

In addition, within the same sample of 30 inspections, some reports, when making an evaluative statement or judgement, also referred to a specific level, a qualification or a

proportion of the student body. And, if no specific level of qualification was cited, it was presumed that the judgement referred to the whole of the provision.

Based on this initial analysis and for the purpose of recording the textual analysis systematically, 81 codes were generated using the structure set out in Table 3.5.

Having moved past the hypothesis testing stage, it was found that two other student groupings were consistently reported against: students aged 19+, and those work-based learners studying for an apprenticeship. For those aged 19+ the type of reporting was similar to that for those aged 16-18, so a further 81 codes were generated, covering all of the possible permutations outlined above. However, the way the inspection report identified the evaluation and judgements made on apprenticeship provision was considerably different, in that hardly any reports categorised this cohort by age. Furthermore, the apprenticeship category in the inspection reports analysed has the same reporting status as a qualification (rather than the aspects that make up an apprenticeship framework having been assessed individually). To record the textual analysis consistently for this provision, 27 codes were generated.

Therefore, in order to analyse each Ofsted report consistently, it was essential to devise a coding structure, against which key judgements and evaluative statements could be attributed. At the design and initial stage of analysis, 189 codes were identified and used in order to capture all measurable judgements and commentary found within each individual inspection report relating to student success rates. The design of the coding structure used the criteria set out in Table 3.5. The full list of codes can be found at Appendix 1:

Criterion	Description
Age	Whether the report made reference to students aged 16-18 or 19+. If the report referred to overall success, then the appropriate codes for both 16-18 and 19+ were used
Performance Level	Three categories were identified as being used in the Ofsted inspection reports, which were used to determine the level of performance in the form of success rates and their performance. They showed whether success rates and students' performance were: <ul style="list-style-type: none"> • High, above national average, successful or if they achieved well; • In-line with, slightly or just above/below the national average; and, • Below the national average, too low (as in the success rates are...), too few (as in too few students complete...), do not complete (as in too many students...)
Proportion	Three categories were identified as being used to describe the proportion of students, to which a judgement has been ascribed. They are: <ul style="list-style-type: none"> • All, the majority or most students • The minority or few students • Some or equal
Level or qualifications	In some Ofsted reports, especially where there is variation in the performance of students, the judgements made or where a contribution to a judgement is offered, students can be described by level or qualification. The descriptions are as follows: <ul style="list-style-type: none"> • Foundation • Intermediate • Advanced • Apprenticeships • AS/A Levels • All courses (used in summary statements or where there is no detailed commentary) • Functional Skills • GCSE English and mathematics
Provision	There are two broad categories to which all of the above applies, both of which underpin them: <ul style="list-style-type: none"> • Classroom based learning (by 16-18, 19+ or all ages) • Work-based Learning, apprenticeships (all ages)

Tables 3.5 The coding framework for the analysis of inspection reports regarding the success rates for students

To illustrate how these codes have been used to analyse the text within an inspection report, the table below (Table 3.6) offers five examples. These examples use the text taken directly from an inspection report within the scope of my study and the code(s) which have been attributed to it for analysis.

Example	Text from an Ofsted inspection report	Codes
1	Too few learners aged 16-18 successfully complete their qualifications, in particular on study programmes at levels 1 & 2. Whilst most AS- and A-level learners stay to the end of their course, pass rates have dropped and for AS-level subjects they are very low.	60 (16-18, too few complete, the majority (of age group), all courses) 55 (16-18, too few complete, the majority (of age group), Foundation) 56 (16-18, too few complete, the majority (of age group), Intermediate) 59 (16-18, low, the majority (of the cohort) A-Levels)
2	Success rates for students aged 16-18 have been low for the previous two years, but in 2012/13 increased markedly and are now above the national rate.	6 (16-18, above national average, all, all courses)
3	Success rates for apprenticeships are inadequate. Numbers of apprentices completing their qualification in the timescale planned for them when they started their training have declined over the past two years and are very low	183 (WBL, low, all, all (apprenticeship) courses)
4	Learners aged 16-18 perform well. Success rates for learners aged over 19 on advanced-level courses have risen and are now well above national averages. However, their performance at foundation and intermediate level, though much improved, has only recently reached national averages.	6 (16-18, above national average, all, all courses) 84 (19+, above national average, all (on advanced courses) advanced) 109 (19+, in-line with the national average, all (on foundation courses), foundation) 110 (19+, in-line with the national average, all (on intermediate courses), foundation)
5	The development of most learners' English and mathematical skills during lesson and through assessment is poor. The proportion of learners achieving their functional skills qualifications decreased 'significantly' last year and is extremely low.	61 (16-18, low, all, functional skills) 142 (19+, low, all functional skills)

Table 3.6 Examples of how inspection report text regarding student success rates have been allocated a code for analysis

Having analysed all 165 inspection reports in the scope of my study, there are cases where judgements and/or evaluative statements are missing and therefore cannot to be coded. Where this is the case, the main reasons are that the commentary is being descriptive, or that improvement or decline has been cited, but no starting or end position has been stated.

Table 3.7 gives three examples of text taken from inspection reports used in my study, where an analysis code cannot be allocated and the reason(s) why.

Example	Text taken from an Ofsted inspection report	Commentary
1	The development of learners' English and mathematical skills is not consistently good across the college. Much work has been done to give a high profile to improving learners' literacy and numeracy, and success rates in functional skills qualifications have improved.	No code has been allocated as it is unclear what the current success rate is or where it sits in relation to the national average.
2	Outcomes for learners require improvement. They have improved since the last inspection when they were inadequate. College data show that a higher proportion of learners remain on programme than at the same time last year, although this varies across subject area.	No code has been allocated as there is no evidence offered as to the rate and impact of improvement or success in relation to the national average. It does give an indication that student retention is better, but no assumption has been made by the inspector that there is a related relationship to an improved success rate.
3	The proportion of apprentices who achieve their qualification has improved over the last three years, although too many take too long in doing so.	No code has been allocated as it is unclear as to the starting point of the three-year trend of improvement and/or the current level of performance.

Table 3.7 Examples where a textual analysis code cannot attributed with regard to student success rates

Once all of the 165 inspection reports used in my study had been analysed, and the data input into SPSS, it became clear that a number of evaluative statements and judgements regarding student success, and their corresponding codes, were used in only a minority of the reports. These were not subject to further analysis or used within my study, because:

- Very few inspection reports reported on the proportion of students as part of their judgements (removing 126 codes from the analysis);
- Where a secondary evaluative statement on student success was recorded, it was often preceded by a judgement based on overall performance. These second tier or more detailed analyses are not a consistent feature of the inspection reports in my study (removing 36 codes from the analysis); and,
- There is quite consistent reporting on A levels, functional skills and GCSE mathematics and English. However, whilst these were intended to be in scope for my

study, they were removed and are suggested areas for future research as identified in Chapter 6 (removing 18 codes from the analysis).

Therefore, the analysis of success rates which form part of my study uses the most regularly and consistently found evaluative statements and judgements made by inspectors, as identified in the inspection reports. The judgements and their associated codes are identified in Tables 3.8, for students aged 16-18 & 19+ and in Table 3.9 for apprentices:

Code	Age	Judgement	Proportion	Level
6	16-18	That students aged 16-18 achieved well, that their success rates were above the national average or that they were high	Most/The Majority	All Courses
33		That students aged 16-18 achieved in line or broadly in line with the national average, or that their success rates were slightly above or slightly below		
60		That success rates for students aged 16-18 were poor, below or 'significantly' below the national average or that too many students did not complete their course		
87	19+	That students aged 19+ achieved well, that their success rates were above the national average or that they were high	Most/The Majority	All Courses
114		That students aged 19+ achieved in line or broadly in line with the national average, or that their success rates were slightly above or slightly below		
141		That success rates for students aged 19+ were poor, below or 'significantly' below the national average or that too many students did not complete their course.		

Table 3.8 The codes used in the analysis of success rates judgements for students aged 16-18 and 19+

Code	Provision	Judgement	Proportion	Level
165	WBL Apps	Apprentices/students on an apprenticeship programme achieved well, that their success rates were above the national average or that they were high	Most/The Majority	All
174		Apprentices/students on an apprenticeship programme achieved in line or broadly in line with the national average, or that their success rates were slightly above or slightly below		
183		That success rates apprentices/students on an apprenticeship programme were poor, below or 'significantly' below the national average or that too many students did not complete their course.		

Table 3.9 The codes used in the analysis of success rates judgements for apprenticeships

3.7 Success rates

3.7.1 What are success rates

Success rates (now called achievement rates) are, as their name suggests, the instrument which identifies how successful a student is in obtaining the qualification they were enrolled onto. There are two factors used in the calculation of a success rate. They are:

- Retention: did the student stay for the duration of their programme or did they 'drop out'; and,
- Achievement: did the student achieve their qualification

The SFA (2014, p10) states that 'success rates are calculated as the number of learning aims achieved divided by the number of starts.' They can also be calculated by multiplying the achievement rate by the retention rate. This calculation is illustrated in Table 3.10.

Number of people starting a course	Number of people retained	% of those retained	Number of people achieving their qualification	% of those achieving (from those retained)	Success rate % (Ret x Ach)
10	8	80	8	100	80

Table 3.10 An illustration of how success rates are calculated, as per SFA guidance.

Over the lifetime of the Handbook (Ofsted, 2012), the principal method for calculating success rates has remained the same. This basic formula, outlined above and which has been in place since 2004/05, is still used. The presentation of success rates, within a college's QSR and across the sector via the National Achievement Rate Tables, is considerable and their analysis is multi-layered.

The presentation of success rates in any one college includes and is broken down by:

- Age;
- Level of qualification;
- Length of qualification;
- Qualification type;
- Gender;
- Learning difficulty or disability; and,
- Ethnic background.

For the purposes of self-assessment and self-improvement, this level of data can be useful in identifying specific areas of the college which are under-performing. However, whilst this validated data is available to inspectors, and might be explored by them during an inspection, the summary reports which have been analysed indicate that only a few categories are systematically and consistently reported on. They are:

- Age;
- Length of qualification; and
- Level of qualification.

Age, for classroom based learners, is separated into those aged 16-18 and those aged 19+. For those on an apprenticeship, there are three age categories: 16-18, 19-23 and 24+. However, when being reported on in an inspection report, it is more common to refer to 'apprentices' generally, without specifying age.

The length of qualification refers to the time period over which a qualification is expected to be delivered. There are three classifications identified below along with their technical description:

- Long: The expected length of a long qualification is 168 days (24 weeks) or more;
- Short: The expected length of a short qualification is more than 34 days but less than 168 days (between 5 and 24 weeks); or,
- Very Short: The expected length of a very short qualification is 34 days or less (less than five weeks).

Qualification levels are determined by a process governed by the Office for Qualification Regulations, before being approved and recorded on the Skills Funding Agency's learning aims reference application (LARA). Each qualification has a level classification. Table 3.11 identifies the five qualification levels, as they were used in the reporting of success rates across the timeline of my study, and a brief description of what is included is shown for each level.

Qualification Level	Description
1	Level 1 and entry level qualifications are pre-GCSE foundation and pre-foundation programmes. This category will also include programmes for students with learning difficulties and/or disabilities.
2	Level 2 are intermediate qualifications which would include GCSEs, BTEC First/BTEC Level 2 Diplomas and NVQ2. In some qualification areas, this level would be required for a licence to practise.
3	Advanced level qualifications. Examples of level 3 qualifications would include A levels, BTEC Extended Diplomas, Access to Higher Education and Foundation Art programmes.
H	Higher levels include QCF levels 4 and 5 such as HNCs and HNDs.
U	Level unknown, mixed level or not applicable

Table 3.11 Qualification levels and descriptions.

Qualification type is also reported on by inspectors, but not consistently. Where this is the case, reference is usually made to qualifications such as AS & A levels, Functional Skills and GCSE mathematics and English for students aged 16-18 and Access to Higher Education for students aged 19+. The success rates for each of these qualifications are cited within the QSR report.

The success rate formula was revised in 2013/14 with the aim of bringing together and therefore consistently recording and reporting classroom based learning and apprenticeships. In the main, this included an additional measure of timeliness for classroom based learners, recording the number of students who completed and achieved their qualification by the planned end date. In addition, although the formula for the calculation of success rates as outlined in Table 3.10 remained the same (except for title changes i.e. success rates became achievement rates etc), there were considerable changes to the structure of reporting, affecting qualification length. Table 3.12, taken from the SFA (2014, p3) explains the different methods of calculating, and commentary on how this has changed as part of the new methodology.

QSR method	Summary description as was	New QSR methodology
Overall	<p>The current methodology for employer-based provision (including Apprenticeship) with the results published in the Statistical First Release (SFR), FE Choices and the National Success Rate Tables (NSRT)</p> <p>Success is simply the achievement of the qualification aim, based on the later of the planned or actual leaving date (hybrid date).</p>	Remains the same
Timely	<p>Currently used as a secondary and supporting method for employer-based provision only (published in the NSRT for the first time in 2014)</p> <p>Success is counted if the qualification achievement occurs on, before, or within 90 days after the planned end date.</p>	Remains the same
Classroom-Based provision	<p>Currently used only for the classroom-based provision. A method that counts success as achievement of a qualification at any time within the same academic year as the planned end date.</p> <p>The results are published as per the overall methodology.</p>	Now includes a timely measure, used as a secondary and supporting method. Success is counted if the qualification achievement occurs on, before or within 90 days after the planned end date.

Table 3.12 Changes to the success rate methodology 2013/14

The changes to duration in the new success rate methodology are more profound. The new methodology no longer includes the timeframe over which a qualification is delivered, but instead is reported against an agreed group of qualifications, totalling 13 (SFA 2015, p4).

As a result of these changes affecting the success rate data towards the back end of my study, it has not been possible to analyse how inspectors used data and information when evaluating student performance and making judgements for students aged 19+. This is due to the major difference in the presentation and reporting of data between the old and the new methodologies with regards to duration. This only applies to the inspections that took place between January and July 2015. However, whilst the changes in success rate methodology also affect the data for students aged 16-18, they have less overall impact as the overwhelming majority of students in this age group study full-time, regardless of

qualification type. Therefore, long level success rates under the old methodology have been used for the purposes of my study to equate to all qualification types under the new methodology. Again, this only affects the inspections that took place between January and July 2015 (n=18 inspections).

3.7.2 Where has the success rate data come from

Success rate data for all colleges, sixth form colleges and training providers is published by the government and can be accessed at <https://www.gov.uk/government/collections/sfa-national-success-rates-tables> Validated success rate data, along with guidance published by the SFA, is available on this web-site covering the academic years 2012/13 to 2015/16. It is this data, housed within these datasets, which is used by inspectors when inspecting colleges, although each college will have an individual report, specific to its organisation (rather than having to extrapolate the data from the national dataset).

For the colleges which were inspected during the timeframe of my study, their success rates were extrapolated from the national reports, not only for the year or year prior to inspection but for all the academic years 2010/11 to 2014/15, with the exception of adult learners where information is up to only 2012/13. This data has been transposed from the national success rate table to SPSS at individual college level.

3.7.3 How are the success rates attributed to an inspection

Ofsted (2012a, p7) states that only validated success rate data should be used to evaluate student performance. It also states that this data is provided via the QSR and the NSRT, both produced by the data service. As part of the procedure of validation, a college is sent its data prior to it being published, to check for any inaccuracies or anomalies. Each year, for the academic years covered in my study, the national success rate datasets (QSRs and NART) have been published in January for the previous year's cohort. For the purpose of my study, because of either incomplete data or a lack of validated data being available to inspectors for inspections that took place between September and December, the last available published QSR has been used. And it is on this data that student success has been evaluated and judged. Table 3.13 identifies the relationship among the QSR year (latest available validated data), the Year of Inspection (when the data was used in an inspection) and the Academic Year in which the inspection took place.

QSR Year (validated success rate)	10/11	11/12		12/13		13/14
Year of Inspection	2012 September – December	2013 January – July	2013 September – December	2014 January – July	2014 September – December	2015 January – July
Academic Year	12/13		13/14		14/15	

Table 3.13 The relationship among QSR report and both the calendar and academic year of inspection

So, as an example, for an inspection that took place in November 2013, the success rate data housed within the QSR report from 11/12 has been used by inspectors. This is because it is the latest validated success rate data available.

Based on the results of the textual analysis, the most common and consistent categories against which student success is evaluated and judged are:

- Long level 16-18;
- Long level 19+; and,
- Overall apprenticeships.

As a result, the success rates for the categories listed above, for the inspections used in my study, covering five academic years (with the exception of those aged 19+ where 3 years have been recorded), have been recorded in SPSS.

3.8 The national average

National averages are calculated 'using the totality of provision offered in England' (Ofsted 2012a, p5) using the same formula as outlined in Table 3.10. Therefore, a national average can exist when analysing success rate data using all or a combination of the methodological features shown above, such as duration, age and qualification.

Ofsted (ibid.) state that, where appropriate, inspectors should use the 'all' national average. This average brings together the mean average success rate from a large student population. Having analysed the inspection reports and, as identified earlier in this chapter,

using the textual analysis of all inspection reports used in my study, the most common national averages used are:

- Long level 16-18;
- Long level 19+; and,
- Overall apprenticeship.

As a result, these national averages have been transposed from the NSRT for each of the corresponding groups and across each inspection year into SPSS.

However, it should be noted that in a small number of inspection reports, reference is made to other national averages. Where this is the case, and in some it is unclear, inspectors refer to the national average for colleges within an area of high social deprivation. In addition, success rates for students on specialist provision within a college of further education, such as A levels or land-based, are compared to the national averages for students in designated sixth form or agricultural colleges.

3.9 Accessing Non-publicly Available Data

Of the data used by inspectors to evaluate and make judgements on the various aspects of outcomes for learners, it is only when success rates are concerned that the dataset is comprehensive and publicly available. In addition, this information, and value added data (although the latter applies to only a small cohort of further education college students) are the only two sets of data which are externally validated.

However, Ofsted (2012a, p8), state that inspectors will also use a provider's own data when seeking evidence on:

- Attendance;
- Value Added
- Progression; and,
- Destination.

Therefore, a freedom of information request (FoI) was sent to all 165 colleges, requesting information on the above aspects of their inspection (see appendix 2 and appendix 3). The

FoI requests were sent out to colleges over three days, 3, 4 & 5 October 2016. The number of colleges who responded to the FoI was 83, or just over 50%, although the response rate for each of the questions asked varied. The exact number of colleges who responded to the individual questions is cited in the detail below.

3.10 Attendance

There is no national guidance on how a college should collect and analyse attendance data and it is therefore unclear what attendance data is used by inspectors to evaluate the standard of attendance in one college when compared to any other.

As part of the freedom of information request, colleges were asked to provide their attendance data from different points in the reporting cycle and for different provision.

Colleges were asked:

- What was the College's overall reported attendance (cumulative) at the point of inspection? (A1)
- What was the College's overall reported attendance (cumulative) at the year end prior to inspection? (A2)
- What was the College's reported attendance (cumulative) at the point of inspection for English and mathematics (Functional Skills and GCSE)? (A3)
- What was the College's overall reported attendance (cumulative) at the year end prior to inspection for English and mathematics (Functional Skills and GCSE)? (A5)

The greatest response rates from colleges were against the first two questions (A1 & A2). The attendance data was received from 65 colleges (39.4% of the total sample) covering inspections regarding attendance at the point of inspection (A1). The attendance data was received from 67 colleges (40.6% of the total sample) covering inspections regarding attendance at the year end prior to inspection.

Attendance, within the inspection reports analysed for my study, indicates that in-year attendance data, provided by the college, is used by inspectors as part of their evidence base to evaluate a college's standard. There is also evidence to suggest that inspectors measure attendance from direct observation, allowing them to compare what they see against what the college says. For example, of the 58 (35% of the total sample) responses

to the question 'Did the inspection team share with the Principal or Nominee that there was a difference between the reported attendance (used by the college at the point of inspection) and the reported attendance by the inspection team?' (A7) only a very small number (n=8) said yes. This suggests, therefore, that the in-year attendance data provided by colleges for this study (A1), which was also presented to the inspection team for evaluation, is both valid and reliable. Thus, it is this data only which has been used to analyse the consistency with which inspectors evaluate and make judgements about attendance.

The responses to the questions regarding English and mathematics are less robust. Although there were 58 responses regarding the reported attendance at the point of inspection for English and mathematics (A3) the evaluative statements and judgements within the reports were differentiated by qualification i.e. functional skills and GCSE rather than as a combined English and mathematics. Therefore, the data could not be attributed with any accuracy to an inspector's judgements.

3.11 How are the reports analysed for attendance

All inspection reports within the scope of my study have been analysed and those where inspectors have made an evaluative statement and/or judgement on attendance have been recorded. Of the 165 inspections in my study, 111 (67.3%) reports made an evaluative statement on attendance, which have been categorised into three judgement areas as outlined in Table 3.14, along with their SPSS coding.

SPSS Code	Description	Example from an Ofsted report
1	Attendance at the college was high/good or students attended well	Learners' attendance and punctuality are good
2	Attendance at the college required improvement/was average or satisfactory	Attendance rates have improved over recent years and are satisfactory
3	Attendance at the college was low/poor/variable/inconsistent or below average	The overall attendance was too low

Table 3.14 Description of attendance judgements found in inspection reports

3.12 Value Added

The measurement of value added and student progress is complex. In colleges of further education, there is only one value added system which is validated externally to the college – although this one system is facilitated by a number of private companies and the Department for Education. However, this system only caters for a small population, less than 7% of the total further education college cohort, namely for students aged 16-18 on level 3 programmes. For the remaining students, it is expected by Ofsted (2012a, p11) that each college will have its own system which measures student progress in relationship to their starting points. It is therefore not possible to analyse with what level of consistency inspectors use data and information to evaluate the progress made by students, as the mechanism by which systems and data captured is individual to each college.

Even where the data has been collected centrally, for the recognised value added system used by Ofsted – the DfE Level 3 Value Added table (L3VA) - it was not publicly available at individual college level. Therefore, no analysis has been provided on how consistently inspectors used this data to evaluate the progress made by students.

However, what was asked of colleges were two questions (V1 & V2), with two supplementary questions (V1Y & V2Y). They were:

- At the point of inspection, did the College formally use and report on value added as a KPI? (V1)
- If the answer is yes to V1, what value added tool is used at the College i.e. ALPS, LAT, college devised? Please specify (V1Y)
- At inspection, did the inspection team explore value added data and information? (V2)
- If the answer is yes to V2, how much weight was placed on it as a tool for measuring student outcomes? (V2Y)

The number of inspections covered by the responses to the first question (V1) was 69 (41.8% of the total cohort) of which 48 stated that they did use and formally report on value added, whilst 21 stated that they did not. Of those 48 inspections of colleges who did use a value added system at the point of their inspection (V1Y), the tools they used included the

LAT (the predecessor to the L3VA), ALPS, ALIS, College devised or a combination. (A detailed presentation of the systems used by colleges is in Table 4.14).

The number of colleges who responded, indicating if inspectors explored value added as part of their inspection, covered 66 inspections (V2). Of these responses, the majority (n=55) stated that value added was explored whilst 11 stated that it was not. When asked about the weight that they (the respondent(s) from the college) believed was placed on value added as a tool for measuring outcomes (V2Y), the majority (n=38), reported that it was hardly any. The responses from the remaining inspections (n=17), stated that a 'significant' weight had been placed on value added by inspections.

3.13 How are the reports analysed for value added and the progress made by students

All inspection reports within the scope of my study have been analysed and those where inspectors have made an evaluative statement and/or judgement on value added and the progress made by students have been recorded. Of the 165 inspections in my study, 99 (60%) reports made an evaluative statement on value added, which have been categorised into three judgement areas as outlined in Table 3.15, along with their SPSS coding.

SPSS Code	Description	Example from an Ofsted report
1	The majority of/a high proportion of students make better than expected progress or a 'significant' majority of students exceed the progress expected of them	The large majority of students make good progress relative to the qualifications they achieved before coming to college.
2	The majority of students make expected progress or their progress is in line with what was expected of them	Learners on advanced level courses make expected progress based on their prior attainment.
3	Students do not make the expected progress or their progress is too slow. Too many students fail to make adequate progress	...too many students do not achieve in line with expectations given their prior qualifications on entry to the college.

Table 3.15 Description of value added judgements found in inspection reports

3.14 Progression Data

Progression is an instrument defined by Ofsted (2012a, p3) which measures the proportion of students who progress from one level of course to the next with the same provider. Whilst that explanation seems straightforward, there was no systematic collection or reporting of progression data beyond its use within individual colleges. Therefore, in order to analyse the judgements made by inspectors and their use of data to underpin such judgements, the data presented at inspections with regard to student progression was sought.

There are two aspects which have been considered when requesting the data from colleges: the overall percentage rate of progression, and the type of data that the inspectors used. Whilst a number of questions in the FoI request looked for detail regarding the progression rate by age or whether there were gaps in any particular curriculum or subject areas (P3, P4 & P5), the inspection reports did not, in any great number, make reference to such detailed analyses. As a result, the data provided by respondents to these questions has not been used in my study. In the main, progression is evaluated and reported by inspectors for the entire student body; at least there is no evidence to suggest otherwise.

In addition, it was also important to identify the type of data presented to inspectors. Colleges have the ability to collect and use either intended progression data (a projection of likely progression) or actual progression data. The difference can be quite stark with regard to the proportion of known outcomes and to its reliability. Therefore, to inform with what consistency progression data was used by inspectors, the responses to the following two questions were used.

- At inspection, what type of progression information was used? Intended progression **(I)** (normally collected by the course tutor or teacher prior to the student leaving the college in the summer) or Actual progression **(A)** (provided by MIS/central data taken directly from the ILR)? (P1)
- At inspection, what figure was presented to the inspection team that indicated positive progression? (P2)

Of the number of colleges who responded to the FoI request, which covered 67 inspections (40.6% of the total sample), 47 stated that the progression data presented to the inspection team for evaluation was actual data, whilst the respondents for the remaining 20 inspections

stated that intended data was used (P1). In addition, the number of respondents who provided an overall figure of positive progression was 59 (35.8% of the total sample).

3.15 How are the reports used to analyse progression

All inspection reports within the scope of my study have been analysed and those where inspectors have made an evaluative statement and/or judgement on attendance have been recorded. Of the 165 inspections in my study, 144 (87.3%) reports made an evaluative statement on progression, which have been categorised into three judgement areas as outlined in Table 3.16, along with their SPSS coding.

SPSS Code	Description	Example for an Ofsted report
1	A high proportion of, or the 'significant' majority of, students progress through the levels to higher study or into work	Learners make good progress from lower-level courses to higher.
2	A good proportion of, or the majority of, students progress through the levels to higher study or into work. The collection of progress data needs to be improved	Most students progress to higher level courses at the college. The collection of progression data requires improvement.
3	The progression of students is too low or not enough are recorded as having positive destinations. The data available is incomplete and too many students' destinations are unknown	Students' progression to higher levels of study within the college is not high enough.

Table 3.16 Description of progression judgements found in inspection reports

3.16 Destination Data

Destination is an instrument defined by Ofsted (2012a, p3) which measures the proportion of students who progress from one level of course to the next with the same provider. Whilst that explanation seems straightforward, there was no systematic collection or reporting of destination data beyond its use within individual colleges. Therefore, in order to analyse the judgements made by inspectors and their use of data to underpin such judgements, the data presented at inspections with regards to student destination was sought.

There are two aspects which have been considered when requesting the data from colleges. The overall percentage rate of destination and the type of data that the inspectors used. Whilst a number of questions in the FoI request looked for detail, regarding the destination

rate by age or whether there were gaps in any particular curriculum or subject areas (D3, D4 & D5), the inspection reports did not, in any great number, make reference to such detailed analyses. As a result, the data provided by respondents to these questions has not been used in my study. In the main, destination is evaluated and reported by inspectors for the entire student body; at least there is no evidence to suggest otherwise.

In addition, it was also important to identify the type of data presented to inspectors. Colleges have the ability to collect and use either intended destination data (a projection of likely destination) or actual destination data. The difference can be quite stark with regard to the proportion of known outcomes and to its reliability. Therefore, to inform with what consistency destination data was used by inspectors, the responses to the following two questions were used.

- At inspection, what type of destination information was used? Intended Destination **(I)** (normally collected by the course tutor or teacher prior to the student leaving the college in the summer) or Actual Destination **(A)** (provided by MIS/central data taken directly from the ILR)? (D1)
- At inspection, what figure was presented to the inspection team that indicated positive destination? (D2)

Of the number of colleges who responded to the FoI request, which covered 63 inspections (38.2% of the total sample), 39 stated that the destination data presented to the inspection team for evaluation was actual data, whilst the respondents for the remaining 24 inspections stated that intended data was used (D1). In addition, the number of respondents who provided an overall figure of positive destination was 56 (33.9% of the total sample) (D2).

3.17 How are the reports used to analyse destination

All inspection reports within the scope of my study have been analysed and those where inspectors have made an evaluative statement and/or judgement on attendance have been recorded. Of the 165 inspections in this study, 139 (84.3%) reports made an evaluative statement on destination, which have been categorised into three judgement areas as outlined in Table 3.17, along with their SPSS coding.

SPSS Code	Description	Example for an Ofsted report
1	A high proportion of, or the 'significant' majority of students' destinations to higher levels of study or into employment is shown	Most college leavers gain employment in jobs that relate closely to their studies, while the remainder move on to higher education, often to prestigious universities.
2	A good proportion of, or the majority of students' destinations to higher study or in to work is recorded. The collection of destination data needs to be improved	Students' destinations to higher levels of learning and into employment require improvement.
3	The positive destinations of students are too low or not enough are recorded as having positive destinations. The data available is incomplete and too many students' destinations are unknown	Data on the destinations of learners outside the college is not yet complete and destinations are unknown for more than a quarter of students.

Table 3.17 Description of destination judgements found in inspection reports

3.18 Limitations of the methods used in this research

The use of secondary data in doctoral research according to Goes & Simon (nd, p1) has some considerable advantages with regard to its availability, accessibility and consistency. Thus, the data on which the analysis, discussion and conclusions are made, is more likely to be part of a complete and larger dataset. Clark and Cossette (2001, p113), however, conclude that, although there are considerable benefits to the use of collecting secondary data for the use in primary research, such as those outlined above, there are some major methodological issues in secondary analysis. These are:

- Sampling;
- Measurement; and
- External validity.

Therefore, the remainder of this chapter uses these methodological issues as its structure, not only to identify the limitations of the research methodology, but to identify the success with which they have been mitigated.

3.19 Sampling

Selection biases: there are two aspects of my study which might be impacted by the accusation of selection bias, namely the timeframe in which the inspections analysed took place, and the number of inspections.

The choice of timeframe for the inspections analysed in my study was based on two facts. The first was that this was the first cohort of inspections in further education colleges carried out exclusively by Ofsted and the second was that it covered one complete inspection cycle. The second cohort of college inspections covering a whole inspection cycle is yet to complete.

Representativeness of original sample: as a study of inspections over time, it was essential that the sample size was sufficiently large to ensure that the findings were representative. The original sample, as defined by the search function of the Ofsted web-site, provided 185 inspections. For the reasons outlined earlier in this chapter, 20 inspections were withdrawn from the study. Therefore, the remaining 165 inspections analysed in my study represent 89.2% of the original sample, or if inspections of colleges designated as specialist are not included ($n=9$), my study represents 95.3% of the original sample.

Sufficient variability of key concepts: there is the possibility that the success rate data used by inspectors to evaluate student performance is not that identified in my study. Inspectors may have used a college's own data, especially if the inspection took place before the publication of the QSR (September - December), or if a college offered its un-validated QSR (which the college might have had for checking purposes) which it shared with inspectors.

To avoid assumptions on what success rate data inspectors used or did not use, and to ensure that the analysis was consistent, I applied Ofsted's guidance, which states that it uses only validated success rate data, provided by government agencies (2012a, p7&8). The publication dates for such reports are publicly available.

Reductions in analytical sample due to missing data: as outlined earlier in this chapter, there are three sources of data, two statistical and one textual, that are the foundation of my study. Of the original 185 inspections identified at the outset, a number ($n=11$) were discounted based on an incomplete history with regards to success rates. For the 165

inspections analysed, there was a complete set of data for 16-18 success rates. For those aged 19+, due to a methodology change, 13.3% (n=22) of inspections could not be analysed. Similarly, due to unavailable success rate data, 6.7% (n=11) of inspections could not be analysed for WBL apprenticeships.

Whilst the number of colleges who responded to the FoI request covers 50.3% (n=83) of inspections, the variation in responses to the individual questions ranges from 35% to 41.8%.

Selection of sub-groups from a larger primary data set: no analysis on any sub-grouping has been included in this study.

3.20 Measurement

Shortcomings of original researchers' instruments: the secondary data collected for this study, both success rate data and the data obtained directly from colleges, uses the same instruments and measurements as originally used by inspectors. The calculation which identifies a student success rate, cited in Table 3.10, has remained constant throughout the timeline of my study.

Conceptual slippage in proxy measures: no additional variables have been added to the secondary data used in my study. Whilst consideration and a substantive amount of analysis was undertaken to recreate comparable success rates for students aged 19+ following the change to the success rate methodology, this data was in fact not used. This was because the relationships between that success rate and the judgements made by inspectors would have been speculative at best and subject to challenge on the grounds of reliability and validity.

3.21 External Validity

Original conditions or measurement may be at odds with conditions of interest: the original conditions, as far as they can be assessed, have been applied to my study. Success rate data (and its constituent parts) is created by the college throughout the year (as working unvalidated data) and once collected and validated, following the end of an academic year, it is produced formally and put into the public domain every year. The publication of this data is

not dependent on a college being inspected. The success rate data in this study has been used in the same way.

The data regarding the remaining aspects of outcomes for learners was collected from colleges under the premise that it was the same data that they provided to inspectors as part of their inspection. Therefore, it too has been used, as far as can be assessed, under the same conditions as pertained during the original inspection.

Time elapsed since original data collection: my study assesses with what consistency inspectors use data and information to evaluate student performance and to make judgements on outcomes for learners. As explained above, the timescale for my study is predicated on the analysis of a complete inspection cycle under one inspection body. Whilst these inspections are now between two and five years old, this is the only set of further education college inspections that can be studied.

3.22 Conclusion

Much of the literature reviewed in Chapter 2 represents a body of research that looks at the relationship between inspection and its impact on individuals, organisations and the education systems in England and across Europe. However, what is now apparent is that in many of the studies cited in Chapter 2 they are limited in one or more factors such as research design, method or the response rates from participants. This has resulted in many instances where quite generalist conclusions have been made and which purport to be representative of the wider education system or to their specific sector, for example the secondary school sector. Whilst the authors of these studies might be correct in their assertions as to the impact of inspection on individuals, organisations and in the wider education system, they are arguably not based on (enough) evidence.

My study, by comparison to most of those in Chapter 2 is large (possibly one of the largest studies of its kind in this field) – in that it analyses all of the inspections that took place in further education colleges between September 2012 and July 2015 – one entire inspection cycle. Moreover, the fact that it considers how consistently inspectors use data and information to inform judgements in outcomes for learners, and to do so using secondary data for primary research, means that the findings have been arrived at scientifically. This method, having been used to test the hypothesis at the beginning of this chapter and

subsequently used for the collection and analysis of both quantifiable statistical and textual data, means that it is reliable and can be tested by others. Or, as outlined in Chapter 6, it can be used to re-run the same analysis and answer the same questions when the current inspection cycle of further education colleges is completed.

Finally, as a result of the scientific research and analysis approach taken, it is hoped that my study will lead the way to similar studies, not only in the further education sector but in both the primary and secondary sectors too. Studies regarding inspection and its impact that can be tested and developed by others on a similar scale; studies like this one that are values neutral; and, studies like this one where the findings, as cited in Chapter 4 and discussed further in Chapter 5, are based solidly on evidence and which are truly representative.

Chapter 4: Data Analysis

4.0 Introduction

If transparency and consistency are two of the principles of inspection (Ofsted 2012, p6), then the data and information analysed through this chapter seek to confirm whether this is indeed the case where inspectors have evaluated and judged outcomes for learners. To ensure that the analysis is comprehensive, each of the individual components which make up the theme on outcomes (Ofsted 2012, p43 & Ofsted 2012a, pp8-9) has been individually analysed in the sections below, namely:

- Success Rates (for students aged 16-18, 19+, and WBL (apprenticeships);
- Value Added;
- Attendance;
- Progression; and,
- Destination.

Each section against which the data is analysed is consistent through this chapter in that it takes the judgement/evaluative statement made by inspectors directly from the inspection report as its starting point. It is against these judgements that the data (which is almost never cited in an inspection report) has been applied, with the aim of identifying how consistently inspectors use data and information to substantiate the judgements made. The data used has been drawn from two sources in line with the guidance set out by Ofsted (2012a), namely the Department for Education (for success rate data) and from the colleges themselves (for the remaining aspects).

It should be noted that one component housed within the outcomes theme has not been considered in this thesis. The criterion 'learners develop personal, social and employability skills' (Ofsted 2012, p43) requires the inspection team to consider a range of skills and knowledge most of which do not lead to a qualification or any form of formal assessment. As a result, the evidence for the judgements made by inspectors against this criterion is likely to be obtained from direct observation, meetings with students, staff, and others, and through the review of planned activities, and therefore it offers no opportunity for critical analysis as part of this research project.

In the section analysing the inspection outcomes and success rates, there are two inspection cycles where there is missing data. At the time of collecting the success rate data for this thesis, the success rate data for WBL (apprenticeships) was not publicly available as part of the 10/11 QSR. This would be the apprenticeship data used by inspectors in the inspections that took place between September and December 2012. In addition, the methodology for calculating success rates for students aged 19+ was changed by the Skills Funding Agency (SFA, 2014). This change was fundamental, and thus it means that there is no comparable data for those inspections that took place between January and July 2015. However, the number of inspections that have had to be eliminated from this research is small and, therefore it does not compromise the analysis and findings for either cohort.

4.1 The Sample

The timeframe used for this analysis is from September 2012 to July 2015, which represents a complete inspection cycle under one handbook and common inspection framework.

The number of inspections that took place across this period when using the provider type selection criteria on the Ofsted web-site was 185. Twenty individual inspections have been discounted on the basis that:

- The institution was designated a special college such as a land-based college;
- The institution was an adult education institution with a very small number of students aged 16-18, providing mostly part-time courses; or,
- Due to merger there is no longer historical/comparable data. These inspections have been removed from the cohort leaving 165 valid inspections of which the reports and datasets have been used in the analysis below.

Table 4.1 identifies the number of inspections which have been used in the analysis of this chapter, confirming the sample of 165, and subsequently distributed by:

- The academic year in which the inspection took place;
- The number of inspections that fell into each year; and,
- The number of inspections by year, distributed by the grade awarded for outcomes for learners

Academic Year	Number in the Sample	OfL Grade 1 ¹	OfL Grade 2 ²	OfL Grade 3 ³	OfL Grade 4 ⁴
12/13	62	2	25	28	7
13/14	62	5	29	28	0
14/15	41	2	10	26	3
Total	165	9	64	82	10

Table 4.1 Total number of inspections in scope, by academic year and inspection grade for OfL

Where information on attendance, value added, progression and destination are not held centrally and/or publicly, Table 4.2 records both the total number of colleges that responded to the freedom of information requests in absolute numbers (83) and the percentage rate (50.3%). Moreover, the same recordings are made when categorised by region (using the Ofsted's regional groupings).

However, it should be noted that the number of responses identified below does not necessarily mean that all colleges answered all of the questions. Where the analysis in this chapter concerns itself with data and information provided by colleges, the number of responses to the question asked is recorded in the accompanying table.

		Regional Area								
		South East	South West	London	Yorkshire Humber	North West	North East	East of England	West Midlands	East Midlands
Colleges in the sample	165	26	14	25	14	20	12	20	25	9
No of FoI Responses	83	18	10	11	7	8	6	9	9	5
% of regional sample	50.3	69	71	44	50	40	50	45	36	56

Table 4.2 Total number of responses received from the FoI, broken down by geographic area

4.2 Success Rates

Success rates are central to the colleges' and to the whole further education sector's measurement of success. In simple terms, they identify how many students stayed on their course and how many achieved the qualifications that they enrolled on. A combination of the retention rate and the achievement rate results in the success rate (YPLA, 2011). These success rates are at the epicentre of the 'outcomes for learners' judgement made by inspectors. Whilst guidance to both colleges and inspection teams states that success rates

¹ OfL Grade 1 = A judgement of 'outstanding' awarded for Outcomes for Learners

² OfL Grade 2 = A judgement of 'good' awarded for Outcomes for Learners

³ OfL Grade 3 = A judgement of 'requires improvement' awarded for Outcomes for Learners

⁴ OfL Grade 4 = A judgement of 'inadequate' awarded for Outcomes for Learners

'are only one measure' to be considered as part of the broader theme (Ofsted, 2012a p4), they are considered an 'important measure of performance' (ibid., p8).

Student success rates play a pivotal role in the measurement of a college's performance and they arguably play an equally pivotal role when evaluating, making a judgement and awarding a grade for outcomes for learners. Indeed, one of the main principles of all inspection reports is to provide a comprehensive picture of the standard of education and training at any particular institution – with the aim of informing the public, some of whom will be considering a range of post-16 options (in conjunction with their parents, teachers and others). In doing so, they have the possibility of looking at the latest inspection report in order to support their choice.

The three sets of tables below (one for 16-18, one for 19+, and the other for WBL) show:

- the extrapolated analysis from each of the inspection reports, coded by what the inspection team have said about the college;
- the publicly available, validated data, which would have been used by the inspection team and on which their judgements were formed;
- and, the national average.

It should be noted that all aspects of a college's data (and therefore the national dataset) have a corresponding national average, for example for different course types such as A Levels or for a particular level such as Level 2 or Level 1. However, the guidance issued to inspectors (Ofsted, 2012a) identifies clearly the national average to be used, which in this case is the 'all' national average for each of the three categories i.e. 16-18 year olds, 19+, and WBL.

The guidance goes on to state that, whilst the national average should be used as a national comparison, it does not use or apply 'threshold values' and for the purpose of inspection 'there are no benchmarks' (Ofsted 2012a, p10). However, the language used by the inspection team and included in their reports does indicate a college's performance and on many occasions makes reference to that performance against the national average. Therefore, in order to ensure consistency and transparency, reference will be made to the national average in the context of both the textual and the data analyses. Moreover, examples which have been taken directly from published Ofsted reports within the scope of

this thesis will be used to illustrate this stance. In the guidance (Ofsted 2012a, p5) and used in the success rate tables below, the national average used is the 'all' national average for each of the success rate categories. This is the national average calculated on the totality of the provision offered in England and published within the QSR report. As identified in this chapter, the national averages used are:

- The national average for students aged 16-18 on Long Qualifications (all levels)
- The national average for students aged 19+ on Long Qualifications (all levels)
- The national average for students on apprenticeships (all ages and levels), overall success

As one of the most high profile cohorts within a further education college, it is not surprising that in most Ofsted inspection reports used in this study, the 'outcomes for learners' section begins with the standard of achievement for students aged 16-18.

For the purpose of this study and to fully understand how student success rate data is being used to evaluate colleges' academic performance, make judgements, and present to the public (in an accessible and consistent manner) what it believes those standards to be, it should also be assumed that inspectors are doing so in a way which allows for comparisons to be made with other colleges, regardless of location, lead inspector, number of students and so on.

4.3 Inspection Cycles

Throughout this section, which analyses the relationship among success rates, national averages and the judgement made by inspectors, references are made back to the accompanying data tables. Where this is the case, both the QSR Year and the Academic Year are cited in order for the appropriate data:

- QSR Year is the year from which the latest validated data was available at the point of inspection; and,
- Academic Year is the year in which the inspection took place.

Therefore, as an example, where QSR 10/11 Academic 12/13 are cited, they refer to the latest available validated dataset for the inspection that took place in 12/13.

4.4 Success Rates, Textual Analysis and National Averages for Students Aged 16-18

The tables below (Tables, 4.3, 4.4 and 4.5) bring together the individual components which underpin the judgements made about the standard by which students achieve their qualifications. The components are:

- The validated success rate data used by the inspection team for students aged 16-18 (found in the available QSR report at the point of inspection) on Long Qualifications;
- The national average used by the inspection team for students aged 16-18 on Long Qualifications (found in the available QSR report at the point of inspection) and;
- The judgement made by the inspection team and published within each college's Ofsted report (these have been analysed and subsequently coded).

It should be noted at this stage that the number of colleges in this sample totals 140 (out of the total cohort of 165). This is due to there being no clear judgement on the performance of students aged 16-18 in 25 inspection reports.

4.4.1 Analysis

There are 53 colleges in this sample where student performance has been evaluated as being above the national average or high. The analysis in Table 4.3 references a college's inspection report where the performance of students aged 16-18 was above the national average and/or high. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p46) it could be argued that an associated judgement on the performance of a cohort of students where the majority perform 'well above the norm' or where the majority perform above the national rate will be outstanding or at least good.

If logic is applied to the grade characteristic, cited above, then the success rate data in Table 4.3 should be above the national average, and if it is, by a clear margin. However, what is identified below presents a picture that is far from clear. Across all years, with the exception of QSR 13/14 Academic 14/15, the differences between the lowest success rate and the highest success rate are between 7.9% and 13.7%. Moreover, with the exception of the

penultimate year, the relationship between success rates and the national average are also variable. Within this set of 53 inspections, there are 7 examples of where the success rate is below the national average. In addition, there are a further 16 examples of where the success rate is on or within a small margin of +/- 1.5% of the national average, indicating that approximately 43% of the inspections in this sample do not enjoy a success rate which clearly denotes that student performance is above the national average or indeed high.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13	4	78.9 – 86.8	7.9	79.7	-0.8 to +7.1
11/12	2013	12/13	11	77.1 – 90.3	13.2	82	-4.9 to +8.3
11/12	2013	13/14	10	77.7 – 89.1	11.4		-4.3 to +7.1
12/13	2014	13/14	20	77.8 – 88.6	10.8	83.8	-5.0 to +5.3
12/13	2014	14/15	4	84.1 – 87.4	3.3		+0.3 to +3.6
13/14	2015	14/15	4	75.8 – 89.5	13.7	78.4	-2.6 to +11.1

Table 4.3 The data range of colleges inspected against the documentary analysis (Code 6):
The success rate for 16-18 year olds is above average/high

The analysis in Table 4.4 references the sample of 46 college inspection reports where the performance of students aged 16-18 was judged to be in line with the national average, slightly above it or slightly below it. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p47), the description for student success rates that require improvement states that 'the majority of groups of learners are likely to be in line with similar groups of learners nationally'.

There is a broad pattern in the data in Table 4.4. The difference between the lowest and highest success rates is relatively consistent, with most years recording the difference at just above 6% to just above 7%, although there is an exception in QSR 13/14 Academic 14/15 where the difference is 13.9%. The relationship between success rates and the national average is also relatively consistent, in that in every year they start below the national average and in most years, do so at around -5% and -6%. At the top end of the same relationship, with the exception of QSR 10/11 Academic 12/13 and QSR 13/14 Academic 14/15 (where success rates are above the national average), success rates are in line with the national average, slightly above and slightly below.

Of the 46 inspections, 18 recorded a success rate below the national average. In particular, of the 7 inspections in QSR 12/13 Academic 13/14, 5 recorded success rates below the

national average. Of the remainder of the whole sample, the majority recorded a success rate in line with the national average and a small minority recorded success rates above it.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average) %	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13	4	76.1 – 83.5	7.4	79.7	-3.6 to +3.8
11/12	2013	12/13	17	77.2 – 84	6.8	82	-4.8 to +2.0
11/12	2013	13/14	6	76.8 – 81.3	4.5		-5.2 to -0.7
12/13	2014	13/14	7	77.9 – 84	6.1	83.8	-5.9 to +0.2
12/13	2014	14/15	6	78 – 85.1	7.1		-5.8 to +1.3
13/14	2015	14/15	6	72.4 – 86.3	13.9	78.4	-6 to +7.9

Table 4.4 The data range of colleges inspected against the documentary analysis (Code 33): The success rates for 16-18 year olds is in line with the national average or just above/below it

The analysis in Table 4.5 references 41 college inspection reports where the performance of students aged 16-18 was low or below the national average. In the grade characteristics for 'outcomes for learners (Ofsted, 2012 p47) where success rates are 'consistently low, too variable or in significant decline' they are identified within the inadequate judgement.

The data range between the highest and lowest success rates for 16-18 year olds is most extreme in this set of inspections with a range in most years of between 13.8% and 15.8%. The shape of this data range is mirrored when comparing success rates with the national average. Across all years the relationship between success rates and the national average, at the bottom end, is considerably below the national average, which is to be expected. However, within this data range there are many examples of success rates being in line or above the national average. Of the 41 inspections in this set, 18 have a success rate in line with the national average and 5 record a success rate above it.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13	1	70	-	79.7	-9.7
11/12	2013	12/13	12	71.2 – 87	15.8	82	-10.8 to +5.0
11/12	2013	13/14	0	N/A	N/A		N/A
12/13	2014	13/14	14	73.6 – 89.4	15.8	83.8	-10.2 to +5.6
12/13	2014	14/15	6	75.3 – 89.1	13.8		-8.5 to +5.3
13/14	2015	14/15	8	72.4 - 81	8.8	78.4	-6.0 to +2.6

Table 4.5 The data range of colleges inspected against the documentary analysis (Code 60): The success rates for 16-18 year olds is below the national average, too low or poor

The analysis outlined above is stark. There are some considerable data ranges, whether it is between the lowest and highest success rates in any one year or whether it is regarding the relationship between success rates and the national averages. In all judgement areas for students aged 16-18, there is some considerable overlap, across the years within the individual judgement areas and across all three judgement areas. As a result, this indicates that the use of the national average in too many cases is not considered at all or is misapplied or that another average is used. From the analysis above, of the 140 inspections within this cohort, 64 (or 45.7%) have data that is arguably out of step with the judgement made and noted within the inspection report.

4.5 Success Rates, Textual Analysis and National Averages for Students Aged 19+

Students aged 19+ (or often referred to as adult students) form a sizeable often the largest cohort within the further education college sector. Many adult students study on a part-time basis, on what is technically referred to as either a short or a very short course, the timeframe of which is less than 12 weeks or less than 5 weeks respectively. Both short and very short courses will include provision such as 1 day or 2 day courses. Where this is the case, success rates and conversely, national averages, are consistently high.

Therefore, similarly to the analysis for students aged 16-18, the data and textual analysis below (Tables 4.6, 4.7 and 4.8) concern themselves with 19+ students, who studied for a Long Qualification (more than 24 weeks), a high proportion of whom would have studied for the whole academic year.

The components used for the analysis of this cohort are:

- The validated success rate data used by the inspection team for students aged 19+ (found in the available QSR report at the point of inspection) on Long Qualifications;
- The national average used by the inspection team for students aged 19+ on Long Qualifications (found in the available QSR report at the point of inspection) and;
- The judgement made by the inspection team and published within each college's Ofsted report (these have been analysed and subsequently coded).

Of the 165 inspection reports analysed as part of this study, there were only 100 that made a specific and clear judgement on the performance of students aged 19+.

4.5.1 Analysis

The analysis in Table 4.6 references the 32 college inspection reports where the performance of students aged 19+ was above the national average and/or high. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p46) it could be argued that an associated judgement on the performance of a cohort of students where the majority perform 'well above the norm' or where the majority perform above the national rate will be outstanding or at least good.

The table below shows a mixed picture denoting the range of success rate data from the highest to the lowest. In QSR 10/11 Academic 12/13, QSR 11/12 Academic 13/14, and QSR 12/13 Academic 14/15 the range is relatively small. However, in QSR 11/12 Academic 12/13 and QSR 12/13 Academic 14/15, the range is much greater at 10.6% for both inspection cycles.

The relationship between success rates and the national average is quite disparate. In three of the five inspection cycles (QSR 10/11 Academic 12/13, QSR 11/12 Academic 13/14, and QSR 12/13 Academic 13/14) success rates are in line with the national average; in the remaining two years (QSR 11/12 Academic 12/13 and QSR 12/13 Academic 14/15) they are below, and at -8.9% in the latter year, considerably below. In addition, when considering the whole set, with the exception of one cycle (QSR 11/12 Academic 12/13), the relationship between success rates and the national average does not suggest a strong indication of student success rates being high. Moreover, of the 32 colleges inspected within this set, 14 have recorded a success rate in line with the national average. This is particularly marked in QSR 12/13 Academic 13/14, where of the 13 inspections, 8 record a success rate which is within 3% of the national average and therefore, arguably, in line with the national average, rather than considerably above it.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average) %		
				Success Rates	Difference	National Average	Relationship to Success Rates	
10/11	2012	12/13	3	77.3 – 80.6	3.3	77.1	+0.2 to +3.5	
11/12	2013	12/13	8	76.2 – 86.8	10.6	79.7	-3.5 to +7.5	
11/12	2013	13/14	4	79.7 – 84.8	5.1		+/- to +5.1	
12/13	2014	13/14	13	82.1 – 87.0	4.9	82	+0.1 to +5.0	
12/13	2014	14/15	4	73.1 – 83.7	10.6		-8.9 to +1.7	
13/14	2015	14/15	N/A					

Table 4.6 The data range of colleges inspected against the documentary analysis (Code 87): The success rates for 19+ year olds is above average/high

The analysis in Table 4.7 references 37 college inspection reports where the performance of students aged 19+ was in line with the national average, slightly above it or slightly below it. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p47) the description for student success rates that require improvement states that 'the majority of groups of learners are likely to be in line with similar groups of learners nationally'.

It is quite positive to note that that in the majority of inspection cycles, the only exception being QSR 11/12 Academic 12/13, there is a relative similarity in the success rate data range of between 6% and 7.6%. In a similar vein, the relationship between the success rates and the national average on the whole shows a consistent trend, in that it starts below the national average and finishes in line, just below or just above it – with the exception of QSR 11/12 Academic 12/13 where it finished above it at +5.2%. However, where the relationship (between success rate and the national) starts below the national average, it does so, in two inspection cycles (QSR 11/12 Academic 13/14 and QSR 12/13 Academic 14/15) starting considerably below it. In addition, it should be noted that of the 37 college inspections in this set, 8 record a success rate below, rather than just below, the national average.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average) %		
				Success Rates	Difference	National Average	Relationship to Success Rates	
10/11	2012	12/13	4	73.7 – 80.0	6.3	77.1	-3.4 to +2.9	
11/12	2013	12/13	14	75.2 – 84.9	11.2	79.7	-4.5 to +5.2	
11/12	2013	13/14	4	73.0 – 80.0	7		-6.7 to +0.3	
12/13	2014	13/14	10	77.8 – 83.8	6	82	-4.2 to +1.8	
12/13	2014	14/15	5	73.1 – 80.7	7.6		-8.9 to -1.3	
13/14	2015	14/15	N/A					

Table 4.7 The data range of colleges inspected against the documentary analysis (Code 114): The success rates for 19+ year olds is in line with the national average or just above/below it

The analysis in Table 4.8 references 31 college inspection reports where the performance of students aged 19+ was low or below the national average. In the grade characteristics for 'outcomes for learners' (Ofsted 2012, p47) Ofsted states that where success rates are 'consistently low, too variable or in significant decline' they are identified within the inadequate judgement.

The data range between the highest and lowest success rate across the whole cohort is consistent to within two percentage points with the exact same range in the two years with the largest samples (QSR 11/12 Academic 12/13 and QSR 12/13 Academic 13/14). When comparing these success rates to the national average there are some variables. As might be expected in this judgements category, success rates, when compared to the national average, are low or considerably low in the majority of inspections. However, of the total set of 31 inspections, 10 have recorded success rates which could best be described as being in line with the national average and one, in QSR 12/13 Academic 14/15, as recording a success rate considerably above it.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average) %		
				Success Rates	Difference	National Average	Relationship to Success Rates	
10/11	2012	12/13	2	65 + 77.9	12.9	77.1	-12.1 to +0.8	
11/12	2013	12/13	10	65.9 – 78	12.1	79.7	-13.8 to -1.7	
11/12	2013	13/14	1	75.4	N/A		N/A	
12/13	2014	13/14	13	71.7 – 83.8	12.1	82	-10.3 to +1.8	
12/13	2014	14/15	5	78.8 – 89.6	10.8		-3.2 to +7.6	
13/14	2015	14/15	N/A					

Table 4.8 The data range of colleges inspected against the documentary analysis (Code 141): The success rates for 19+ year olds is below the national average, too low or poor

It is clear from the above analysis that, for students aged 19+ on long qualifications, there is some relationship between the textual analysis and the validated data used by the inspection team. Where that relationship is closest is where the performance of 19+ students is in line, just above or just below the national average. However, even in this judgement category, there are examples of inconsistency, especially where success rates are at their lowest. The relationship among the text within the inspection report, success rates and the national average is less reliable, where student performance has been judged to be too low or poor, and even less so, where it is judged to be high.

When considering the relationship between success rate and the national average for this whole cohort of 100 inspections, it is suggested that in 32 cases there is a potential

disconnect between the data and the judgements made. Moreover, there are too many examples across the judgement areas and across the years where success rates shown as an absolute and as compared to the national average, overlap, and where they thereby have the ability to lessen one's confidence in the judgements made.

4.6 Success Rates, Textual Analysis and National Averages for WBL Students (Apprentices)

Work Based Learning (WBL), or more specifically apprenticeships, have played, and continue to play, a central role in the provision offered by colleges of further education. This, at least in part, can be attributed to apprenticeships being part of a centralised government policy which is underpinned both by a financial framework (which has developed during this study and continues to change) from which colleges can grow financially, and by a national recruitment target, of 3 million by 2020.

Below are three tables (Tables 4.9, 4.10 and 4.11). Each one is based on a judgement or evaluative statement made by the inspection team and recorded in the inspection report. For this category, WBL refers to apprenticeships only, as the most substantive WBL programme. Of the 165 inspection reports used in my study, 126 cited a judgement on the standard of performance regarding apprenticeships.

In order to identify what relation there is between the judgement made by the inspection team and the validated data available at the time of the inspection, the tables also include:

- The validated success rate data used by the inspection team for apprentices (found in the available QSR report at the point of inspection), all ages, all levels;
- The national average used by the inspection team for apprentices, all ages, (found in the available QSR report at the point of inspection) and;
- The judgement made by the inspection team and published within each college's Ofsted report (these have been analysed and subsequently coded)

4.6.1 Analysis

The analysis in Table 4.9 references the 50 college inspection reports where the performance of apprentices was above the national average and/or high. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p46) it is suggested that an associated judgement on the performance of a cohort of students where the majority perform 'well above the norm' or where the majority perform above the national rate will be outstanding or at least good.

The relationship between success rate data and the national average in Table 4.9 is anomalous, with the exception of the QSR 13/14 Academic 14/15 where success rates, at their starting point, are above the national average. Excluding this exception, in all years the relationship between success rates and the national average, at their starting points, is below, and in two cases, noticeably so. In these two cases, QSR 11/12 Academic 12/13 and QSR 12/13 Academic 14/15, the data range between the lowest and highest success rate is 27.9% and 28.1% respectively. Moreover, with regard to the relationship between the success rate and the national average, at their starting points, they are considerably below, by -13.1% and -18.2%. Furthermore, for the same two years, the number of inspections where the judgement made in the inspection report of apprentices performing to above the national average, but where the success rate is below it, is close to a third for QSR 11/12 Academic 12/13 and is half of the set for QSR 12/13 Academic 14/15.

When considering this sample as a whole, it is worth noting that of the 50 inspections cited in Table 4.9, one fifth record success rates that are below the national rates and when these are in line with the national rate, this level increases to just under a quarter.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13		N/A			
11/12	2013	12/13	15	60.7 – 88.6	27.9	73.8	-13.1 to +14.8
11/12	2013	13/14	6	68.0 – 84.9	16.9		-5.8 to +11.1
12/13	2014	13/14	17	63.6 – 84.9	21.3	72.5	-8.9 to +12.4
12/13	2014	14/15	6	54.3 – 82.5	28.1		-18.2 to +9.9
13/14	2015	14/15	6	73.6 – 79.8	6.2	70.3	+3.3 to +9.5

Table 4.9 The data range of colleges inspected against the documentary analysis (Code 165). The success rate for WBL is above the national average/High

The analysis in Table 4.10 references the 28 college inspection reports where the performance of apprentices was in line with the national average, slightly above it or slightly below it. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p47) the description for student success rates that require improvement states that 'the majority of groups of learners are likely to in line with similar groups of learners nationally'.

Similarly to the analysis in Table 4.9, there are huge sizable data ranges between the lowest and highest success rates in this group, for example in QSR 11/12 Academic 12/13, in QSR 11/12 Academic 13/14 and in QSR 12/13 Academic 14/15 of 19.5%, 21.6% and 20.6% respectively.

With regards to the relationship between success rates and the national average, there is some indication, with the exception of QSR 11/12 Academic 12/13, that at the top end they are broadly in line, just above or just below it. However, at their starting points, in the majority of years, notably the first three in the table, the success rates are below or considerably below the national average at -8.3%, -16.1% and -18.6%. In addition, taking into account the description of the judgement (that WBL/apprentice performance is in-line, just above or just below the national average), there are actually six inspections within this set (a fifth) where success rates are considerably below the national average.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13		N/A			
11/12	2013	12/13	11	65.5 – 85.0	19.5	73.8	-8.3 to +11.2
11/12	2013	13/14	4	57.5 – 79.3	21.6		-16.1 to +5.5
12/13	2014	13/14	10	53.9 – 74.5	20.6	72.5	-18.6 to +2.0
12/13	2014	14/15	1	70.7	N/A		-1.8
13/14	2015	14/15	2	68.4+74.4	6	70.3	-1.9 to +4.1

Table 4.10 The data range of colleges inspected against the documentary analysis (Code 174). The success rate for WBL is in line with the national average or is just above/below

The analysis in Table 4.11 references the 48 college inspection reports where the performance of apprentices was low or below the national average. In the grade characteristics for 'outcomes for learners' (Ofsted, 2012 p47) Ofsted states that where success rates are 'consistently low, too variable or in significant decline' they are identified within the inadequate judgement.

The data range and the relationship between success rates and the national averages, particularly at the lower end, are quite extreme. There are several inspections in every year where the performance of apprentices is very low, which might be expected within this category. However, of the 48 inspections, only half record success rates below the national average rate. The other half either record success rates in line with the national average or above it. Where the relationship differs most, is in QSR 11/12 Academic 12/13, where 35% of inspections record a success rate on or above the national average and in QSR 12/13 Academic 14/15 where 46% of inspections have a success rate on or above the national average.

QSR Report Year	Year of Inspection	Academic Year	Number in Sample	Data Range (Success Rates) %		Data Range (National Average)	
				Success Rates	Difference	National Average	Relationship to Success Rates
10/11	2012	12/13		N/A			
11/12	2013	12/13	14	59.9 – 75.6	15.7	73.8	-13.9 to +1.8
11/12	2013	13/14	5	49.1 – 88.1	39		-24.7 to +14.3
12/13	2014	13/14	13	37.5 – 72.2	34.7	72.5	-35.0 to -0.3
12/13	2014	14/15	8	50.6 – 82.3	31.7		-21.9 to +9.8
13/14	2015	14/15	8	41.3 - 70	28.7	70.3	-29.0 to -0.3

Table 4.11 The data range of colleges inspected against the documentary analysis (Code 183). The success rate for WBL is below the national average, too low or poor

The performance of apprentices outside of any comparison is highly variable, with the most notable and consistent success rate ranges - most above 20%, some above 30% and one at nearly 40% - between the lowest success rate and the highest in any specific year. This then gives rise to an equally variable relationship between the success rates and the national average.

There are too many examples across all of the judgement areas where the data does not support the inspectors' published findings. For example, where the inspection team have identified as part of the inspection that the performance of apprentices is high or in line with the national average, four (of the ten combined years) start with a negative relationship between the success rate and the national average of more than 10 percentage points. Moreover, where it might be suggested that the judgement of apprenticeship is high, resulting in a wholly positive relationship between success rate and the national average (Table 4.9), it is noted that in every year, with the exception of QSR 13/14 Academic 14/15, success rates at their starting point are below the national average.

4.7 Student Attendance: College Data, Textual Analysis and its Use in Inspection

How well students attend their programmes of study, or not, is a central tenet of the 'outcomes for learners' theme. For colleges, and indeed all educational institutions, there is believed to be a clear link between a student's attendance and punctuality and how likely they are to stay on their course (retention) and to pass their course (achievement). Poor attendance can be the first indicator to identify whether a student is 'at risk' and whether they need some form of support – academic, financial or personal.

To date, there is very little information on attendance in the further education system. In addition, unlike the school sector where data is collected, analysed and made publicly available and where comparisons can be made and minimum expectations set, there is no national collection of attendance data from the college sector. As a result, there is no national average publicly available and no possibility of comparing one college's attendance statistics against another's. This means, that when a college is inspected and its attendance is reported and judged, the inspection is based on the experiences, perceived standards and opinions of the inspection team and individual inspectors.

Similarly, the way that student attendance in schools is recorded is set out in legislation (DfE, 2006). This piece of legislation includes detailed instructions on when a register should be taken, in what form and the four categories for attendance recording which include provision for those who are present, absent, attending an approved educational activity or unable to attend due to exceptional circumstances (*ibid.*, 2006 p2). There is no such guidance for the further education sector which is evidenced in the various attendance and punctuality policies publicly available. For example, in four attendance and punctuality policies from colleges that fall within the scope of this research, the number of register marks available to teachers ranges from three to eight – from the most simplistic of present, absent and late, to those that include a code for behaviour and one for an authorised absence. This level of inconsistency also applies to the expectation of when a register should be taken. In the pupil registration regulations (DFE, 2006, p2) students should be registered at the beginning of each of their morning sessions and at least once in the afternoon. By contrast, the guidance given to staff in the aforementioned college policies is also variable, with statements requiring that e.g. a register should be taken at the beginning of the session or on the day of the session and recorded no later than midday on the day after.

The above introduction to attendance and the ways it is recorded, analysed and presented within the further education sector aims to place the information provided by the colleges into context, the data of which is shown below.

The information summarised in Table 4.12 cites the number of responses from the freedom of information request against the two questions regarding attendance. The two questions sought to identify both what the college attendance rate was at the point of inspection, i.e. shared with the inspection team and representing the in-year position, most likely up to the week before the inspection (A1), and the final attendance rate for the completed previous year (A2).

Question	Number of Responses	OfL Grade 1	OfL Grade 2	OfL Grade 3	OfL Grade 4
What was the College's overall reported attendance at the point of inspection? (A1)	65	6	28	30	1
Data Range %		87.7 - 95	83 - 93	82 - 93	91
What was the College's overall attendance at the year end prior to inspection? (A2)	67	6	30	29	2
Data Range %		85 - 91	80.5 - 90	80 - 91	87 & 88

Table 4.12 Total number of responses received from the FoI regarding attendance

What is clear from the above table is that there is very little difference, if any, in the reported attendance by colleges at the time of their inspection. And it is this data that would have been used by inspectors to inform their judgements. The only consistent factor in this regard is the lowest attendance figure for those whose 'outcomes for learners' were judged to be outstanding and those who were judged to be good – just under five percentage points lower. However, this picture does not follow through to a difference between those judged to be good and those judged to be satisfactory, where the difference at the starting point of the range has been reduced to just one percentage point. Indeed, of those colleges judged to be good, ten fell into the data range of those judged to be outstanding (above 87.7%) and for those that were satisfactory the number was seven.

There is a similar, almost identical, pattern for the colleges' overall attendance at the year-end prior to the inspection. Across all grade categories, the starting point of the data ranges is down, which is in line with the findings of the Ofsted/AoC report on improving attendance and punctuality (Ofsted, 2013). This research identifies that over the course of an academic year, attendance generally declines (ibid. p9). However, whilst the relative difference in

starting points remains similar between those graded as outstanding and those graded as good, there are two notable differences: the data range between good and satisfactory colleges is now statistically very similar, and the majority of colleges judged to be good or satisfactory have similar attendance data to that experienced in those colleges judged to be outstanding.

A second layer of analysis on attendance data and its use in inspection is summarised in Table 4.13 below. This table correlates the judgements and comments made by inspectors as stated in the published Ofsted reports with the data provided by the college which would have informed the judgements made. It should be noted that the number in each of the categories in Table 4.13 is lower than the number of colleges who responded to the FoI request; this is because not all inspection reports make reference to attendance or if they do, it is not done in the form of a judgement. In the overwhelming majority of reports published, no data is presented on attendance to support the judgements made.

Academic Year	Documentary Analysis	Number of Colleges in Sample (A1) ⁵	Data Range %	Number of Colleges in Sample (A2) ⁶	Data Range %
12/13	High/attend well	6	86 – 92	7	86 – 90
	Requires Improvement/in-line	7	82 – 89.7	7	80 – 89.9
	Low/poor	3	85 – 91	3	81 – 87
13/14	High/attend well	6	85.2 – 90	5	85.6 – 87.2
	Requires Improvement/in-line	4	82.7 – 89.2	5	82 – 87.3
	Low/poor	7	83.7 – 91	7	80.5 – 87
14/15	High/attend well	3	93 – 95	4	83 – 90
	Requires Improvement/in-line	3	88 – 92.4	4	88 – 88.1
	Low/poor	4	82 – 88.7	4	84 – 87.5

Table 4.13 Presentation of attendance data by academic year, documentary analysis and the type of attendance data

Where an explicit judgement has been made in an Ofsted report, it generally falls into one of three categories: where attendance is high or where students attend well, where attendance requires improvement or is in-line with expectations, and where attendance is low, too low or poor. In a small number of reports, where attendance either requires improvement or is

⁵ A1 represents the attendance data up to the point of inspection (in-year)

⁶ A2 represents the attendance data at the year-end prior to the colleges inspection (whole year)

low, some reference is made to attendance being below the college's own target. On no occasion is the college's target published within their inspection report.

When assessing the attendance data at the point of inspection (A1), against the judgements made in the individual inspection reports, there is no direct relationship in the academic years 12/13 and 13/14 although the attendance range is slightly smaller where attendance has been noted as high. However, for the same years, there is a considerable overlap in the data where the judgement on attendance is 'requiring improvement' and where it is judged to be low, noting that the bottom of the range in 12/13 where attendance requires improvement is 3% lower than where attendance is low. Similarly, in 13/14, there is only a one percentage point difference across the same categories.

In the academic year 14/15 there appears to be the outline of a trend, albeit a very tight one with a slight overlap (of less than one percent) between those requiring improvement and those with a judgement of having had poor attendance.

When considering the judgements of attendance against the data provided by colleges at the year-end, and for the year prior to an inspection, there is no clear trend. In each academic year, there is some overlap. For example, in 14/15, the starting point of the attendance data for those judged to require improvement is eight percentage points higher than for those whose attendance was judged to be high. And in 12/13 the top of the range for colleges whose attendance was high and required improvement is the same as they are across all judgement areas in 13/14.

There is arguably very little or no systematic consistency between when attendance data is considered by the inspection team and the judgement made. In the context of the further education sector, where attendance is not collected locally or nationally, and where there is no statutory guidance on the marking of registers, colleges have the possibility of seeking guidance from the Handbook for the Inspection of Colleges (Ofsted, 2012). However, it is clear from the information held within the handbook that there are no hard descriptors or ranges of attendance data that equate to a grade or judgements. Attendance is mentioned in the section 'grade descriptors' and it states under each judgement that attendance should be:

- Consistently good (in 'outstanding');

- Good (in 'good');
- Mostly at an acceptable level (in 'requires improvement'); and
- Are generally low (in 'inadequate').

However, there is no further guidance on how, in the form of data or other information, each of these attendance statements is to be interpreted.

4.8 Value Added: College Data, Textual Analysis and its Use in Inspection

The progress that students make whilst at college (and sixth forms) is another key performance measure that Ofsted inspectors consider, judge and report on as part of the inspection when considering 'outcomes for learners'. Ofsted states that there are three areas in which progress should be evaluated. In summary, they are:

- For students aged 16-18 on Level 3 programmes, known as value added (L3VA)
- For students aged 16-18 on intermediate and foundation programmes
- For adult students

The underlying premise of the three categories above is that the progress of an individual student (and therefore of groups of students) is measured over a period of time, normally the duration of their [full-time] course of study. This takes into consideration their qualifications at the beginning of their course indicating the standard by which they should achieve their current qualification. However, in reality, this process only applies to the first category for students aged 16-18 on Level 3 (or advanced) programmes. For this cohort of students, there have been several well established systems used by many colleges and sixth forms to predict/target-set achievement based on a student's GCSE results. The original systems were established for A-Level provision only, but were subsequently expanded to include vocational courses such as the then BTEC National Diplomas (which formed the majority of the Level 3 cohort studying in colleges of further education). Ofsted (2012a, p11) states that, covering the duration of this study, they have used the Learner Achievement Tracker (LAT) as the main value added and progress measure for advanced-level learners aged 16-18 for graded qualifications. Prior to this (and indeed throughout the 2012-2015 inspection cycle) colleges were using a range of value added systems including the more well established ALIS and ALPS and in some cases, well and less well established institutionally devised systems.

Measuring value added and progress for students aged 16-18 on intermediate and foundation programmes and for adults is less systematic. Ofsted (ibid.) states that in both cases, as part of the inspection process, it will be looking for evidence on how individual colleges capture a student's prior attainment and initial/diagnostic assessment, how they use that information for target setting and of course, how they measure progress against agreed targets. As a result, and for the purpose of this thesis, these last two categories have not been considered, neither was the data requested as part of the FoI due to the lack of a standardised framework by which any results could have been analysed.

Ofsted, (2012a, p11) states that from 2009 'Ofsted have used the Learner Achievement Tracker as the main value added and progress measure for advanced-level learners aged 16-18 in schools and colleges'. Therefore, for the purpose of consistency and transparency, the data used to support judgements and commentary in this area should be supported by a single data source, arguably in a similar vein as that used for success rates. However, the information included in Table 4.14 below, taken from the result of the FoI request to all colleges inspected within the timescales of this study, indicates that this is not the case.

The question in Table 4.14 was actually a follow-up to another question which was: At the point of inspection, did the College formally use and report on value added as a KPI? Of the total of 69 responses, 48 responded that they did (69.6%) and 21 stated that they did not (30.4%). Of those that responded stating that they did use and formally report on value added, the system or process that they used was varied. The Learner Achievement Tracker was used, either as a single system or in conjunction with another reporting tool, in 24 cases (50%). However, the use of the LAT as a single value added system only occurred in 8 colleges of those that responded to the questionnaire (16.6%), i.e. in a minority of colleges.

Where other value added systems were reported to have been used in colleges at the point of inspection, ALPS was cited as being used in the majority of cases, either as a standalone system, in 15 colleges (31.25% of those colleges that responded stating that they used a VA system), or in conjunction with another system in 26 college (54.2% of those colleges that responded stating that they used a VA system); it was therefore being used by the majority of colleges within this sample. Although small in number, it should be noted that a few colleges (eight), at the point of their inspections, were using a college devised scheme.

What type of value added tool was used by the college at the point of inspection?	
Value Added Tool	Number of Responses
ALI	2
ALPS	15
ALPS & ALI	1
College devised	6
LAT	8
LAT & ALI	4
LAT & ALPS	10
LAT & College devised	2
Didn't have/use one	21
Total Responses	69

Table 4.14 FoI responses on what type of value added tool was used by the college at the point of inspection (V1Y)

So, if it can be seen from the analysis above that the systems used to measure value added were many and varied, or indeed that many colleges didn't use a system at all, was value added explored as part of the inspection process and if so, what weighting was applied in relation to 'outcomes for learners'?

The 66 responses received to the question about whether inspectors explored value added data and information (Table 4.15) suggest that, in the majority of cases (55), value added data and information were explored with the college by the inspection team. However, the mere fact that they were explored bears no relation to how much weight the college thought that the inspection team placed on them; indeed, where colleges responded that they were not using a value added system at the point of inspection, the question arises as to what data and information were being explored?

At inspection, did the inspectors explore value added data and information?	
Yes	No
55	11

Table 4.15 FoI responses on whether the inspection team explored value added data and information

With regard to the question of how much weight the inspection team placed on value added, as a tool for measuring student outcomes as opposed to just exploring the data, thirty seven (of fifty five) or just over two thirds of respondents stated that they believed that hardly any weight was placed on value added as a tool for measuring student outcomes, with the remaining third believing that their inspection team gave 'significant weight' to the data and information on value added – as recorded in Table 4.16.

How much weight was placed on value added (by inspectors) was a tool for measuring student outcomes?	
'Significant Weight'	Hardly Any
17	38

Table 4.16 FoI responses on how much weight was placed on value added as a tool for measuring student outcomes

Having considered the responses directly from colleges about the system that they did or did not use at the point of inspection, about whether value added was explored by the inspection team and about how much weight they thought was given to it by the inspection team, it is important to look at what and how Level 3 value added was judged and reported on.

Table 4.17, taking information from the published Ofsted reports for the 165 colleges in this sample, analyses two aspects of how Level 3 value added is judged and reported on. As can be seen in Table 4.17, the summary judgements fall easily in to three distinct categories and the second aspect records where no reference or no judgement was made.

The analysis is stark. Of the inspection reports where a judgement was made and recorded (n=99), only 17 (17.2%) identified that students made better than expected progress when compared to their starting points, that in nearly 50% of this sample, students made either uneven progress or did not make the progress expected of them and that just over a third of the sample made the expected progress. Furthermore, no judgements were made or recorded in 66 colleges which fell within the total sample. This apparently concludes that the majority of 16-18 year old students studying for a Level 3 qualification, in many cases a pre-university qualifying qualification, failed to make better than expected progress.

Summary Judgement	Number in Sample
Students make better than expected progress	17
Students make expected progress or the progress that students make is in-line with what is expected of them (based on their starting points)	34
Students make uneven progress or students do not make the expected progress (based on their starting points)	48
No judgement made regarding Level 3 Value Added/Progress Measure	66

Table 4.17 Documentary analysis of judgements made regarding Level 3 Value Added/Progress Measure

Finally, if the progress that students make whilst studying for their qualification is at the heart of an education system, as it could be argued it should be, then it could be reasonable

to assume that there might be a relationship between the progress that students make and the overall grade awarded for outcomes of learners.

Table 4.19, which uses the same set of colleges analysed in the above table (Table 4.17), captures and re-distributes the summary judgements under the grade awarded for 'outcomes for learners'. What this table identifies is, that for this cohort of college inspections, that there is no relationship between the progress made by students aged 16-18 at Level 3 and the overall grade awarded for outcomes. With the exception of the three colleges awarded a Grade 1 (where the numbers are too small to render them reliable) the majority of students are judged to be not making better than expected progress. For example, where 'outcomes for learners' have been awarded a good grade and where the grade characteristic states that 'the majority [of learners] are making better than expected progress on their learning programmes' (Ofsted 2012, p46), 22 colleges out of the 33 in this cohort awarded a good grade failed to meet the requisite standard. Similarly, where a college has been awarded a grade 3 and where students should, according to the grade descriptor for this category, be making satisfactory progress with only a small minority making less than satisfactory progress (ibid.), in 32 colleges out of 50, student progress has been cited as being uneven or as where they do not make the expected progress.

Summary Judgement	OfL Grade 1	OfL Grade 2	OfL Grade 3	OfL Grade 4	Total
Students make better than expected progress	2	11	1	2	17
Students make expected progress or the progress that students make is in-line with what is expected of them	0	13	17	4	34
Students make uneven progress or students do not make the expected progress	1	9	32	6	48

Table 4.18 Analysis of Level 3 value added judgements by OfL grade

4.9 Progression: College and Type of Data and Textual Analysis

Progression, in Ofsted terms (Ofsted 2012a, p8) is when a student moves 'from one level of course to the next within the same provider'. It is acknowledged in the same Ofsted document (ibid.) that the information used by the inspection team to make an evaluative statement and/or judgement will be provided by the college and that the college will have its own systems to 'record, monitor and evaluate' that performance.

Of the 165 colleges that were written to under a Freedom of Information Request, 67 responded to the question regarding the type of data used by the inspection team. Of the 67 responses, 20 indicated that the type of progression data used in the inspection was intended and the remaining 47 indicated that the type of progression data used in the inspection was actual. Intended progression is a prediction, as opposed to actual progression.

At the outset, it is clear that any evaluative statements and/or judgements on how well (or not) students progress within the college are based on two different methodological approaches to data collection.

Question	Total Number in Sample	Intended	Actual
At inspection, what type of progression information was used? Intended progression or actual progression?	67	20	47

Table 4.19 College responses to the type of data used by the inspection team to assess the standards of student progression

Below, Table 4.20 shows the judgements made by inspectors about how well students are progressing to higher levels of study within the college and the data provided by the college which would have been used by the inspections to support such judgements. The data and information were further analysed by the academic year in which the college was inspected.

It is obvious from the textual analysis of the 50 inspections, whose colleges provided the progression data, that in the overwhelming majority of cases, the judgement on how well students progress to higher levels of study is positive. Just over three quarters of the inspections across the sample recorded a judgement stating that a high number of students progress within the college to the next level of study.

However, the data to support such judgements is extreme with a data range of 54.8%, 68.3% and 59.6% in each of the academic years. And, although the numbers of inspections is low where judgements have been made that the progression of students requires improvement, is low or where too much data is unknown (12), the supporting data is variable and on occasions is better than that found in the most positive judgement.

There are two further areas worthy of note. The first derives from the fact that of the 50 inspections in Table 4.20, just over a quarter (28%) of the judgements made are based on intended/predicted progression, the majority of which support the most positive judgement. However, the progression rates, with the exception of three inspections in 14/15 are high, from 76% - 95%. This suggests that in those inspections where intended progression data is used, the judgement is more likely to be a positive one.

The second is the use of intended/predicted progression data throughout the academic year. Whilst an argument could be made for the use of intended/predicted progression data for inspections that take place early in the academic year, in September and October, (and only where it is used to support the actual progression data from the previous year), there is evidence to suggest that it has been used throughout the year. For example, out of the 12 inspections in 12/13 and the 11 inspections in 14/15 where the number of students progressing was deemed to be high, 4 in 12/13 (33%) and 5 in 14/15 (45%) were based on intended/predicted progression data, and the majority of these inspections took place between November and June.

Academic Year	Number in Sample	Documentary Analysis	Data Range (Progression Rates) %	
			Progression Rates %	Difference %
12/13	12	High number of students' progressing within the college to the next level of study	35 – 89.8	54.8
	1	Progression to the next level of study within the college is satisfactory or requires improvement	50	N/A
	2	Progression to the next level of study within the college is too low or too much data is unknown	74.8 + 93	18.2
13/14	15	High number of students' progressing within the college to the next level of study	27 – 95.3	68.3
	2	Progression to the next level of study within the college is satisfactory or requires improvement	54 + 76	22
	4	Progression to the next level of study within the college is too low or too much data is unknown	69 – 91	22
14/15	11	High number of students' progressing within the college to the next level of study	35.4 – 95	59.6
	2	Progression to the next level of study within the college is satisfactory or requires improvement	37 + 54	17
	1	Progression to the next level of study within the college is too low or too much data is unknown	45	N/A

Table 4.20 Presentation of progression data by academic year and documentary analysis

The analysis above suggests some clear discrepancies in the data used to underpin the judgements made regarding progression, with considerable differences against every judgement made by the inspection team and across every academic year (where there is more than one inspection). In too many cases, where the judgement made is that the number of students progressing to the next level of study is high, the supporting data is low. This inconsistency is reversed for those inspections which identified progression as being too low or where too much data is unknown. In these cases, especially in years 12/13 and 13/14 (although the sample is small) there are examples of progression rates being very high.

4.10 Destination: College and Type of Data and Textual Analysis

In addition to the progression of students from one level to another within the organisation (as analysed above), Ofsted also considers and makes evaluative judgements on the progression of students from the provider to destinations including to higher education, further training and employment (Ofsted 2012a, p8). Similarly to the data used for progression, it will be collected, analysed and evaluated using a system unique to each college.

Of the 165 colleges within this study, 63 responded to the question on destinations regarding the type of data used by the inspection team at the point of their inspection. Of the 63 responses, 24 (38%) responded that the type of data used by the inspection team was intended (information normally provided by students approaching the end of their studies) and therefore a projection of intent. For those colleges who had collected and presented to the inspection team actual destination data, the percentage of known outcomes is likely to be greater in a number of cases, due in part to the accuracy of former students' contact details, their responses and the staffing resources needed to collect such information.

There is greater complexity when considering projected destination data, as information is likely to come from more than one source. Unlike progression where students are still within the same organisation and can therefore be identified from internal information systems, destination data can include:

- Information from UCAS on the number of students who took up their place at university or higher level institution

- Information fed back directly to colleges by former students
- Information fed back directly to colleges, where an employer has been identified within the college’s system

When considering destinations as a whole, a large proportion of the sample in Table 4.21 stated that the information used at their inspection was projected, signalling the use of at least two types of data used by inspection teams.

Question	Total Number in Sample	Intended	Actual
At inspection, what type of destination information was used? Intended destination or actual destination?	63	24	39

Table 4.21 College responses to the type of data used by the inspection team to assess the standards of student destination

Table 4.22 shows the judgements made by inspectors about how well students are progressing after their studies to destinations including higher education, further training and/or employment together with the data provided by the college which would have been used by the inspection team to support the judgements made. The data and information were further analysed by the academic year in which the college was inspected.

Across all years, but especially in 13/14 and 14/15, the majority of inspections concluded that a high proportion of students recorded a positive destination. Of the 50 inspections where the destination data has been provided, 33 recorded the highest judgement, although the data suggests a different picture – one of some considerable variety. For example, in 13/14 for the 14 inspections that stated that a high proportion of students recorded a positive destination, the data from the lowest to the highest destination rate has a range of 68%. Similarly in 14/15 for the same category, the range is 44.2%. And, in most of the other categories, when the number of inspections in the sample is small, the range is still worthy of note, often between 10% and 20%.

Of the 47 inspections in Table 4.22 where we have the data used by the inspection team provided by the college, nearly half (22) use intended destination data. Unlike that presented for progression, this data is far more varied in its range. Furthermore, where intended destination data has been used by the inspection team, in 18 of these inspections (82%), the conclusion has been that a high proportion of students recorded positive destinations even through some of lowest data presented is within this sample.

It is also worth noting that intended destination data was used by the inspection team in 20 of the 22 inspections, between December and June. It might be reasonable to assume here that actual destination data would have been available to inspectors or if this was not the case, the judgement awarded might have indicated that too much data was unknown. From both the textual and data analyses, this clearly was not the case.

Academic Year	Number in Sample	Documentary Analysis	Data Range (Destination Rates) %	
			Destination Rates %	Difference %
12/13	8	High proportion of students record positive destinations	60 – 79	19
	3	A good proportion of students record positive destinations or the number of students recording a positive destination requires improvement	72 – 85	13
	4	Positive destinations are too low or to many are unknown	76 – 96	20
13/14	14	High proportion of students record positive destinations	25 – 93	68
	2	A good proportion of students record positive destinations or the number of students recording a positive destination requires improvement	66 + 76	10
	2	Positive destinations are too low or to many are unknown	86 – 99	13
14/15	11	High proportion of students record positive destinations	51.8 – 96	44.2
	1	A good proportion of students record positive destinations or the number of students recording a positive destination requires improvement	36	N/A
	2	Positive destinations are too low or to many are unknown	55 + 58	3

Table 4.22 Presentation of destination data by academic year and documentary analysis

There is no discernible trend in the analysis above which indicates a clear and transparent use of data to support the judgements made with regard to student destinations. Much of the data provided by colleges overlaps, and in all years, especially when the most positive judgement is recorded, the data range is extreme. Moreover, the use of intended destination data not only includes huge variations, but it is over represented in inspections with the most positive judgements and it is used by the inspection team when it would be reasonable to expect actual destination data to be available. This gives rise to an unsteady grounding for the judgements made by inspectors.

4.11 Conclusion

There is a compelling argument to be made, based on the analysis in this chapter, that the relationship between the judgements made by inspectors and published in the inspection report, and the data on which they are based is inconsistent. Moreover, there is a severe lack of clarity and transparency within the reports themselves. The judgements and evaluative statements on student performance and success rates within the inspection reports do not present the data on which they are founded, either as an absolute or as compared to the national average. Furthermore, it is not stated anywhere in the published report what dataset is being used, and whilst the analysis above is based on a strict interpretation of the term 'validated data', it is conceivable that local or college based data might have been used. If this is the case, it is possible that further inconsistency between the use of success rate data and judgements ensues.

This apparent lack of consistency follows through to the other criteria considered by inspectors under 'outcomes for learners'. Due to the lack of a national framework for the measurement of attendance within the further education sector, it is clear from colleges' individual policies that the collection and analysis of data are based on different parameters; there is evidence that suggests that the attendance data used by the inspection team also differs between one inspection and another. The same can be applied to both progression and destination data, the collection and analysis of which are based on a college's individual system, and which are made more complex by the type of data used, building into the inspection system a potential systemic lack of comparability.

This analysis strikes at the heart of inspections. 'Outcomes for learners', is the inspection theme which has the potential to be objectively and consistently measured, in a way that is free from local interpretation or inspector subjectivity. Arguably, it is also the theme which gives the clearest understanding to all stakeholders of student performance and academic standards. As a result, these findings raise fundamental questions about how data is used and consistently interpreted by inspectors, what data is being used, and whether key stakeholders are clear about the data on which judgements are made.

Chapter 5: Discussion

5.0 Introduction

My aim in this chapter is to extend the knowledge and understanding about the collection, analysis and use of performance data and how that data is used by inspectors across an inspection system to evaluate student performance and make judgements on 'outcomes for learners' in the further education sector.

To date, all of the areas concerned with school and college inspections that I have found in the research literature deal with the impact of inspection. None to date seem to raise questions about the process or about the way data are used. Therefore, in this discussion the bulk of the literature referred to is claims made by Ofsted itself about the consistency, reliability and transparency of data-based/driven judgements.

As a result the constituent theme, from which the analysis and accompanying narrative in this chapter are derived, is taken directly from the inspection handbook (Ofsted, 2012), although its reach might be applied to the current inspection framework (Ofsted, 2015). It is in itself a critical evaluation of those who evaluate others, citing areas that require improvement.

The chapter sub-headings below are the key research questions identified in Chapter 1 and that are dealt with in my study.

5.1 How consistently are data and information used by inspectors to evaluate and make judgements on student success?

At the heart of any educational institution must be the aim that every young person or adult who has enrolled onto a qualification or study programme - after they have applied, been selected and obtained any entry criteria or condition of entry - will successfully achieve their qualification. It is hoped that that this will be an aim shared equally between the student and the institution in which they will study. It is therefore reasonable to suggest that when key stakeholders, including potential students and their parents, are looking for a further education college in which to study, they are able to contrast and compare institutions with ease, and know that when they are looking at two or more inspection reports the basis on which student achievement is evaluated is consistent. Indeed, Ofsted (2009, 2012 & 2015) holds the principle that inspections will be 'transparent and consistent'.

My study has used the validated data which would have been available at the point of the college's inspection. This success rate data would have been the evidence on which inspectors based their findings in order to evaluate student performance and on which they make judgements; both the handbook (Ofsted 2012 p33) and the supplementary guidance (Ofsted 2012a, p7) make this clear. Furthermore, Ofsted (*ibid.*, p8) states that the use of this validated data enables inspectors to compare success rates with national performance – again suggesting that all colleges in this aspect of an inspection will be assessed against an agreed set of criteria.

From Chapter 4, it is clear that there is considerable variation across inspection teams when evaluating the success rates of students, across all three student cohorts and across judgement areas.

Success rates

For colleges, success rates, like A level and GCSE results for a secondary school, are the most public demonstration to the outside world of how well they teach and support their students to achieve, and indeed of their academic standards. Like schools, but with much less visibility, recognition or understanding in the wider community, success rates are often cited in a college's marketing literature or as part of its recruitment campaign seemingly to inform prospective students, parents and guardians and to gain some sort of competitive

advantage. However, there is normally very little context to the success rate and moreover, it is also possible to see just the achievement rate being cited; this is often much higher than the success rate and will only indicate the achievement rate of those who were still enrolled at the end of their qualification and it will not take into consideration the number of people who dropped out along the way. This presents a very unclear and confusing picture.

It can be posited that Ofsted's approach to the use of success rates, when evaluating student performance and on which judgements are made, is equally unclear and indeed confusing. Ofsted inspection reports clearly articulate a threshold evaluating student performance – whether success rate are high, or in line or below national rates. Therefore, it could reasonably be argued that the data that underpins these judgements might follow a similar pattern, especially as there is very little, often no, data within the inspection reports to indicate otherwise.

Moreover, what we do see in the textual and data analyses, is a lack of clarity regarding the consistent application of success rate data to evaluate student performance. There are many examples where the data range in any one year and across each of the cohorts is considerable and overlapping, and when that data is applied to the evaluative statement and judgements, the results are nebulous. This therefore calls into question the accuracy of the judgements made by inspectors and subsequently the true position of a college with regards to how well it supports students to achieve the qualification(s) that they are studying for and their overall academic standard – the two questions young people and adults might be seeking answers to when looking for a college in which to study.

National average

The use of the national average as a marker against which student success is evaluated and judged is variable and in too many areas considerably so. Ofsted (2012a, p10) states that the application of national averages is only a guidance and that there are no threshold values or benchmarks, despite, as has been shown in my study, the use of evaluative language signifying such a value. Ofsted (ibid.) also makes clear that the analysis of success rates should begin with the 'all' national average, suggesting a common approach to the relationship between the assessment of success rates and the national average across all inspections. However, this guidance goes on to suggest that contextual factors may result in a different national average being used. So, for a reader to assume a like for like comparison

between colleges as a way of assessing academic standards, such a comparison would need to be based on the assumption that they (students, parents, teachers etc) have access to the data used by inspectors and that the inspectors are clear about which national average is being used.

In the overwhelming majority of reports analysed, the 'all' national average is used, although the terminology identifying it varies and includes the term 'similar colleges', 'general further education colleges' or 'sector averages'. For example:

'Overall, for the last three academic years, the number of learners successfully completing their course has improved and is now just above the national rate for similar colleges' (IR 124).

and

'Overall long course success rates have steadily improved over the last three years and are now in line with sector averages' (IR 94).

However, it should be noted that there are references to other national averages being used, but only in a very small number of inspections.

'Outcomes for learners are good and have improved over the past three years. The proportion of learners successfully completing their courses improved...and was above the average for colleges with learners from similar socio-economic circumstances' (IR 101).

It is likely, where socio-economic factors are being considered, that the national average will be lower than the average for the sector. Conversely, as identified by Ofsted (2012a, p10), one of the other areas where an alternative national average could be used, is where there is specialist provision within a general further education college, such as A levels. Where this is the case, inspectors have the option to apply the 'sixth form' national average, which, due to possibly higher entry criteria, will have a higher national average than that for studying A levels in the further education college sector.

What is not clear from any of the Ofsted guidance documents or its handbook, is what the decision making process is for the application of a national average outside of the 'all' national average. There is no clarity on who decides if one college's A level offer should be judged against other further education colleges, or against specialist sixth form colleges, or

indeed whether a college which caters for students from an area of high-ranking social and economic deprivation should be measured against colleges within similar communities. Thus, it could be argued that not only will the readers of a college's inspection report be oblivious of the fact that student performance may be evaluated against the national average, but there is also the distinct possibility that they will be equally oblivious of what national average is being used.

Therefore, if one of the key purposes of inspection is 'to help inform [users] about the providers they use or about providers they may want to use in the future' (Ofsted 2012 p5), then there is a lack of consistency and transparency regarding:

- how success rates directly relate to the judgements made by inspectors;
- which national average inspectors are using to evaluate student performance; and,
- which contextual factors inspectors are applying, as these can render any meaningful comparison between one inspection and any other inspection unsound.

When it comes to the use of data and information and their use in the evaluation, judgements and reporting on the achievement of students, Ofsted's claim of consistency is flawed.

5.2 How is attendance evaluated by inspectors? Is the evaluation consistent across inspections?

Ofsted (2013, p5) is clear that 'no national measures of attendance rates at FE colleges exist' and it therefore interpreted that an inspection team 'are unable to compare their [the] data with those of their [a] statistical neighbour'. However, the evaluation of attendance by inspectors remains a highly important aspect of a college's inspection, demonstrated by the fact that, of the 165 inspection reports analysed in this study, 111 make some kind of judgement about attendance.

To the public in general, attendance in the further education sector is probably akin to success rates – in that assumptions can be made about what they believe attendance is, when in fact, the picture behind the one line judgement in an inspection report is far more complex. Therefore, if attendance continues to be such an important aspect to be evaluated and reported on by inspectors – whether it be as an indicator that suggests how well

students are enjoying the course, the likely retention and achievement rates, or how well the college is preparing students for the world of work – it could be argued that further clarification or more specifically a form of national guidance is needed.

Ofsted (2012a, p8) acknowledges that providers will have their own attendance data, which in the absence of any national guidance on how and when this data should be collected and analysed, pre-supposes that its evaluation by inspectors will be subjective and therefore any like for like comparisons between colleges must be negligible. Below are two illustrations which might call into question the validity of a college's attendance data, which inspectors evaluate and on which they present their judgements to the public as a standard:

- **Compliance:** I have made the assumption, based on my professional experience, that in all colleges there will be a set of criteria and/or guidance, no matter how basic or indeed formal, regarding the recording of attendance. From the attendance and punctuality policies identified and analysed in Chapter 4, it is clear that there is some, if not considerable, variation between colleges when it comes to how and when registers are to be marked. However, what is included in these policies and what can only be assumed by inspectors (as there is no published evidence to suggest otherwise) is that all registers do get marked. It is not unreasonable to suggest that in large colleges of further education, with huge student populations, multi-campus operations and a range of study modes, there will be some level of non-compliance with the marking of registers or indeed, where a register has not been marked for a more legitimate reason, such as a tutor being ill or having left the organisation, there will be examples of student attendance not having been recorded. Alternatively, the register may be completed after the point in time when it should have been taken and possibly by someone else.
- **Timeliness:** by contrast to the previous point, if all teachers and trainers within a college comply with the college's own policy, does that mean that attendance is being recorded in a timely way? Although small in sample, the attendance and punctuality policies used in this study act as an example which could easily be representative of the sector as a whole, whereby there are a variety of expectations on the timeliness of capturing attendance data. For some it is immediate (assuming that they have access to electronic registers), whereas for others, whilst attendance is expected to be recorded in the lesson, it does not have to be reported until mid-day for morning sessions and at the end of the day for afternoon sessions and, for

the remainder, within a time period of up to 24 hours. So in reality, there is at least the possibility that for some attendance will be unreported for a period of time and that for those with the longest reporting period attendance data has the ability to become less reliable.

In addition, Ofsted does not make clear what attendance is and how it is being measured. In the Ofsted reports analysed as part of this study, there is very little information to support the one or two sentences allocated by inspectors to the judgement(s) on attendance. In most reports, there is a judgement on overall attendance with a subsidiary judgement, normally attributed to English and mathematics. However, as we have seen, further education colleges are far more complex. Most colleges will offer courses that last as little as a two hours; many colleges will have an offer for part-time adults and in some colleges, these people will form the largest student group; and other colleges will offer multiple-year courses such as foundation degrees and degrees. Furthermore, there is further complexity when considering attendance for those on an apprenticeship – does the attendance data only include the time that these students spend in college or does it also include the time spent with the trainer? And, for young people on study programmes, is attendance on the compulsory work experience aspect of their programme recorded?

Only in a small number of the inspection reports analysed in this study is reference made to the fact that the college's measure of attendance is overstated. In no report is it indicated by how much or the difference between the data held by the college and the data identified by the inspection team. And, whilst in this small sample, where college's attendance rates and attendance observed by inspectors are different, there is no commentary or further judgement about the accuracy or appropriateness of how the college collects its data.

Therefore, it is realistic to assume that if a member of the public is reading a college's inspection report or seeking to make comparisons between two or more colleges with regards to attendance, they will have little or no idea of what attendance is actually being measured or whether the attendance data on which the judgement is based is reliable.

In nearly all inspection reports analysed in this study, attendance is allocated only one or two lines. There is very little or no information on how this judgement has been made. Whilst it is acknowledged by Ofsted (Ofsted 2012a, p10) that providers have a context within which they operate and that that context might be worthy of consideration when assessing

student success rates, there is no such acknowledgement that a college's context might also affect attendance. Therefore, it is unclear whether inspectors are evaluating attendance at a college (based on their own experience as a practitioner and/or inspector) against a college with a similar context, or whether they are evaluating a college's attendance against a perceived norm or expectation (in the absence of any national average). For example, where a college has its own (or leases a) fleet of buses, could it be reasonable to assume that attendance might be higher than at a college where travelling to college requires public transport across a city and is more complex? Or could it be considered that if a college is deemed to be in an area of high social-deprivation, young people and adults might have to make a choice between attending college as opposed to finding work with the ability to pay for travel, in contrast to finding a college in an area with low socio-economic deprivation where there might be greater access to personal/parental transport?

Finally, in addition to attendance reporting, colleges and inspectors will seek to confirm how punctual students are to lessons and if they are late, how that is recorded and challenged and how it might affect learning – not only for the individual who is late but for the rest of the class. However, in their analysis on how to improve attendance and punctuality Ofsted (2013, p14) asserts that the college managers interviewed not only reported 'difficulties with recording punctuality accurately' but that they could not agree on what 'constituted lateness'. As a result, it is again unclear about the reliability of data colleges are providing to Ofsted, which Ofsted is using to evaluate punctuality. If, as cited in their report (ibid.), one college states that any student not in class at the beginning of the session is late, and another states that anyone arriving before 10 minutes into the lesson is late and constitutes a late mark, parity in the way that this data is evaluated and reported is questionable.

So, as one of the key aspects of a college's inspection, it is unclear how attendance is evaluated and judged by inspectors. With no national guidance on what constitutes attendance or punctuality, and no national collection of attendance data (from further education colleges) and consequently no national average to provide a marker against which colleges or inspectors can consistently evaluate the standard of attendance, it is not an unreasonable assertion to suggest that the basis on which attendance is judgement by inspectors is flawed. Especially when the very data on which this assessment is based is unreliable.

5.3 To what extent is value added data used by inspectors to evaluate and make judgements on student progress?

'Learners enjoy learning and make progress relative to their starting points' (Ofsted 2012, p43). This is one of the key aspects of an inspection which inspectors are asked to consider. Their findings will contribute to the judgement 'all learners achieve and make progress relative to their starting points and learning goals' (ibid.). Therefore, it might be reasonable to assume, that if a judgement is to be made on 'all' learners' progress relative to their starting points, there must be an expectation by Ofsted that colleges systematically collect entry data for all students and have a tool which uses that data to project a student's outcome.

However, as a result of my study, it is now known that, at the point of their inspection, at least 21 colleges did not have or did not use any form of value added – which represents just over 30% of colleges that responded to my study. And, despite Ofsted stating that only the DfE Level 3 Value Added table, the successor to the Learner Achievement Tracker (LAT) will be used as the data source on which the above judgement will be based (Ofsted 2012a, p11), the majority (65%) of colleges responded that they didn't use the LAT/L3VA at the point of their inspection. This therefore suggests that the extent to which value added is evaluated by inspectors, for the only cohort which has a national value added system, is questionable for the following reason: if a college does not use any form of value added system to identify what progress students are expected to make, but is then inspected and measured against a national value added tool, it is reasonable to assume that the data used by inspectors is procedural i.e. that the college does nothing with the data to change or affect its outcome. Therefore, it is also reasonable to suggest that if this data shows that students make worse than or no expected progress, it might be highly influential on the overall judgment on 'outcomes for learners'. However, there is little evidence to suggest that this is the case.

In addition, using again Ofsted's principles of clarity and transparency, it is difficult to imagine what a parent, young person or adult might reasonably understand value added to be – what is measured, how it is evaluated and indeed how to compare the judgements made about student progress in one inspection report with any other.

This is made even more complicated by the fact that, as stated above, only one student cohort, that of students aged 16-18 on level 3 programmes, which is a small cohort of the further education sector (just under 7%), has a nationally recognised tool for projecting and reporting on a student's expected and actual progress. It is expected by Ofsted, (2012a, p11), that for the remaining student groupings there will be a value added system designed by each individual college which captures prior attainment which in turn will be used to set targets. Furthermore, it is also expected that a college will monitor a student's progress against their agreed targets and it will be this process that inspectors will evaluate.

Ofsted (ibid.) cites two other student groups for which progress, when compared to their starting points, will be evaluated and judged – those studying at level 2 and below, and adults. It is acknowledged by Ofsted (2012a, p11) that compared with the relationship between prior attainment i.e. GCSEs and those studying for a level 3 qualification, this relationship does not necessarily exist for students on an intermediate or foundation level programme. Ofsted suggests that this relationship is more complex. In addition, Ofsted (ibid.) also acknowledges that, due to the time that might have passed for an adult student between leaving school and enrolling onto a college course, GCSEs are no longer a reliable foundation on which to predict a student's potential success; in both cases it is the responsibility of the provider to design a system to measure a student's progress.

Therefore, it is argued that it may be impossible for any reader of an inspection report to know and understand what data and information are being used by the college and evaluated by inspectors to assess how well (or not) students make progress –both for those aged 16-18 on courses below level 3 and for adults on all courses. There is no indication in any of the inspection reports analysed by me for this study to explain how a college measures progress for these two student groupings, and it might also be a reasonable assumption that given the situation with regard to a value added system for level 3 students, a number of colleges might not have any process in place.

Having considered how student progress is measured (or not) for the cohorts above, there is arguably a failure by Ofsted to support its 'all' learners statement and how they 'make progress relative to their starting points' (Ofsted 2012, p43). There is no provision in the handbook (Ofsted 2012) which suggests what progress made by apprentices relative to their starting point is expected. Whilst it might be that part of their programme might fall under one of the above groupings and their associated system, it does not state if progress in any

other part of their apprenticeship should be measured and subsequently evaluated. Furthermore, there is also an underlying assumption (Ofsted, 2012 & 2012a) that a student's progress can be measured only against qualifications that are graded. Many students, including students with learning difficulties and/or disabilities, who might not be studying for a qualification at all, or students who are studying for a technical qualification which is assessed on a pass/fail basis, are seemingly not covered by Ofsted's guidance. And, although the latest handbook states that 'Inspectors will consider the progress of learners in all types of provision, not just those who have taken or are about to take examinations of national tests' (Ofsted 2015, p50), there is no explanation as to the basis by which this will be achieved, no guidance on an agreed standard and no explanation within inspection reports of how student progress is measured.

It should be noted that there is some provision within the handbooks (Ofsted 2012 & 2015) that progress, in addition to the use of a value added framework within colleges, can be identified through direct observation by inspectors. However, whilst it is not unreasonable to suggest that progress, or at least a snap-shot of progress, can be evaluated by inspectors via direct observation, it is reasonable to expect that any judgements made are based on an adequate sample size. Many further education colleges will have thousands if not tens of thousands of students on a range of courses with various modes of study. However, as can be seen from the three inspection reports used in this study (below), the actual number of inspectors who could have made an assessment on how well students progress, in proportion to the size of college (in student numbers) indicates that any such sample is disproportionately low. This proportion is reduced because subject inspectors (who form the majority of an inspection team) will, according to the inspection timetable (Ofsted 2012, pp29-30), be inspecting provision for just under three days.

Inspection Number	Number of Students			Number of Inspectors	Ratio (inspectors: students)
	Total Number of students ⁷	Number of students aged 16-18 Level 3 ⁸	Number of students where VA is assessed through direct observation ⁹		
43	11,536	361	11,175	10	1:1,117
61	1,941	428	1,513	7	1:216
162	8,764	1,435	7,329	12	1:611

Table: 5.1 Ratio of inspectors:students regarding direct observation of student progress

The table above does not consider any additional complexity, such as multi-campus inspection or delivery of provision for a college via sub-contracted partners, which reduces even further the time available for inspectors to observe students directly.

Measuring a student's progress in relationship to their starting point is a clear aim and most likely achievable. Having the ability to measure all students' progress in relationship to their starting points is complex. Indeed, it is made difficult due to the lack of any national criteria or guidance, which affects the majority of the further education sector's student body.

Moreover, even where there is a national framework such as the DfE L3VA there is credible evidence to suggest that it was not used by a considerable number of colleges in this study at the point of their inspection.

Finally, there is too much that is unknown when it comes to the evaluation and judgements by inspectors. It is not known what population of the student body is directly observed and therefore how representative any findings are; it is not known what provision is included or excluded by inspectors and it is not known whether a college has a VA system (or series of systems) and if it does, how effective it is at identifying a student's potential.

Therefore, the lack of clarity about the points made in the above paragraph, in addition to the lack of clarity within the inspection reports used in this study regarding the evidence on which judgements are made, make it highly likely that the wider public who use these reports to compare institutions will have very little or no idea about what value added actually means and, moreover, how it is evaluated. More importantly, the judgements

⁷ As cited in the college's inspection report

⁸ As cited in the college's inspection report and where L3VA data applies

⁹ The total number of students minus those aged 16-18 on level 3 programmes

themselves, in the absence of the same points of clarity (cited above) might also be called into question; in addition, as suggested by the majority of colleges' responses, inspectors place hardly any weight on their contribution when assessing 'outcomes for learners' anyway.

5.4 How are progression and destination evaluated by inspectors and is this evaluation consistent across inspections?

Success rates can communicate how well a college supports its students to achieve their qualifications. And, if value added can inform a potential student or parent how well a college supports its students to achieve their academic potential, then progression and destination must be the tool whereby users identify how well a college prepares a student to progress to a higher level qualification and/or into employment.

Therefore, it is arguable that should a young person, adult, parent or any other such stakeholder look to an inspection report to form an opinion on which college best supports people to progress into further education and/or employment, they should do so knowing that they are comparing like with like. Otherwise, it stands to reason that if this aspect of an inspection is evaluated and judged using different criteria, data or parameters then any comparison between providers is arguably meaningless. It is clear that there is considerable variation in the type and reliability of data used when evaluating the progression and destination of students.

Type of data: it is expected by Ofsted that all colleges of further education will collect data, in some form and from at least some of its student body, regarding progression and destination (Ofsted 2012). Therefore, all colleges will have data which can be used by the college for the purpose of self-assessment and which inspectors can use as evidence to evaluate the extent to which a college successfully prepares its students for progression to further education and/or work. However, what is not made clear in any of the inspection reports used in this study is what type of data is being used and its robustness.

Similarly to success rates, colleges will use progression and destination data to promote to the wider public how successful they are. Again, similarly to success rates, it is possible that only certain aspects of that data will be used, probably where it provides the most positive picture. Where it is used in promotional material it is highly likely that, based on the lack of

any national guidance on the collection of progression and destination data, the data presented will not necessarily be based on the same parameters as others are. Furthermore, any such data presented to the public is not be accompanied by an explanation of how this data is collected, what exactly it is measuring or indeed how it is validated. This makes like for like comparisons extremely difficult.

It is therefore reasonable to suggest that any such like for like comparisons across colleges, based on the content of an inspection report, are equally as difficult or indeed impossible. There is no evidence within the inspection reports analysed in this study that suggest that inspectors are clear about what type of data is being used to evaluate a college's performance with regard to progression and destination. However, it is known that for the data collected as part of this study directly from colleges, both intended and actual progression and destination data is used by inspectors. Intended data, which usually has a much higher response rate, acts as a projection of a student's intention before they have completed their course, and is therefore un-validated. Actual data is likely to represent a truer picture, normally as a snap-shot in time, whereas for destination data at least there is likely to be a smaller, and in some cases, a much smaller response rate. Actual progression and destination data is more likely to be subject to some form of validation process. Therefore, it is perfectly reasonable to suggest that users, looking to an inspection report to find out how successful a college is at progressing students to further education and/or employment, will be ignorant of the different types of data used by inspectors and its resulting inconsistencies.

Sample size: when any data or aspect of an inspection is evaluated, it is reasonable to suggest that that the sample should be of a sufficient size to be representative of the whole. That way assurance is given to the user that any judgements made, in this case regarding student progression and destination, are based on a considerable body of evidence. However, the body of evidence which underpins the judgements made by inspectors and cited in the reports analysed in this study, is absent, and whilst there is commentary in some inspection reports that the college has not collected sufficient data (see example below), it is still unclear what proportion of the student body data on destination has not be systematically collected.

'...managers do not collect information systematically on destinations and, as a result, the college is unable to evaluate fully the extent to which its provision enables

students to progress successfully into employment and further learning or training, or identify areas for improvement'(IR 153):

Furthermore, due to an absence of guidance, there is no explanation about what groups of students' progression and destination data should be collected. The majority of further education college inspection reports make statements that might suggest that progression and destination data is collected from all students – or at least, as outlined in the example above from inspection 153, that it is expected to happen. (There is no indication to suggest that the collection of destination information is missing from any one particular grouping, curriculum area, or course). A further example is cited below which suggests that inspectors are evaluating the progression and destination of all learners.

'Internal progression of students from lower to higher-level courses within the college is good. This reflects the well-planned and appropriate curriculum pathways the college has successfully developed, particularly at intermediate level. As a result, success rates at intermediate level have improved greatly. Progression into employment and higher education is good. The proportion of students gaining suitable employment is much higher than might be expected from local economic circumstances'(IR 30).

However, in most colleges, including in the two examples cited above, there is a wealth of provision, across many levels and delivered via a range of delivery modes. So, would it be expected that a student on a half-day professional development course would have their progression or destination information collected? If so, what would be the intended destination? Whilst this might be an extreme example, many students within a further education college will be studying on a part-time basis – on one day, one week, 10 week or 18 week courses for example. Therefore, at what point is it expected that progression and/or destination information be collected? It might therefore be reasonable to assume that the data on which inspectors evaluate progression and destination data and information is collected, in the main, from students on full-time or substantive programmes. In many colleges, full-time students will be the minority student grouping, and this, as a result, raises questions about proportionality and sample size. And, where data is collected by colleges for its full-time students and used as a basis for judgements, there is no reference within the inspection reports used in this study to clearly indicate the proportion of the full-time cohort from which this data has been collected. Furthermore, there is no commentary within the report which, in addition to not making any reference, also misses out on the opportunity to explore the proportions of data collected, compared between intended progression/destination information and actual progression/destination information.

Therefore, it is reasonable to surmise that each college will have a different point at which they start to collect and analyse progression and destination data, which in turn makes comparability difficult and evaluation by inspectors inconsistent. And, whilst efforts are now being made to collect and present destination data at the end of Key Stage 5 (DfE, 2016), similarly to L3VA, this only applies to students aged 16-18 studying for A levels or a level 3 vocational qualification. Nationally, in 2012/13, this group represented 6.93% of the further education college student population and is clearly therefore no representative measure of the sector as a whole.

In a small number of the inspection reports, and as stated in the example on p129 from college 30, there is evidence that inspectors are considering the socio-economic context in which the college and/or their students operate. However, it is unclear to what extent this or any other such context within which the college operates is considered by inspectors, and to what extent, moreover, the consistency with which inspectors accept that a context or situation may affect a college's progression and destination picture from what must be a perceived norm (in the absence of any national benchmark) is also considered.

Therefore, how do inspectors evaluate and judge a college's success in preparing its students to progress to a higher level qualification and/or into employment? And how can this information help users about the provider? It is reasonable to suggest that the data on which inspectors can evaluate this aspect consistently from one college to another is inadequate. In each report there is an absence of any detail which would give the reader an idea of what basis inspectors evaluated progression and destination upon. It is not known which student groups are prioritised to be measured, if it is not the entire student body (which it is highly unlikely to be); it is not known whether the data used is a projection of intention or actual destination and whichever it is, what proportion of outcomes is known. And finally, it is unclear whether inspectors accept a college's proposition that their college operates outside a perceived norm and should be evaluated and judged accordingly. Therefore, it is reasonable to suggest that a young person, adult or parent reading two or more inspection reports to assess how well one college supports its students to progress onto a higher level qualification and/or into employment, will be doing so unaware that the basis on which judgements are made is inconsistent and potentially unreliable.

5.5 How transparent are inspectors about the type and validity of data and information used when evaluating, judging and reporting on outcomes for learners?

The focus of this thesis is on how data and information are used by inspectors to evaluate student performance and make judgements on outcomes for learning. To achieve this, there has also been some quite considerable analysis and commentary on the type of data that is being used, what the expectations are for its collection, and whether it is reliable. All of this underpins the levels of confidence that users may have in the accuracy of an inspection and their ability to compare like for like, as regards a college's performance compared to any other's. This, therefore, has the ability to impact on a student's or potential student's choice of institution.

However, whilst discussion in this chapter has been framed around the first purpose of an inspection, as cited by Ofsted (2012, p5), namely '[to] provide users with information about the quality of provision in England; to help inform them about the providers they use or about the providers they may use in the future', it is also worth highlighting the remaining two purposes of inspection (ibid.). They are:

- [to] help bring about improvement by identifying strengths and areas for improvement, highlighting good practice and judging what steps need to be taken to improve provision further
- [to] provide the relevant secretaries of state and other stakeholders with an independent public account of the quality of education and training, the standards achieved and how efficiently the provision is led and managed

It must therefore be reasonable to assume, as it has been earlier in this chapter, that there is commonality among inspections and that that commonality is transparently understood. For indeed, if Ofsted, through its annual report and through other mechanisms, reports on the state of the further education sector (or indeed any other sector) to the secretary of state, it would be perplexing if much of its analysis, especially on 'outcomes for learners', is based on subjective or unreliable data.

What has been identified, with regards to the use of data, its transparency and validity, in my study, can be summarised as follows:

Achievement rates (formerly success rates): achievement rates are the most validated dataset the further education sector has. It records all students who have enrolled onto a funded qualification, those who have been retained and those who have passed their qualification. Throughout an academic year, this data is submitted regularly to the Education and Skills Funding Agency (and its predecessors) and it is also subject to audit. There is clear guidance as to how and when this data should be collected.

However, its use in inspection is variable. There is evidence to suggest that the use of the national rate (formerly national average) arguably should, but does not, provide the measure from which achievement rates are evaluated. Indeed, it is unclear in some cases which national rate is being used and no clarity on how any deviation from the 'all' national is decided. Hardly any reports state the achievement rate on which their judgements are based.

Attendance and punctuality data: there is no consistent approach to the collection and analysis of attendance and punctuality data and indeed, there is no evidence to suggest a single criterion for what 'late' actually means. Inspectors expect that colleges will have their own attendance and punctuality procedures and it is on this basis that the data produced will be used as evidence within an inspection. There is little evidence to suggest that inspectors challenge a college's attendance data, except where it is explicitly stated in the inspection report, and there are very few examples of this happening. It is unclear if inspectors consider a college's context (as they might do for achievement rates) with regard to attendance and punctuality or indeed if the attendance of one part of the study body is weighed against another. In the overwhelming majority of inspection reports, there is no attendance or punctuality data presented.

Value Added: there is some nationally produced data which identifies the expected progress a student should make, and a summary report which identifies the number of students who achieved or exceeded that progress measure. However, this dataset only applies to less than 7% of the further education student cohort. Its impact is lessened by the fact that there is evidence of non-use of this system, for 16-18 year olds on level 3 programmes, by colleges at the point of their inspection. For the remaining 93% of the further education sector's student body, colleges are expected to devise and implement their own value added or progress measuring tools. No inspection report in this study identifies what that tool is, how

it is used and for what population of the student body it applies to. And, in both cases (nationally produced data and local data), beyond the evaluative statement, there is no data presented within the report.

Progression and destination: similarly to value added, there is now some nationally produced data on student destination. It caters for the cohort of students aged 16-18 on level 3 programmes, which is a minority of the student body across the sector. It is unclear for the remaining student population whose progression and destination data should be collected, especially for part-time students or those studying for a short professional development type qualification – although the text within inspection reports suggests that the data and information for an entire college community have been evaluated. There is evidence which identifies an inconsistent approach to the type of data used by inspectors and no identification of the proportion of known outcomes. In the overwhelming majority of inspection reports, no data is presented to support the judgement on the progression and destinations of a college's students.

The user should be confident that when they read an inspection report, especially the one section which is data rich and has the potential to be objective, they can do so knowing that there is comparability across inspections. They should be able to know instantaneously, without having to visit the national dataset, the proportion of students who achieve their qualifications and how that compares with the national average. That user should also know how well students attend their classes, as an indicator of expectations and preparation for work, and how much students enjoy going to that college. They should be able to compare attendance rates on a like for like basis with other colleges and again against the average.

An inspection report should also indicate, clearly, to this user the proportion of students who achieve and exceed their progress targets and know that only one system is used for this measurement. This user will be able to compare how successfully one college supports its students to progress academically against another one, and to decide which college might be better for them to attend. And finally, this user will be able to compare one provider against another regarding its positive progression and destination rates and knowing that, due to a sector wide agreement on how this data should be collected and validated, they do so with complete assurance.

However, this picture of a utopia is somewhat in the future, but it should be the basis on which all colleges collect data and against which all colleges self-assess and are inspected. My study has identified that where data is collected nationally, it is not used consistently when inspectors evaluate and make judgements. This, in turn, has the ability to be extremely unhelpful to anyone using an inspection report as a basis on which to make a decision on which college to attend. Moreover, where a college is expected to devise, implement and maintain a system which aims to capture essential aspects of a student's experience – such as the realisation of their potential, aspiration and futures – there is very little ability for these areas to be evaluated consistently or for data to be validated across the sector.

An inspection report should be more than informative (Hatton, 2017). An inspection report should evaluate with consistency the standards of one college in relation to the standards of the sector. An inspection report should be helpful. An inspection report should not raise more questions than it answers.

5.6 Implications of the findings

My study supports the fact that there is a case for change in how inspectors use data and information when evaluating student performance and judging 'outcomes for learners'. However, this change is predicated on there being a standardised framework for the collection, analysis and presentation of data for all aspects of an inspection where judgements are based on objective evaluation.

The purpose of inspection is clear. It is there to inform. To inform users and potential users about the standard of a provider; to inform the secretary of state about the state of the sector in England and to inform providers what their strengths are and what are areas that require improvement. However, what should not be underestimated is that Ofsted's ability to inform and the possibility for others to take it seriously, are predicated on the credibility and confidence of the inspection system to evaluate a provider's performance, accurately, consistently, transparently and objectively.

It is all too common for people who work in the further education sector, and the education sector more widely, to dread an inspection. It is looked at as something to fear, something to plan for and something to anticipate; it is constantly in the mind-set of teachers and

managers alike. No matter how frequent an inspection is or how many times they are experienced by teachers and managers, they are unknown. Unknown because of a different inspection team and unknown because the themes being inspected are different from the last inspection. The handbook is different and the weighting of whether 'outcomes for learners' or teaching, or learning and assessment is the most considerable contributor to leadership and management and overall effectiveness, has changed (but may change back).

Therefore, standardised data with clear threshold values, linked to evaluative criteria, has the ability to remove such fear (at least from one aspect) or the element of surprise from an inspection. This will be achieved by colleges having a greater understanding and an ability to self-assess student performance against sector norms and knowing, with confidence, how its performance relates to its neighbours'. In turn, a college's assessment of itself and what inspectors state about its standards of student performance, its strengths and areas requiring improvement, should be closely aligned.

As a result, the exploration and validation of data which take place during an inspection and its inconsistent presentation and its subjective analysis are removed. Thus, more time and resource from the inspection team can be spent in the classroom, directly observing the standard of teaching, learning and assessment, and speaking to students about their experience – creating a larger sample on which to evaluate and judge this aspect of inspection. Moreover, there is also a reasonable argument to be made, that this additional time could be spent supporting teachers to improve – suggesting teaching and learning strategies based on best practice from around the sector.

A further, and arguably the most important, implication of the findings as a result of my study is that users will be able to compare one college with any another on a like for like basis.

Chapter 6: Conclusion and Recommendations

6.0 Introduction

This chapter takes into consideration the data in Chapter 4 and the analysis in Chapter 5 and presents a range of recommendations which it is hoped will improve the transparency and consistency in evaluating and reporting on the performance of students and the judgements made by inspectors. Moreover, it is also intended that the principles behind these recommendations will support the self-assessment process, not only to enable a framework that will ensure better accuracy, but also to enable a clear synergy between what the college says about itself and what inspectors say about outcomes for students.

Furthermore, this chapter also identifies areas for possible future research. This research is intended to further understand the relationship between the use of data, the evaluation of student performance, and inspection across the further education sector. It concludes with an example of what a part of an inspection report might look like, if the suggested recommendations were implemented.

6.1 Conclusion

The consistency with which inspectors use data and information to evaluate student performance and make judgements on 'outcomes for learners' requires improvement.

The scope of my study was to identify with what consistency data and information are used by inspectors to evaluate student performance and make judgement on 'outcomes for learners' in further education colleges. To achieve this:

- the inspection reports from one complete inspection cycle have been analysed;
- success rate data and national averages from the same inspection cycle have been collected and catalogued; and,
- additional data used by inspectors, produced locally (by colleges) has been sought directly from the 165 colleges in scope of this study

It is not news that on any one day of a college's life and even more during an inspection, a college will produce a substantial amount of data – not just on all the aspects of student

performance highlighted in this study, but covering many more areas of a college's normal operation – in finance, HR and facilities for example. The way that this data is collected, re-produced and analysed in some cases will be set out externally to the college and as such will be directly comparable with other colleges, the sector as a whole, and possibly other educational, public or private institutions. These areas can be expected to be audited objectively and benchmarked with consistency.

However, there is evidence in this study to suggest that even when the data is collected and analysed using a national set of parameters and complying with pre-determined rules, it is not used by inspectors to evaluate student performance consistently or, it could be argued, accurately. The use of national averages, as a foundation from which success rates can be evaluated, is sporadic and success rates themselves are not constrained to the evaluative statements made by inspectors. There are too many examples where the judgements made by inspectors are unsubstantiated by the success rate data. Examples include colleges where success rates for a cohort of students are judged to be good, but the data is either in line with or below the national average or indeed, below that of a college judged on the same criteria to require improvement or to be inadequate. There are also other examples of inconsistency regarding similar permutations. Furthermore, there is a severe lack of clarity as to which national average is being used, whether it is the 'all' GFE national average or a specialist national average or indeed if the context of a college has been taken into consideration.

In addition, too much data and information is collected, re-produced and analysed locally, by individual colleges against what can only be described as, in the absence a standardised framework or guidance, presumed norms. And, more importantly, it is this data that is being used by inspectors to evaluate student performance and make judgements on further education colleges across the country. Therefore, the ability to compare and make judgements on most aspects under the theme of 'outcomes for learners' becomes limited, unexplained and wholly subjective, despite that fact that these judgements are used and presented as statistical fact.

The inspection report itself fails to offer the user or potential user of a college assurance that what they are reading can be compared to another college on a like for like basis. As identified above, the validated data on which judgements are made is unclear and the use of a college's own 'local' data for such aspects as attendance, value added, progression and

destination means that by its very definition, its comparability with other colleges is limited. And, regardless of what data is used, it is hardly ever cited in an inspection report, further limiting like for like comparisons.

It should also be noted that how inspectors report and what inspectors report on, varies from inspection to inspection. There are examples where evaluative statements and clear judgements are replaced by a descriptive commentary on the college's position, and there are other examples where one or more of the aspects being inspected under outcomes are not reported on at all.

Ofsted aims to provide fair, impartial, authoritative and comparable judgements about the quality of providers, to inform the choices made by parents, students, potential students and employers (Coffield 2017, p14; Ofsted 2014, p6), and it does so for 'outcomes for learners', clearly, consistently and transparently. The outcome of this my study indicates that inspectors are not consistent in the way that they use data and information to evaluate student performance accurately and that inspection reports lack sufficient detail to allow for any meaningful analysis.

Finally, whatever my study suggests, are these discrepancies sufficiently big enough to matter, especially when BIS (2013, p7) identified that only 9% of learners indicated that the provider's reputation was the main reason for choosing where they studied? Every school, college and training provider under Ofsted's purview should expect to be evaluated consistently, on a like for like basis and as Ofsted (2012, p6) declare 'without fear or favour', regardless of whether it impacts on a student's choice of where to study. To do otherwise is tantamount to anarchy.

6.2 Recommendations

The following set of recommendations aims to make, not only the process, but the reporting of inspection more consistent, reliable and transparent. If implemented, they have the potential to remove fear, subjectivity and perceived unfairness from this aspect of an inspection and to replace these with confidence in the ability of a college to accurately self-assess and, for those using an inspection report, to compare one provider to another with the assurance that they are comparing like with like. Moreover, if implemented, there is the

opportunity to re-distribute inspectors away from time validating and discussing data, back into the classroom to observe teaching and learning.

Some of these recommendations are easy to implement. Those which seek to collect and reproduce data at a national level will require further thought on how this might best be achieved.

6.2.1 Of the 165 college inspections used in this thesis, plus the twenty that fell outside the scope, there are no examples of where the success rate data is cited within the report to support the judgements made. And, as suggested in this study, validated success rate data, on which the judgements were founded, is inconsistent and has some wild variations. Therefore, to ensure complete transparency, where a judgement is made regarding achievement rates, the corresponding validated achievement rate should be cited. In addition, if inspectors include a statement(s) on the progress made over time for any particular cohort and/or level, then the related achievement rates should be cited.

Recommendation 1 It is suggested that in every inspection report the achievement rate (previously success rate) used by inspectors is included against each key judgement.

6.2.2 The relationship between the achievement rate and national rate for any student cohort, qualification or level should be made clear, so that all stakeholders, when reading an inspection report, have the ability to place the performance of students into a context that is comparable across the sector. Of the inspections analysed within this thesis, the majority refer to the national rate. However, a number make reference to other national rates such as 'similar colleges' or 'colleges with similar socio-economic background' where the national rate might be lower. Therefore, to ensure consistency, Ofsted, in line with its own guidance, should use the national average without applying any contextual factors and cite it within all inspection reports. Furthermore, where reference has been made to achievement rates over a period of time, the appropriate national rate should also be used. As a result, this will give a clear indication of the relationship between the national rate and the achievement rate in any one of the years cited within the report.

Recommendation 2 It is suggested therefore, that where a judgement on achievement rates is made, it includes the national rate (previously national average) and where appropriate, cites the difference in percentage point terms.

6.2.3 To further support Ofsted's aim of being transparent and consistent, inspectors should make clear where they are getting their achievement data from. Validated data, in the form of the QAR, is usually published in the January after the end of the previous academic year. On previous occasions, this has been delayed until March. For inspections that take place between September and January the only validated achievement rate data is held within the QAR from the previous year. There are no references to date sources or timeliness in any of the inspection reports. Furthermore, it is in these reports where the national rates are housed and again, neither the national rate itself nor the data source is included in the inspection reports (see Recommendation 2). Therefore, each report should clearly indicate the QAR being used by inspectors.

Recommendation 3 It is imperative, therefore, for inspectors to identify the dataset used (QAR) from which the achievement rates and the national rates were taken. This has the ability to assure that any future audit or quality assurance process assesses the reliability and validity of judgements made by inspectors, based on exactly the same data.

6.2.4 There are many examples in the inspection reports analysed for this thesis which indicate threshold values being applied to success rates. However, it has also been identified that, whilst these threshold values are indicated within the text of the report, normally in relation to the national average, the same does not necessarily apply to the accompanying data. Indeed, it has been suggested that, due to the inconsistent application of achievement and national rates by inspectors to inform their judgements, no indicative threshold exists.

As a result, such an identification of datasets not only gives confidence to colleges and key stakeholders that the use of achievement rate data is being applied objectively, with consistency and fairness, but it also has the ability to help colleges accurately self-assess student performance with the same level of objectivity. Ofsted should therefore set the criteria for data threshold values for achievement rates, with accompanying evaluative statements and judgements. These threshold values should identify how well (or not) all learners achieve, by age, duration and level, as well as other categories, when compared to the national rate.

By applying the achievement and national rate data to the inspection commentary and assuming that threshold values have been agreed, there should be a clear relationship between the achievement rate and national average for the judgements of 'requires improvement' and of 'inadequate'.

Recommendation 4 It is argued that it is important for colleges and for inspectors to have clear data threshold values for the measurement of achievement against the national rate for each judgement category. Not only does this have the ability to ensure consistency across inspections but it will also enable transparency.

6.2.5 Taking into account Ofsted's aim to be transparent and to 'provide users with information about the quality of provision in England' (Ofsted 2012, p5), there are a considerable number of inspection reports within this cohort which do not give a clear or unambiguous statement on student performance. Rather, there are examples of where commentary refers to the improvement or decline in success rates, absent of any starting point or relationship to the national rate, or indeed of any statement that a college's achievement rates are better (or worse) when compared to the previous inspection. This lack of clear judgements applies to all three student cohorts, namely 16-18, 19+, and WBL.

Recommendation 5a In order to improve consistency across inspections and comparability, it is suggested that all inspection reports include a clear summary judgement, using Ofsted's grading scale – one for each of the student groupings identified above, but also for the other areas within the current handbook (Ofsted, 2015), such as traineeships, provision for learners with high needs, and full-time provision for 14-16 year olds (where applicable).

Recommendation 5b It is argued that all key areas should be reported on and should include a judgement, accompanied by the supporting data (outlined above). A person reading one report looking for the standard of achievement for 16-18 year olds, for example, should be able to look at any other inspection report and know that a judgement for this age group will be there.

6.2.6 Un-validated, in-year data has the ability to give context about the direction of travel that the college, a specific student grouping or substantial qualification is apparently taking, especially if it is compared to the same point in the previous year. With particular regard to retention, it is clearly, however, only one part of the achievement rate, and for any

judgement to be made, it can only be done by inspectors either forecasting (or validating the college's forecasts) and/or assuming that the pass rate (the second component) will be the same if not better than it was in the previous year. There are a small number of inspection reports in this sample where the judgement made about the then success rates or student performance is evidenced by commentary on in-year data.

Similarly, throughout the academic year in every college there will be groups of students who have completed their qualification. For the majority, it will be because they have studied for a short or a very short qualification – some lasting no more than a day, and delivered over a term. Again, even when the achievement rates are known, they will only be indicative of the overall performance of the college. As a result, it is suggested that where this is the case, it should be commented on, citing the context within which this data sits, but with no judgement or evaluative statement attached. Furthermore, any judgement based on this sample data should not support a judgement of the whole college or cohort.

Recommendation 6 To avoid inspectors making judgements on student performance or indeed, anticipated outcomes, it is suggested that the use of un-validated achievement rate and in-year retention data, presented at inspection, should be commented on (if there is something remarkable about it) but not used to judge student performance.

6.2.7 If there is no guidance, no comprehensive collection of attendance data and therefore no national rate, it must be unclear to colleges how and against what they are measured. The same must surely apply to the inspection team. Furthermore, inspectors, in reality, have potentially three un-validated sources of data on which to form an opinion and make a judgement: attendance for the whole of the previous year prior to the inspection, in-year data up to the point of inspection, and the attendance that inspectors witness when undertaking observations and learning walks (assuming they have sight of the register).

It is known from the information provided by colleges that attendance data does not always relate to the judgement made. In addition, there are a number of examples from the wider cohort that the attendance witnessed by inspectors was different, always lower, than attendance data presented to them by the college. Therefore, to ensure consistency across the college sector, and indeed across the pre-16 and post-16 educational phases, especially as the participation age for staying on in education is now 18, the guidance on attendance as outlined in legislation (DFE, 2006) might be extended.

It is noted, however, that this is not necessarily an easy or speedy resolution. Therefore, it might be opportune for Ofsted to set out how it wants attendance data to be analysed and presented as part of the inspection process. Furthermore, it might be that it formally collects attendance data from classroom observations and learning walks. Subsequently, Ofsted could publish, as part of its regular data releases, national attendance rates for further education colleges (by cohort) and the overall national attendance rate as witnessed by inspectors.

Recommendation 7 It is suggested that, as happens for schools, national guidance be issued on the marking, collection and analysis of attendance data.

6.2.8 As a direct relation to the outcome of Recommendation 7, and to further support transparency and consistency, it would be beneficial to colleges and to inspectors to outline the threshold values for both attendance and punctuality for the purposes of self-assessment and inspection. These threshold values should be based on student attendance for the whole college and key cohort sub-sets, such as 16-18, 19+, and WBL, working from the national rate derived from the systems outlined in the previous recommendation, and applied to the Ofsted grading scale and associated judgements.

Additionally, when a judgement is made by inspectors regarding attendance (and it is suggested that this become a standard in every inspection) it should be accompanied by the attendance data – for the college as a whole and for each sub-cohort.

For any provider being inspected, it is important to know on what basis a judgement is being made.

Recommendation 8 It is suggested that inspectors identify threshold values for attendance and punctuality, against a national rate. And that, if, when analysed, this data identifies variation on the grounds of context for example, they issue guidance as to in what circumstance this might be considered during an inspection. As this is a recommendation which will require a longer lead-in time, in the interim it could be that inspectors collect, analyse and re-produce the attendance data taken from direct observation during an inspection for the purposes of such benchmarking.

The learner achievement tracker has been an online tool for measuring value added since 2005 (YPLA, 2010). Ofsted (2012a, p11) states that this tool and its successor will be used by inspectors to measure the progress students make compared to their starting points. The evidence now suggests that through the research period of this thesis, a number of value added systems were used, including those devised by individual colleges. Moreover, a considerable number of colleges were not using a value added system at the point of inspection.

Ofsted (ibid.) is clear about the value added tool that they will use during inspection, although there is evidence of deviation from their own guidance. Therefore, to promote consistency and comparability, it is suggested that inspectors and colleges use only one tool to measure the progress students make when compared to their starting points. Guidance on the use of this measure should include threshold values using Ofsted's evaluative criteria, again to assist with both the accuracy of self-assessment and consistency between inspections.

Finally, the progress that students make should be an essential and serious component when measuring 'outcomes for learners'. The majority of respondents in this study thought that hardly any weight was placed on value added by inspectors at their inspections and it might be therefore, taking in to consideration the number of colleges that did not have a value added system at the point of inspection, that colleges themselves do not give enough credit to this measurement of student performance.

Recommendation 9 It is suggested therefore, that a value added system with clear and unambiguous threshold values be given prominence by inspectors and colleges alike. And for it to be referenced within each inspection report, not only referenced by judgement but with its underpinning data.

6.2.10 Having validated and consistent data on the progress that students make should not be limited to those aged 16-18 on Level 3 programmes, but should include all levels for students aged 16-18 and those 19+ on full-time or substantive programmes.

There is little reference to any measurement system of student progress other than those cited in Recommendation 9, namely Value Added 16-18 year olds studying for a Level 3 qualification. Therefore, it is feasible that direct comparisons and judgements between one

college and any other might be variable – the extent of which variation is unknown and cannot be verified.

Ofsted (2012a, p11) acknowledges that measuring student progress for young people studying for qualifications at Level 2 or below, or for adults (who may enter the further education system with a range of qualifications and experiences), is more complex. However, there is an expectation that colleges have systems in place for both student groups. Therefore, it is not inconceivable that a national system be devised for these more complex groups; if every individual college is expected to devise its own system, there must be examples which could be used to provide a solution.

To have such a system (or series of systems), it is argued, will not only create a dataset on which expectations for both inspection and self-assessment can be measured, but it should create a dataset from which an individual's academic development is not only identified, but accelerated and extended to beyond the average (where appropriate).

Recommendation 10 Therefore, instead of having hundreds of individual systems that assess the progress made by students - for young people, adults and those on apprenticeships - it is suggested that one system be created (or one system per student cohort) and that the standardised data which this will generate is used by inspectors consistently to evaluate the progress made by students.

6.2.11 Having an understanding of how many students progress throughout the levels within the same institution and/or how many progress to destinations outside a college could be a really important factor for parents and potential students choosing to attend one college over another. Whilst achievement rates will identify the proportion of students who stay on at the college and pass their qualification, progression and destination identify how the college, along with academic achievement, prepares a young person or adult for their next step – further learning, higher education or work. Therefore, the judgements made in these two areas should be unambiguous so as to be informative and the data on which they are founded should be presented within the inspection report.

Moreover, to ensure consistency, avoid contradiction, and establish greater synergy between a college's self-assessment and the findings of the inspection, guidance by the inspectorate could be published on how progression and destination data will be used in inspection (and

should be used for self-assessment) when evaluating actual or predicted data or indeed if only one dataset should be used.

In addition, when making judgements on both progression and the destinations of students, in addition to publishing the underpinning data, guidance should be produced on how such data equates to grading and evaluative statements.

Furthermore, in a number of inspection reports, inspectors have commented that too much data is unknown on the progression and destination of students –with reference sometimes to the whole college and sometimes to a specific cohort. It is reasonable to suggest therefore, that if there is a lack of data on which inspectors and indeed the college as part of its self-assessment process can substantiate a judgement, then that sample size should be known and published. And, it might also be reasonable, in order to ensure consistency, that no judgement or automatic judgement of 'inadequate' be made where the number of unknown outcomes for both progression and destination fails to reach an agreed threshold.

Recommendation 11 It is suggested that the data used by the inspection team to underpin their judgements on both the progression and destination of students is included in each inspection report. Each inspection report should also indicate whether actual or predicted data is used and it should identify what proportion of the student body has a known outcome for both progression and destination. Whatever the threshold, whether it is reached or not, the sample size on which judgements are founded should be cited within the inspection report.

6.3 Contribution to research

The further education sector in England, and indeed the rest of the UK, is limited when compared to the schools and the higher education sectors. In addition, there are no large scale studies which attempt to analyse objectively how and with what consistency Ofsted inspectors use data and information to evaluate student performance and make judgements on 'outcomes for learners'. Therefore, I believe that my study has the ability to make a contribution to the current, limited, literature on the further education sector, to inspections within the further education sector, and to Ofsted more widely in the realms of social policy and educational administration.

Much of the literature in Chapter 2 is based on case study research in either one school or college or across a very small cohort. My study is probably one of very few, if not the only one, to consider an entire cohort of colleges, across one complete national inspection cycle. It has not sought to interpret findings from a single college, fraught with context and individual experience, only then to interpret these findings as being attributable to the sector as a whole. Instead, the fact that my study analyses nationally validated data, published inspection reports and data provided by colleges across an entire inspection cycle potentially justifies the content of this thesis as important, and useful to the research community.

Indeed, whilst this level of analysis on the further education sector and/or Ofsted might not have been carried out before, it is also possible that the data collected directly from colleges is another first. And, whilst this data is taken in good faith and not subject to any external validation, as the first collection and analysis of such data, it provides a platform and a basis for future research.

I have identified a number of areas for future research below. The overarching aims of these studies is to contribute further new knowledge, understanding and analysis of the further education sector, how it is inspected, inspection more generally and to help raise the profile of this under-researched sector.

6.4 Limitations of the research

Having fully considered the data and information used in this thesis, I would suggest that there are two areas which, whilst not affecting the overall findings, conclusions and recommendations, could result in a stronger, more reliable evidence base for future research. The two areas to be considered are:

- The timeliness and reliability of primary data; and,
- The response rate from colleges.

First, my study concerns itself with inspections that took place between September 2012 and July 2015, covering one complete inspection cycle. The primary data sought from colleges via a freedom of information request, was sent out by email at the beginning of October 2016. As a result, I was asking colleges to provide me with data that was between 1.5 and 5 years old. Whilst I have no evidence to suggest that this data is anything other than what

was presented to inspectors, and therefore reliable, I am open to the possibility that there may be cases, especially from colleges at the beginning of the cycle, where data was re-produced, for the single purpose of responding to the FoI request.

Second, and related to the above point, as a result of the time passed between a college's inspection and my request for information, a number of colleges stated that they were not able to respond to the FoI request – either because they didn't keep copies of the data and information that they provided to the inspection team, or because the key people, such as the nominee and other members of the senior management team, who would have been directly involved, no longer worked for the college. This has, therefore, limited the response rate not only to the FoI request more generally to around 50%, but also to the individual questions, which when analysed by year and judgement can result in a small sample size.

With regard to the future areas of research detailed below, especially in the proposed study similar to this one using the latest handbook (Ofsted, 2015), and covering the current inspection cycle, there will still be a delay between the timing of an inspection and the request for data – although this may be reduced to no more than 3.5 years. Moreover, if this thesis and the proposed study outlined below form the basis of a longitudinal study, it may be possible to gather almost instant data – relying on colleges to submit the data used at inspection in real time (or within days rather than years). This also has the ability to increase the response rate – either through a voluntary agreement, possibly through the Association of Colleges, Education and Training Foundation or similar - or again, through a freedom of information request to each college at the point of their inspection report being published.

6.5 Identification of future research areas

6.5.1 How are data and information used by inspectors to evaluate student performance and make judgements on outcomes for learners?

In August 2015, Ofsted launched the further education and skills inspection handbook (Ofsted, 2015), which, much like the handbook used as a framework for this thesis (Ofsted, 2012) sets out guidance on how colleges and other skills providers will be inspected, and is a guide for inspectors.

However, there are some considerable changes from the 2012 handbook. The inspection and reporting on Subject Sector Areas (SSAs) no longer take place. Instead, the findings by curriculum inspectors and other inspectors contribute to the four overall judgement areas in addition to the types of provision inspected. They are:

- 16-18 study programmes;
- Apprenticeships;
- Adult learning programmes;
- Traineeships;
- Learners with high needs; and,
- Full-time provision for 14-16 year olds.

In addition, as a result of the introduction of the 16-18 Study Programmes in 2013 (DfE, 2012) and the new qualification categories for both 16-18 year olds and adults (the latter replacing the former duration measure), the evaluation and judgements on 'outcomes for learners' are arguably more complex. This additional level of complexity gives further reason for inspectors and inspection reports to be consistent and transparent, so that all stakeholders can compare one college with another. It should be noted that whilst consistency and transparency were clear principles of inspection, cited in the previous handbook (Ofsted 2012, p6), this section did not make it into the new handbook. However, as part of his forward to the introduction of the study programmes (DfE 2012, p2), the then Secretary of State for Education stated that there would be 'far-reaching reforms in the way that post-16 education is...monitored and reported' and that 'anyone will be able to judge whether attending a particular...college...is likely to offer them the best opportunity to progress'.

Proposed aims and objectives:

- To critically analyse to what extent validated achievement rate data is used to evaluate student performance and to make judgements;
- To identify and analyse what progression and destination data is being used by inspectors to inform their finding and judgements;
- To identify to what extent there are any inconsistencies in the application of the CIF and if so, with what potential impact; and,

- To make recommendations to further support the accuracy, consistency and transparency in judging 'outcomes for learners'

6.5.2 Short inspections: what relationship is there between student performance and the grade awarded to the college?

Another change to the inspection system for further education and skills, as part of the implementation of the handbook (Ofsted, 2015), was the introduction of short inspections. Short inspections are available to colleges who achieved a good overall effectiveness grade at their most recent inspection and the inspection would take place within three years of the beginning of this inspection cycle.

According to the guidance on short inspections (Ofsted 2015, p28), they can take place over one or two days by one or more inspector. At the end of a short inspection there are three outcomes. That:

- 1) the provider continues to be a good provider:
- 2) the provider remains good and there is sufficient evidence of improved performance to suggest that the provider may be judged outstanding; or,
- 3) the inspection team have insufficient evidence to satisfy themselves that the provider remains good or there are concerns.

With outcomes 2 or 3 the short inspection will be converted to a full inspection.

Most colleges of further and higher education are large institutions, with thousands of students and a financial turnover in the millions. As a result of some recent (and not so recent) mergers, many colleges host tens of thousands of students studying for an array of qualifications across multiple campuses. The nature of the organisation may be complex and the ability for a very small team of inspectors over a very short period of time to evaluate truly the overall effectiveness of such an institution must be limited, unless student performance in the form of 'outcomes for learners' is at the heart of this process. Indeed, in the section titled 'Inspectors' planning and preparation' (Ofsted 2015, p30), the only evidence available to inspectors which might give an indication on a college's current position is the performance data.

Proposed aims and objectives:

- To critically analyse to what extent validated achievement rate data is used to evaluate student performance and to make judgements in short inspections;
- To analyse the inspection letters and identify what other data and information are being used to evaluate student performance;
- To understand and analyse what, if any, relationship exists between the confirmation of the provider continuing to be judged as a good provider and those which have resulted in a conversion to a full inspection; and,
- To identify to what extent there are any inconsistencies in the application of the CIF with regard to short inspections and if there are any, to identify their potential impact; and,
- To make recommendations to further support the accuracy, consistency and transparency in judging student performance in short inspections.

6.5.3 How important are English and mathematics achievement rates when evaluating student performance in further education colleges?

At the heart of the study programme reforms for students aged 16-18 are the mandatory elements of English and mathematics. Colleges of further education have, for a long while, offered English and mathematics qualifications alongside mainly vocational programmes. However, these qualifications, known throughout their respective tenure as core skills, common skills, key skills and now functional skills, were not mandatory; not all students were enrolled, and there is some suggestion that some unqualifiable number was enrolled onto these qualifications at levels lower than their previous attainment or considerably lower than the level of the vocational qualification studied.

Since the implementation of the study programme, it is now expected that all young people and some adults 'where it is valued or set as a requirement by employers' (Ofsted 2012, p9) should study English and mathematics until they achieve a GCSE grade A*-C (now grade 9-4) in both subjects (DfE 2012, p10). Indeed, it is a condition of funding. Other English and mathematics qualifications such as functional skills are to be used as stepping stones to the achievement of a GCSE, apart from provision for students with learning difficulties or disabilities.

As a result of the textual analysis undertaken for this thesis, there is some evidence that suggests that student performance in English and mathematics, on both functional skills and on GCSE programmes, is variable. Where inspectors make reference to functional skills achievement rates, in the majority of reports they are judged either to be in line with the national average or below it across both age groups. For 16-18 year olds studying for a GCSE in English and/or mathematics, 70 out of 117 inspection reports report that GCSE achievement rates are below or considerably below the national rate.

Proposed aims and objectives:

- To critically analyse to what extent validated achievement rate data is used to evaluate student performance and to make judgements in functional skills and GCSE English and mathematics;
- To analyse the judgements, commentary and outcomes for functional skills and GCSE English and mathematics and how they contribute to the overall grade of 'outcomes for learners';
- To identify whether there is evidence to suggest that achievement rates and their associated judgements in English and mathematics act as de facto limiting grades; and,
- To identify to what extent there are any inconsistencies in the application of the CIF and if there are any, to assess their potential impact; and,
- To make recommendations to further support the accuracy, consistency and transparency in judging 'outcomes for learners'

6.5.4 What is the relationship between student performance and the timing of an inspection in the further education sector?

The handbooks for the inspection of colleges (Ofsted 2009, 2012 & 2015) identify that colleges (and other providers within its scope) are selected for inspection based on a risk-proportionate basis. In the latest handbook (Ofsted 2015, p5), Ofsted 'uses a broad range of indicators to selected providers for inspection' and will do so in such a way that it 'can focus its efforts where it can have the greatest impact' (ibid.).

In the handbooks (Ofsted 2012 & 2015), there is a list of indicators which include:

- Self-assessment reports
- Performance data
- Destination data
- Change of leader

The other categories include concerns raised by funding bodies and others, in addition to the views of learners, parents/carers and employers gathered through on-line questionnaires. However, Camden (2017) in an article for FE Week, cited that the take-up by students and employers for these questionnaires was extremely low – less than 2% for students and, out of a sample of 10 inspections, fewer than half employers responding.

Therefore, it is conceivable that student performance data, destination data and change of leader, form the central aspects of the risk assessment process.

Having analysed success/achievement rates and progression and destination data across an inspection cycle, there is some indication that variation (and in some cases, considerable variation) in student performance in colleges with and without a change of leader has not resulted in an inspection. Indeed, there are a small number of cases where a small number of colleges have not been inspected for upwards of six years, despite a change of leader and/or varying success.

Based on these cases, it is worth analysing how Ofsted used data and information to select providers for inspection – with and without a change of leader.

Aims and objectives:

- To identify a large sample of colleges from each Ofsted region, critically analysing student performance data for considerable variations – decreases and sporadic performance – from 2012 to today;
- To investigate and analyse the relationship between student performance and the timing of an inspection;
- To identify and analyse the relationship (if any) between the timing of an inspection and a change of leader;

- To identify to what extent there are any inconsistencies in the application of the CIF with regards to the selection of providers for inspection and if there are any, what is their potential impact; and,
- To make recommendations to further support the accuracy, consistency and transparency when risk-assessing a college for inspection.

6.5.5 Do students who attend better, achieve better?

My study has highlighted that measuring attendance across the further education sector is complex and in reality, incomparable – due to the lack of a consistent framework or national guidance against which attendance can be recorded and analysed. Furthermore, Ofsted (2013, p15) states, in its small scale study, that ‘few colleges carry out any specific analysis of attendance to compare it with success rates or the quality of teaching’. There are no reliable studies identified by Ofsted in their study which analyse the potential relationship between attendance and achievement or indeed the level of that achievement. And, knowing, through a small number of inspection reports that indicate, albeit tentatively, that attendance may be impacted by its socio-economic context, does it actually affect achievement and value added measures?

Proposed aims and objectives:

- To analyse if a student’s attendance at college has a causal relationship with their achievement;
- To analyse if a student’s attendance at college has a causal relationship to how well they achieve, in relation to their starting point. Do they make the progress expected of them?
- To identify whether a differential relationship exists between a student’s attendance and their achievement and progress depending on the socio-economic context within which they live and study;
- To identify two colleges from each Ofsted inspection area: one in an area of low socio-economic deprivation and the other in an area of high socio-economic deprivation (18 colleges in total);
- To identify a cohort of students from each college, following the same qualification at the same level as the foundation of this study; and,
- To make a valid contribution to research.

So what might the 'outcomes for learners' section in an inspection report look like if the recommendations suggested above are implemented? The commentary used is extracted from several inspection reports used in my study, although in places it is edited to ensure that it cannot be traced back to its original source. And, whilst reference is made to the national averages taken from QSR 11/12, references to success rates are invented by way of illustration. In addition, the example below does not consider the reporting of all aspects under 'outcomes for learners'; it addresses only those that have been analysed as part of this study. It should also be noted that as part of the handbook (Ofsted, 2015), attendance is reported under the theme of personal development, behaviour and welfare.

6.5.6 Outcomes for learners – an example

Over the past three years achievement rates for students aged 16-18 on long courses have risen considerably, and at 87%, are now above the national rate by 5%. Historically, the proportion of adults gaining qualifications has been low. However, over the past academic year, managers have implemented an effective action plan to bring about improvements. As a result, achievement rates for students aged 19+ on long qualifications are now at a rate of 80.1%, against a national rate of 79.7%.

The proportion of apprentices who successfully complete their apprenticeship programme has been considerably below the national rate over time. Overall achievement rates for apprentices remain poor at 55.2%, which is 18.6% below the national rate.

Attendance rates require improvement. Whilst attendance is improving for full-time students aged 16-18, where it is now, at 84% and whilst it is above the average attendance rate for inner-city colleges by 2.1%, it remains stubbornly below the average attendance rate for those aged 19+, by -6.4%.

The progress made by students when compared to their starting points is variable. For those aged 16-18 on level 3 programmes, a high proportion exceed what is expected of them and 67% achieve high grades. For young students aged 16-18 on courses at level 2 and below and for adults, the college does not use the qualification on entry data to inform its teaching, learning and assessment strategies. As a result, only 33% of young people and 42% of

adults achieve the outcomes expected of them. Both are below the national rate, by 17% and 13% respectively.

Nearly all students progress to positive destinations from full-time programmes on which there are 2,300 young people and 650 adults. For young people, around 50% progress to higher level qualifications within the college, 25% find employment, and 20% progress to higher education. The destinations of 5% of this cohort are unknown. For students aged 19+, 20% progress to a higher-level qualification within the College, 60% progress to higher education and 10% enter employment. The destinations of 10% of this cohort are known. This information was provided by the college and brings together data from UCAS, MIS and an 'after the course' survey.

The number of apprentices who are retained by their employer on completion of their programme has improved and is now good. Of the 1,000 apprentices completing in the last academic year, 72% have been retained by their employer, whilst 10% have moved to another employer in the same sector.

6.6 Personal Learning

The process of research, analysis and the writing of this thesis has been exhilarating, frustrating and rewarding in equal measure. There have been times when the mere thought of completing this study stopped me in my tracks – especially when I wrote the first word on a blank piece of paper and immediately thought about how I would ever get to the end. So, at this point of reflection on the process I offer the following:

The whole is the sum of its parts. At the beginning of the process, especially during the research and writing of the literature review, I become quite focused on the entirety of the thesis, which became quite an overwhelming feeling, to the point that it started to frustrate my writing. To that end, I started to approach each chapter as a study in its own right and within chapters, smaller essays, answering specific questions. Once I established this as a pattern, the whole process became much less daunting, more manageable and quicker.

Just write something or walk away (briefly). There were several occasions through the writing of this thesis, where I could easily spend hours just staring at the screen or worse, writing and deleting the same sentence over and over and ending the day with fewer words

than I started with. When this happened, and to feel productive, I would move away from my intended task to work on something more mundane, such as organising the bibliography or inputting data. Failing that I walked away to gather my thoughts, returning once the first line had clicked in my head, with the rest of the section following as a result.

I love SPSS! The only experience I had, prior to my study, of analysing and reporting on statistical inspection was as a Vice Principal, preparing reports for senior management team and governors meetings. Inevitably, the data would be produced by the MIS team. So, having to collect, collate, code, input and analyse my own data was a huge learning curve, but one that I enjoyed. And, as someone new to SPSS, I asked for help from one of my supervisors who was extremely helpful and who spent some time with me taking me through it and how to use it. Whilst the books, even the most popular or funniest can be insightful, I found them difficult to penetrate.

Clean the data as you go along. The dataset that I have amassed as part of my study, now holds just under 16,000 pieces of data. With this amount of data there were inevitably going to be some inputting errors. Once the first tranche of data was input (inspection report data) I asked two colleagues to pick out a selection of the physical inspection report to play 'inspection report bingo', enabling me to cross-check the information that they called out to what I had input into SPSS. However, with a small number of the other categories, I failed to check for accuracy during the creation stage, so that when it came to the analysis stage the numbers were out and the painful process of trying to find the offending data began. From that moment on I completed a frequencies analysis to ensure that the number (of colleges) within the sample matched.

7.0 Bibliography

Altrichter, H. & Kemethofer, D. (2015) Does accountability pressure through school inspections promote school improvement? *School Effectiveness and School Improvement*. Vol 26, (1) 32-56

Andrews, D. (2008) *Issue paper 5: Guidance and Careers Education*. Oxford: University of Oxford

Association of Colleges. (2015) *College Key Facts: 2015/16*. London: Association of Colleges

Association of Colleges. (2015a) History of further education [on line]. Available from <http://www.aocjobs.com/article/history-of-further-education-to-90s/> [Accessed 22 March 2016]

Associate of Colleges. (2015c) AoC warns of the end of adult education and training by 2020 [on line]. Available from <https://www.aoc.co.uk/news/aoc-warns-the-end-adult-education-and-training-provision-2020> [Accessed 22 September 2016]

Avis, James (2009) Further education: policy hysteria, competitiveness and performativity. *British Journal of Sociology of Education*. Vol 30, (5) 653-662

Bailey, B. & Unwin, L. (2014) Continuity and change in English further education: A century of voluntarism and permissive adaptability. *British Journal or Educational Studies* Vol 62, (4) 449-464

Ball, S. (1997) Good school/bad school: paradox and fabrication. *British Journal of the Sociology of Education*. Vol 18, (3) 317-316

Ball, S. (1998) Performativities and fabrications in the education economy: towards the performative society. Carter, J. (Ed) *Post modernity and the fragmentation of welfare*. Florence: Routledge

Ball, S. (2008) *The Education Debate*. Bristol: Policy Press

Baxter, J. (2014) An independent inspectorate? Addressing the paradoxes of educational inspection in 2013. *School Leadership & Management*. Vol 34, (1) 21-38

Baxter, J. (2013) Professional inspector of inspecting professional? Teachers as inspectors in a new regulatory regime for education in England. *Cambridge Journal of Education*. 43, (4) 467-485

Baxter, J. & Clarke, J. (2013) Farewell to the tick box inspector? Ofsted and the changing regime of school inspection in England. *Oxford Review of Education*. Vol 39, (5) 702-718

Baxter, J. & Clarke, J. (2014) Knowledge, authority and judgement: the changing practices of school inspection in England. *Journal of Education*. Vol 2, (1) 106-127

Behnke, K. & Steins, G. (2017) Principals' reactions to feedback received by school inspection: a longitudinal study. *Journal of educational change*. Vol 18, (2017) 77-106

- Bell, J. (2003) *Doing your research project: a guide for first-time researchers in education and social science (3rd Ed)*. Berkshire: Open University Press
- BBC. (2011) *EMA: student numbers fall blamed on allowance cut [online]* BBC Available from: <http://www.bbc.co.uk/news/education-15273410> [Accessed 10 June 2016]
- Birdwell, J., Grist, M. & Margo, J. (2011) *the forgotten half: a demos and private equity foundation report*. London: DEMOS
- BIS. (2013) *The impact of further education learning*. London: BIS
- BIS. & DFE. (2016) *Reviewing post-16 education and training: area reviews (waves 1, 2 and 3)*. London: BIS
- Black. T. (1993) *Evaluating social science research: an introduction*. London: Sage Publications
- Blaxter, L., Hughes, C. & Tight, M. (1996) *How to research*. Buckingham: Open University Press
- BoE (1926) *Report of the consultative committee on the education of the adolescent*. London: Board of Education
- Boocock, A. (2014) Increased success rates in a FE college; the product of a rational or a performative college culture? *Journal of Education and Work*, 27 (4), 351-371
- Brimblecombe, N., Shaw, M. & Ormston, M. (1995) Teachers' perceptions of school inspection: a stressful experience. Vol 25, (1) 1-9
- Burke, J. (2016) 'Inadequate' colleges double-digit disaster as English and mathematics policy blamed [online]. Available from: <http://feweek.co.uk/2016/06/16/inadequate-colleges-double-digit-disaster/> [Accessed 17 June 2016]
- Camden, B. (2017) *Ofsted survey tools used by only 2% of learners [online]*. Available from: <http://feweek.co.uk/2017/07/03/ofsted-survey-tools-used-by-only-2-of-learners/> [Accessed 29 August 2017]
- Campaign for Learning. (2006) *Overcoming the academic/vocational divide pre-19*. London: Campaign for Learning
- Campbell, J. & Husbands, C. (2000) On the reliability of Ofsted inspection of initial teacher training: a case study. *British Educational Research Journal*. Vol 26, (1) 39-48
- Carlbaum, S. (2016) Equivalence and performance gaps in Swedish school inspection: context and the politics of blame. *Studies in the cultural politics of education*. Vol 37, (1) 133-148
- Carley, K. (1993) Coding choices for textual analysis: a comparison of content analysis and map analysis. *Sociological Methodology*. Vol 23, (1993) 75-126
- CASE. (2015) *The coalitions' record on further education, skills and access to higher education 2010-2015*. London: London School of Economics

Case, P., Case, S. & Catling, S. (2000) Please show you're working: a critical assessment of the impact of Ofsted inspection on primary teachers. *British Journal of Sociology of Education*. Vol 21, (4) 605-621

Camden, B. *Three month inspection report delay nearly over, Ofsted claims [online]*. Available from: <https://feweek.co.uk/2017/08/04/three-month-inspection-report-delay-nearly-over-ofsted-claims/> [Accessed 29 August 2017]

Cannicott, S., Kempton, J. & McDaniel, S. (2013) *Smarter accountability in further education*. London: Centre Forum

Chapman, C. (2002) Ofsted and school improvement: teachers' perception of the inspection process in schools facing challenging circumstances. *School Leadership & Management*, 22 (3), 257-272

City and Guilds. (2015) *Global economic investment in vocational education and training*. London: City and Guilds

Clapham, A. (2015) Post-fabrication and putting on a show: examining the impact of short notice inspections. *British Educational Research Journal*. Vol 41, (4) 613-628

Clarke, S. & Cossette, S. (2000) Secondary analysis: theoretical, methodological, and practical considerations. *Canadian Journal of Nursing Research*. Vol 32, (3) 109 - 129

Cohen, L., Manion, L., & Morrison, K. (2007) *Research methods in education*. Oxford: Routledge

Cockett, M. (1996) Chaos or coherence, progression and continuity. Halsall, R. & Cockett, M. (Ed.) *Education and training 14-19: chaos or coherence*. London: David Fulton Publishers

Coffield, F. (2012) *Ofsted re-inspected*. Cardiff: NIACE

Coffield, F. (2017) *Will the leopard change its spots? A new model of inspection for Ofsted*. London: UCL IoE Press

Coffield, F., Edward, S., Finaly, I., Hodgson, A., Spours, K., & Steer, R., (2008) *Improving learning, skills and inclusion: The impact of policy on post-compulsory education*. London: Routledge

Courtney, S. (2014) Post-panopticism and school inspection in England. *British Journal of Sociology of Education*. Vol 37, (4) 623-642

Courtney, S. (2012) *Ofsted's revised school inspection framework: experiences and implications*. Paper presented at the BERA conference. Manchester, 2012

Courtney, S. (2013) Head teachers' experiences of school inspection under Ofsted's January 2012 framework. *Management in Education*. Vol 27, (4) 164-169

Crehan, L. (2016) *Cleverlands* London: Unbound

- Cross-Durrant A. & Singleton D., (2003) *The impact of re-inspection monitoring in colleges*. Manchester: Ofsted
- Cuckle, P., Hodgson, J. & Broadbent, P. (1998) Investigating the relationship between Ofsted inspections and school development planning. *School leadership and management*. Vol 18, (2) 271-283
- DCSF. (2007) *Children's Plan 14-19 Expert Group Report*. Nottingham: DCSF Publications
- DCSF. (2008) *Delivering 14-19 reform: next steps*. Nottingham: DSCF Publications
- DCSF. (2008a) *Promoting achievement, valuing success: a strategy for 14-19 qualifications*. Norwich: TSO
- DCSF. (2008b) *Delivering 14-19 reforms in rural areas: final report* Nottingham: DCSF Publications.
- DCSF. (2008d) *Post 16 participation rate at highest rate ever [on line]*. Available from: http://www.dcsf.gov.uk/pns/DisplayPN.cgi?pn_id=2008_0119 [Accessed 28 April 2016]
- DCSF. & DIUS. (2008) *Raising expectations: enabling the system to deliver*. Norwich: TSO.
- de Wolf, I & Janssens, F (2007) Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*. Vol 33, (3) 379-396
- DfE. (2013) *A guide to level 3 value added in 2013 in school and college performance tables [online]*. Available from :DfE available from: http://www.education.gov.uk/schools/performance/2013/16to18_13/A_Guide_to_Level_3_Value_Added.pdf [Accessed 21 March 2016]
- DfE. (2014) *Participation of young people in education, employment or training*. London: DfE
- DfE. (2012) *Study programmes for 16 to 19 year olds: government response to consultation and plans for implementation*. London: DfE
- DfE. (2016) *Education Excellence Everywhere*. Nottingham: DfE
- DfEE. (1999) *Learning to succeed: a new framework for post-16 learning*. London: DfE
- DfEE. (1998) *Schools standards and framework act*. London: HMSO
- DfES. (2004) *14-19 Curriculum and Qualifications Reform: final report of the working group on 14-19 reform*. Nottingham: DFES Publications
- DfES. (2005) *14-19 Education and Skills*. Norwich: HMSO
- DfES. (2005a) *14-19 year olds in our schools: building on success*. London: HMSO
- DfES. (2002) *14-19: extending opportunities, raising standards*. London: HMSO

- DfES. (2002a) *Success for All*. London: DfES
- DfES/LSC. (2007) *Delivering world-class skills in a demand-led system*. Coventry: LSC
- DfES. (2006) *Personalising further education: developing a vision*. Nottingham: DfES
- DFES. (1980) *The Education Act*. London: HMSO
- DFES. (1988) *The Education Reform Act*. London: HMSO
- DFES. (2006) *Education and Inspections Act*. London: HMSO
- DFEE. (2000) *Learning and Skill Act*. London: HMSO
- Doel, M. (2018) *Defining further education: does it matter? [on line]*. Paper presented at the University College London, Institute for Education, London, February 2018
- Ehren, M., Perryman, J. & Spours, K. (2014) *Accountability and school inspections*. London: IoE
- Ehern, M., Gustafsson, J., Altrichter, H., Skedsmo, G., Kemethofer, D. & Huber, S. (2015) Comparing effect and side effects on different school inspections systems across Europe. *Comparative Education*. Vol 51, (3) 375-400
- Ehern, M. & Swanborn, M (2012) Strategic data use of schools in accountability systems. *School Effectiveness and School Improvement* Vol 23, (2) 257-280
- Ehern, M. & Visscher, A. (2008) The relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*. Vol 56, (2) 205-227
- Elliot, A. (2012) *Twenty years inspecting English schools – Ofsted 1992-2012*. London: RISE
- Elmore, R & Fuhrman, S. (2001) Holding schools accountable: is it working? *Phi Delta Kappan*. Vol 83, (1) 67-70
- Fielding, M. (2001) Ofsted, inspection and the betrayal of democracy. *Journal of philosophy of education*. Vol 35, (4) 695-709
- Fidler, B., Earley, P., Ouston, J. & Davies, J. (1998) Teacher gradings and Ofsted inspection: help or hindrance as a management tool. *School leadership and management* Vol 18, (2) 257-270
- Fletcher, M., Gravatt, J. & Sherlock, D. (2015) Levers of power: funding, inspection, and performance management. Hodgson, A. (Ed) *The coming of age for FE*. London: Institute of Education Press
- Fletcher, M. (2015) *Inspection and FE colleges: an AoC discussion paper*. London: Association of Colleges
- Forrest, C. (2015) Inspection and improvement in three further education colleges. *Research in Post-Compulsory Education*, 20 (3), 296-314

Foster, A. (2005) *Realising the potential: a review of the future role of further education*. Nottinghamshire: DfES

Freshminds. (2014) *Careers guidance: guaranteed*. London: AoC/Find A Future

Gillard, D. (2011) *Education in England: a brief history [on line]*. Available from: www.educationinengland.org.uk/history [Accessed 10 February 2018]

Gilleece, L. (2014) Understanding achievement differences between schools in Ireland – can existing data-sets help? *Irish Educational Studies, Vol 33 (1) 75-98*

Gleeson, D. (1996) Continuity and change in post-compulsory education and training. Halsall, R. & Cockett, M. (Ed.) *Education and training 14-19: chaos or coherence*. London: David Fulton Publishers

Goes, J. & Simon, M. (nd) *Using secondary data in doctoral research [on line]*. Available from: www.dissertationrecipies.com [Accessed 10 August 2017]

Gorard, S. (2013) *Research design: creating robust approaches for the social sciences*. London: Sage

Gorard, S. & See, B. (2013) *Overcoming disadvantage in education*. Oxford: Routledge

Government. (2015) *2015 headline performance at key stage 4 – all pupils [on line]*. Available from: <https://www.compare-school-performance.service.gov.uk/schools-by-type?step=phase&geographic=all®ion=0&phase=secondary> [Accessed 24 June 2016]

Government. (2016) *Apprenticeships Overall Headline Success Rates [on line]*. Available from: <https://www.gov.uk/government/statistical-data-sets/sfa-national-success-rates-tables-2013-to-2014> [Accessed 5 September 2016]

Government. (2016) *Apprenticeships Overall Institution Success Rates [on line]*. Available from: <https://www.gov.uk/government/statistical-data-sets/sfa-national-success-rates-tables-2013-to-2014> [Accessed 5 September 2016]

Government. (2016) *Education and Training Overall Headline Success Rate Tables Data 201314 [on line]*. Available from: <https://www.gov.uk/government/statistical-data-sets/sfa-national-success-rates-tables-2013-to-2014> [Accessed 5 September 2016]

Government. (2016) *Education and Training Institution Success Rate Tables Data 201314 [on line]*. Available from: <https://www.gov.uk/government/statistical-data-sets/sfa-national-success-rates-tables-2013-to-2014> [Accessed 5 September 2016]

Government. (2016) *Institute Success Rate Tables Data 201213 [on line]*. Available from: <https://www.gov.uk/government/statistical-data-sets/sfa-national-success-rates-tables-2013-to-2014> [Accessed 5 September 2016]

Government. (2016) *Overall Success Rate Table Data 201213 [on line]*. Available from: <https://www.gov.uk/government/statistical-data-sets/sfa-national-success-rates-tables-2013-to-2014> [Accessed 5 September 2016]

Gravatt, J. (2014) What's gone wrong with the Cinderella sector? *[online]*. Available from: <https://www.theguardian.com/education/2004/may/19/furthereducation.uk2> [Accessed 27 May 2018]

Gray, C. & Gardner, J. (1999). The impact of school inspections. *Oxford Review of Education*. Vol 25, (4) 455 – 468

Grek, S., Lawn, M., Ozga, J. & Segerholm, C. (2013) Governing by inspection? European inspectorates and the creation of a European education policy space. *Comparative Education*. Vol 49, (4) 486–502

Greasley, P. (2008) *Quantitative data analysis using SPSS: an introduction for health and social sciences*. Maidenhead: McGraw Hill/Open University Press

Gustafsson, J., Ehern, M., Conyngham, G., McNamara, G. & Altrichter, H. (2015) From inspection to quality: ways in which school inspection influences change in schools. *Studies in Educational Evaluation*. Vol 47, (2015) 47-57

Harvey, L. (unknown) *Value added [online]* Analytical Quality Glossary available from <http://www.qualityresearchinternational.com/glossary/valueadded.htm> [Accessed 21 March 2016]

Hall, J. (2017) Examining school inspectors and education directors within the organisation of school inspection policy: perceptions and views. *Scandinavian journal of educational research*. Vol 61, (1) 112-126

Hasall, R. (1996) Setting the scene. Halsall, R. & Cockett, M. (Ed.) *Education and training 14-19: chaos or coherence*. London: David Fulton Publishers.

Hatton, P. (2017) *Ofsted reports require improvement [online]*. Available from: <https://feweeek.co.uk/2017/01/23/ofsted-reports-require-improvement/> [Accessed 28 April 2017]

Heaton, J. (2008) Secondary analysis qualitative data: an overview. *Historical Social Research*. Vol 33, (3) 33-45

Higham, J. & Yeomans, D. (2007) Policy memory and policy amnesia in 14-19 education: learning from the past? Raff, D. & Spours, K. (Ed) *Policy-making and policy learning in 14-19 education*. London: Bedford Way Papers

Hodgson, A., Keep, E. & Spours, K. (2007) *Issue paper 1: the new 14-19 diploma*. Oxford: University of Oxford.

Hodgson, A. & Spours, K. (2003) *Beyond A levels: curriculum 2000 and the reform of 14-19 qualifications*. London: Kogan Page

Hodgson, A. & Spours, K. (1997) 1991 white paper to the Dearing report. Hodgson A. & Spours, K. (Ed.). *Dearing and beyond: 14-19 qualifications, frameworks and systems*. London: Kogan Page.

- Hodgson, A. & Spours, K. (2008) *Issue paper 10: General Education in the 14-19 phase*. Oxford: University of Oxford
- Hodgson, A. & Spours, K. (2008a) *Education & training 14-19: curriculum, qualifications & organization*. London: Sage
- Hodgson, A. & Spours, K. (2011) National education systems: the wider context. Hodgson, A., Spours, K. & Waring, M. (Ed) *Post-compulsory education and lifelong learning across the United Kingdom: policy, organisation and governance*. London: Bedford Way Papers
- Hodgson, A., Spours, K., Waring, M., Gallacher., Keep, E. & Rees, M. (2011) Post-compulsory education and lifelong learning across the UK: moving from managed divergence to a new period of uncertainty. Hodgson, A., Spours, K. & Waring, M. (Ed) *Post-compulsory education and lifelong learning across the United Kingdom: policy, organisation and governance*. London: Bedford Way Papers
- Hodgson, A. & Spours, K. (2014) *17+ participation, attainment and progression in London: summary report key findings and recommendations*. London: IoE
- Hodgson, A., Bailey, B. & Lucas, N. (2015) What is FE? Hodgson, A. (Ed) *The coming of age for FE*. London: Institute of Education Press
- Institute of Education/Ofsted (2004) *Improvement through inspection: an evaluation of the impact of Ofsted's work*. London: Ofsted
- Jackson, A. & Wallis, J. (2006) No pain, no gain? Learning from inspection. *Research in Post-Compulsory Education*. Vol 11, (3) 251-266
- Janssens, F. & van Amelsvoort, G. (2008) School self-evaluations and school inspections in Europe: an exploratory study. *Studies in Educational Evaluations*. Vol 34 (2008) 15-23
- Janssens, F. (2007) Supervising the quality of education. Bottcher, W. & Kotthoff, H. (Ed) *School inspection: evaluation, accountability and quality development*. Munster: Waxman
- Jeffrey, B. & Woods, P. (2005) Feeling deprofessionalised: the social construction of emotions during an Ofsted inspection. *Journal of Education*. Vol 26, (3) 325 - 343
- Johnston, M. (2014) Secondary data analysis: a method of which the time has come. *Qualitative and Quantitative Methods in Libraries*. Vol QQML, (3) 619 - 626
- Jones, K. & Tymms, P. (2014) Ofsted's role in promoting school improvement: the mechanisms of the school inspection system in England. *Oxford Review of Education*. Vol 40, (3) 315-330
- Jupp, T. (2015) Leadership and leaders of colleges. Hodgson, A. (Ed) *The coming of age for FE*. London: Institute of Education Press
- Keep, E. (2011) The English skills policy narrative. Hodgson, A., Spours, K. & Waring, M. (Ed) *Post-compulsory education and lifelong learning across the United Kingdom: policy, organisation and governance*. London: Bedford Way Papers

- Kennedy, H. (1997) *Learning Works: Widening Participation in Further Education*. Coventry: Further Education Funding Council
- Lee, J. & Fitz, J. (1997) HMI and Ofsted: evolution or revolution in school inspection. *British Journal of Educational Studies*. Vol 45, (1) 39-52
- Leitch, S. (2005) *Skills in the UK: the long term challenge*. Norwich: HMSO
- Leitch, S. (2006) *The Leitch Review of Skills: prospoerty for all in a global economy – world class skills*. Norwich: HMSO
- Levacic, R. & Glover, D. (2006) Relationship between efficient resource management and school effectiveness: evidence from Ofsted secondary school data inspections. *School effectiveness and school improvement*. Vol 9, (1) 95-122
- Lumby, J. (2003) *Accountability in further education: the impact of UK government policy*. Lincoln: University of Lincoln
- Lumby, J. & Foskett, N. (2005). *14-19 Education: Policy, Leadership & Learning*. London: Sage
- Lumby, J & Foskett, N. (2007). Turbulence masquerading as change: exploring 14-19 policy. Raffe, D. & Spours, K. (Ed.) *Policy-making and policy learning in 14-19 education*. London: Bedford Way Papers, Institute of Education, University of London
- Major, J. (1991) *Education: all our futures*. London: Centre for Policy Studies
- Mamon, S. (2011) *Enrolment falls as EMA withdrawn [online]* Institute of Race Relations. Available from: <http://www.irr.org.uk/news/enrolment-falls-as-ema-withdrawn/> [Accessed 10 June 2016]
- Mansell, W. (2010) *The Diploma's future is in the balance*. London: TES
- Matthews, P. & Sammons, P., (2004) *Improvement through inspection: an evaluation of the impact of Ofsted's work*. London: Ofsted
- McInerney, L. (2016) *Teacher of not, Ofsted's new chief inspector passes the test [online]* The Guardian. Available from: <http://www.theguardian.com/education/2016/jun/21/teacher-ofsted-amanda-spielman-lack-classroom-experience> [Accessed 22 June 2016]
- McInerney, L. (2016a) *Farewell, Sir Michael Wilshaw, the Dirty Harry of Ofsted [online]* The Guardian. Available from: <http://www.theguardian.com/education/2016/dec/20/michael-wilshaw-ofsted-chief-inspector-schools> [Accessed 13 September 2017]
- McNiff, J., & Whitehead, J. (2006) *All you need to know about action research*. London: Sage
- McTavish, D. (2006) Further education management strategy and policy. *Educational management administration and leadership*. Vol 34, (3) 411-428
- Mogalakwe, M. (2006) The use of documentary research methods in social research. *African Sociological Review*, 10, (1) 221-230

- Morley, L. & Rassool, N. (1999) *School effectiveness: fracturing the discourse*. London: Falmer Press
- Nash, I. & Jones, S. (2015) The politician' tale. Hodgson, A. (Ed) *The coming of age for FE*. London: Institute of Education Press
- Norris, .E & Adam, A. (2017) *All change: why Britain is so prone to policy reinvention, and what can be done about it*. London: Institute for Government
- Offord, P. & Burke, J. (2017) 'Good' Ofsted report finally unveiled for recently incorporated FE college, after 90 day wait [online] Available from: <https://feweeek.co.uk/2017/02/16/good-ofsted-report-finally-unveiled-for-recently-incorporated-fe-college-after-90-day-delay/> [Accessed 15 August 2017]
- Ofsted. (2017) *Do two inspections inspecting the same school make consistent decisions? A study of the reliability of Ofsted's new short inspections*. Manchester: Ofsted
- Ofsted. (2015) *Further education and skills inspection handbook: handbook for inspecting further education and skills providers under part 8 of the Education and Inspections Act 2006*. Manchester: Ofsted
- Ofsted. (2012a) *Guidance on the use of data to support judgements on 'outcomes for learners' in learning and skills inspections*. Manchester: Ofsted
- Ofsted. (2007) *Handbook for the inspection of colleges*. Manchester: Ofsted
- Ofsted. (2012) *Handbook for the inspection of further education and skills*. Manchester: Ofsted
- Ofsted. (2009) *Handbook for the inspection of further education and skills from September 2009*. Manchester: Ofsted
- Ofsted. (2013) *Improving attendance and punctuality: strategies, approaches and lessons learned from London colleges*. Manchester. Ofsted
- Ofsted. (2000) *Inspecting schools – the framework*. London: Ofsted
- Ofsted. (2012b) *Official statistics release*. London: Ofsted
- Ofsted. (2015b) *Official statistics release*. London: Ofsted
- Ofsted. (2014a) *Official statistics release*. London: Ofsted
- Ofsted. (2014b) *Official statistics release*. London: Ofsted
- Ofsted. (2016) *Ofsted Annual Report 2015/16: education, early years and skills*. London: Ofsted
- Ofsted. (2009a) *Ofsted inspects: a framework for all Ofsted inspection and regulation*. London: Ofsted

- Ofsted. (2014c) *Raising standards, improving lives: the office for standards in education, children's services and skills*. Manchester: Ofsted
- Ofsted. (2016a) *School inspection handbook*. Manchester: Ofsted
- Ofsted. (2014) *The annual report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2013/14*. London: Ofsted
- Ofsted. (2015a) *The annual report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2014/15*. London: Ofsted
- Ofsted. (2009a) *The framework for school inspection*. Manchester: Ofsted
- Ofsted. (2004) *The future of inspection*. London. Ofsted
- Ofsted. (1993) *The handbook for the inspection of schools*. London: Ofsted
- Ouston, J., Fidler, B. & Early, P. (1997) What do schools do after Ofsted schools inspections – or before? *School Leadership & Management*. Vol 17, (1) 95-104
- Orr, K. (2009) Performativity and professional development: the gap between policy and practice in the English further education sector. *Research in Post-Compulsory Education*, 14, (4) 479-489
- Ozga, J., (2009) Governing education through data in England: from regulation to self-evaluation. *Education Policy*. Vol 24, (2) 149-162
- Panchamia, N. (no date) *Choice and competition in further education*. London: Institute for Government
- Park, J. (2013) *Detoxifying school accountability: the case for multi-perspective inspection*. London: DEMOS
- Parry, G. (2009) Higher education, further education and the English experiment. *Higher Education Quarterly*, 63, (4) 322-342
- Payne, G. & Payne, J. (2004) *Key concepts in social research*. London: Sage
- Pennickx, M., Vanhoof, J., De Maeyer, D. & Van Petegam, P. (2015) Explaining effects and side effects of school inspections: a path analysis. *School Effectiveness and Improvement*. Vol 27, (3) 333-347
- Penninckx, M, & Vanhoof, J. (2015) Insights gained by schools and emotional consequences of school inspections: a review of evidence. *School Leadership & Management*. Vol 35, (5) 477-501
- Perry, A. & Davies, P. (2015) Students count. Hodgson, A. (Ed) *The coming of age for FE*. London: Institute of Education Press
- Perryman, J. (2002) Surviving special measures: case study of a 'fresh start' school. *Improving schools*. Vol 5, (3) 46-59

- Perryman, J. (2005) School leadership and management after special measures: discipline without the gaze? *School Leadership & Management*. Vol 25, (3) 281-297
- Perryman, J. (2006) Panoptic performativity and school inspection regimes: disciplinary mechanisms and life under special measures. *Journal of Education Policy*. Vol 21, (2) 147-161
- Perryman, J. (2007) Inspection and emotion. *Cambridge Journal of Education*. Vol 37, (2) 173-190
- Perryman, J. (2009) Inspection and the fabrication of professional and performative processes. *Journal of Education Policy*. Vol 24, (5) 611-631
- Perryman, J. (2010) Improvement after inspection. *Improving schools*. Vol 13, (2) 182-196
- Plowright, D. (2007) Self-evaluation and Ofsted inspection: developing an integrative model of school improvement. *Educational Management, Administration and Leadership*. Vol 35, (3) 373-393
- Pring, R., Hayward, G., Hodgson, A., Johnson, J., Keep, E., Oancea, A., Rees, G. et al (2009) *Education for all: the future of education and training for 14-19 year olds*. London: Routledge
- Pring, R., Rawlings, E., Egan, D. & Hewlett, M., (2008). *Issue paper 7: The Whole Curriculum 14-19*. Oxford: University of Oxford.
- QCA. (2008) *The Diploma and its Pedagogy*. London: QCA
- Raff, D. & Spours, K. (2007) Three models of policy learning and policy making in 14-19 education. Raff, D. & Spours, K. (Ed) *Policy-making and policy learning in 14-19 education*. London: Bedford Way Papers
- Raff, D. & Spours, K. (2007a) Policy learning in 14-19 education: from accusation to an agenda for improvement. Raff, D & Spours, K. (Ed) *Policy-making and policy learning in 14-19 education*. London: Bedford Way Papers
- Raggatt, P. & Williams, S. (1999) *Government, markets and vocational qualifications: an anatomy of policy*. London: The Falmer Press
- RCS (1906) *Final report of the Department Committee's review on technical education*. London: Royal College of Science
- Ready, D. (2013) Associations between student achievement and students learning: implications for value-added school accountability models. *Educational Policy* Vol 27, (1) 92-120
- Remenyi, D., Williams, B., Money, A. & Swartz, E. (1998) *Doing research in business and management: an introduction to process and method*. London: Sage
- Richards, C. (2012) Ofsted inspection inspected: an examination of the 2012 framework for school improvement and its accompanying evaluation schedule. *FORUM*, 54 (2), 247-272

- Richards, C. (2001) the limitation and possibilities of primary school inspection. *Education Review. Vol 14, (2) 42-47*
- Roberts, N. & Abreu, L. (2016) *School inspections in England*. London: House of Commons Library
- Robertson, A. (2017) *Why the delay for Learndirect inspection report?* [online] Available from: <https://feweek.co.uk/2017/04/28/why-the-delay-for-learndirect-inspection-report/> [Accessed 15 August 2017]
- Roelande, H., Dijkstra, N. & Hoffman, A. (2009) School self-evaluation and student achievement. *School Effectiveness and School Improvement. Vol 20, (1) 47-68*
- Ronnberg, L. (2014) Justifying the need for control: motives for Swedish national school inspection during two governments. *Scandinavian journal of educational research. Vol 58, (4) 385-399*
- Sandbrook, I. (1996) *Making sense of primary inspection*. London: Open University Press
- Saunders, MNK., Lewis, P & Thornhill, A. (2016) *Research methods for business studies (7th Ed)*. Harlow: Pearson Education
- Segerholm, C. & Hult, A. (2016) Learning from and reacting to school inspection – two Swedish case narratives. *Scandinavian journal of education. Vol 62 (1) 125-139*
- Scanlon, M. (1999) *The impact of Ofsted inspections*. Slough: NFER
- Schagen, I. & Weston, P. (1998) Insights into school effectiveness from analysis of Ofsted's school inspection database. *Oxford review of education. Vol 24, (3) 337-344*
- Scott, J. (1990) *A matter of record: documentary sources in social research*. Cambridge: Polity Press
- Seidman, I.E. (1991) *Interviewing as qualitative research: a guide of researchers in education and the social sciences*. New York: Teachers College Press
- SFA. (2014) *Qualification success rates: rationale for the method to be used from 2013 to 2014*. Coventry: SFA
- SFA. (2014a) *Education and training national success rate tables for the 2012/2013 academic year*. Guidance notes. Coventry. SFA
- SFA. (2015) *The effects of the changes to qualification success rate methodologies in 2013/14*. Coventry: SFA
- SFA. (2015a) *Education and training national success rates tables for the 2013/2014 academic year: guidance notes*. Coventry: SFA
- Shaw, I., Newton, P., Aitkin, M. & Darnell, R. (2003) Do Ofsted inspections of secondary schools make a difference to GCSE results? *British Educational Research Journal. Vol 29 (1) 63-71*

- Sinkinson, A. (2004) Manager, moderator, motivator or what? Does ensuring consistency of reporting fall within the role of managing inspector in initial teacher training Ofsted inspections. *Educational review*. Vol 56, (3) 235-246
- Smith, E. (2006) Using secondary data in educational and social research. Maidenhead: McGraw Hill/Open University Press
- Smith, G. (2000) Research and inspection: HMI and Ofsted, 1981-1996 – a commentary. Vol 26, (3/4) 333-352
- Smithers, A. (2002). *Vocational Education: comparing standards, academic and vocational, 16-19 year olds*. London: Politea
- Steadman C. and Singleton D. (2003) *Quality assurance and the evaluation of post-compulsory education inspections 2002 to 2003*. Manchester: Ofsted
- Steadman C. and Singleton D. (2003) *College and other post-compulsory education inspections: the views of providers*. Manchester: Ofsted
- Steele, A. (2000) *Understanding the Ofsted schools inspection process*. London: FEDA
- Sykes A. and Key T. (2004) *Changes in inspection arrangements since 1993*. Manchester: Ofsted
- The national archives (no date) *Technical colleges and further education [online]*. Available from <http://www.nationalarchives.gov.uk/help-with-your-research/research-guides/technical-colleges-further-education/> [Accessed 22 March 2016]
- Thomson, P., Hall, C. & Jones, K. (2010) Maggie's day: a small-scale analysis of English educational policy. *Journal of Education Policy*. Vol 25, (5) 639-656
- Trowman, G. (1996) *Models of the 'good' teacher*. London: Routledge
- UCU. (2016) *Government figures reveal an 11% drop in adults studying further education [on line]*. Available from: <https://www.ucu.org.uk/article/8323/Government-figures-reveal-an-11-drop-in-adults-studying-further-education> [Accessed 23 June 2016]
- University of Exeter (2016) *A history of FE in England: narratives and reputation change between 1944 and 1996 [on line]*. Available from <http://elac.ex.ac.uk/fecolleges/> [Accessed 22 March 2016]
- Unwin, L. & Bailey, B. (2014) Continuity and change in English further education: a century of voluntarism and permissive adaptability. *British Journal of Education Studies*. Vol 62, (4) 449-464
- Waldgrave, H. & Simons, J. (2014) *Watching the watchman: the future of school inspection in England*. London: Policy Exchange
- West, A., Mattei, P. & Roberts, J. (2011) Accountability and sanctions in English schools. *British Journal of Education Studies* Vol 59, (1) 41-62
- Wiggins, K. (2016) *White paper: Ofsted to stop grading quality of teaching [online]*. Available

from <https://www.tes.com/news/school-news/breaking-news/white-paper-ofsted-stop-grading-quality-teaching> [Accessed 17 June 2016]

Williams, S. (1999) Policy failure in vocational education and training: the introduction of national vocational qualifications (1986-1990). *Education & Training*, Volume 41 (Issue 5)

Wilshaw, M. (2016) *Ambition for education*. London: Ofsted

Wilshaw, M. (2013) Ofsted announces measures to tackle under performance [online]. Available from <https://www.gov.uk/government/news/ofsted-announces-measures-to-tackle-under-performance> [Accessed 12 January 2017]

Wolf, A. (2009) *An adult approach to further education*. London: Institute of Economic Affairs

Wolf, A. (2011) Review of vocational education: the Wolf report. London: DfE

Witcher, B. (2006) *What should a PhD thesis look like?* Norwich: University of East Anglia

Yin, R. (2009) *Case study research: design and methods* (4th ed). London: Sage

Young, M. & Leney, T. (1997) From A-Levels to an Advanced Level Curriculum of the Future. Hodgson, A. & Spours, K. (Ed.). *Dearing and beyond: 14-19 qualifications, frameworks and systems*. London: Kogan Page.

Young, M. (1997) The Dearing review of 16-19 qualification: a step towards a unified system? Hodgson, A. & Spours, K. (Ed.). *Dearing and beyond: 14-19 qualifications, frameworks and systems*. London: Kogan Page.

YPLA. (2011) *A field guide to sixth form success rates*. Coventry: YPLA

YPLA. (2010) *Value Added for 16-18 learners: learner achievement tracker handbook*. Coventry: YPLA

Appendix 1: Textual (Inspection Reports) Analysis Codes

Code	Age	Direction	Proportion	Level
1	16-18	High/Above/Average/Increased	Most/The Majority	Foundation
2	16-18	High/Above/Average/Increased	Most/The Majority	Intermediate
3	16-18	High/Above/Average/Increased	Most/The Majority	Advanced
4	16-18	High/Above/Average/Increased	Most/The Majority	Apprenticeship
5	16-18	High/Above/Average/Increased	Most/The Majority	A-Levels
6	16-19	High/Above/Average/Increased	Most/The Majority	All Courses
7	16-18	High/Above/Average/Increased	Most/The Majority	Functional Skills
8	16-18	High/Above/Average/Increased	Most/The Majority	GCSE E&M
9	16-19	High/Above/Average/Increased	Most/The Majority	E&M
10	16-18	High/Above/Average/Increased	Few/The Minority	Foundation
11	16-18	High/Above/Average/Increased	Few/The Minority	Intermediate
12	16-18	High/Above/Average/Increased	Few/The Minority	Advanced
13	16-18	High/Above/Average/Increased	Few/The Minority	Apprenticeship
14	16-18	High/Above/Average/Increased	Few/The Minority	A-Levels
15	16-18	High/Above/Average/Increased	Few/The Minority	All Courses
16	16-18	High/Above/Average/Increased	Few/The Minority	Functional Skills
17	16-18	High/Above/Average/Increased	Few/The Minority	GCSE E&M
18	16-19	High/Above/Average/Increased	Few/The Minority	E&M
19	16-18	High/Above/Average/Increased	Equal/Some	Foundation
20	16-18	High/Above/Average/Increased	Equal/Some	Intermediate
21	16-18	High/Above/Average/Increased	Equal/Some	Advanced
22	16-18	High/Above/Average/Increased	Equal/Some	Apprenticeship
23	16-18	High/Above/Average/Increased	Equal/Some	A-Levels
24	16-18	High/Above/Average/Increased	Equal/Some	All Courses
25	16-18	High/Above/Average/Increased	Equal/Some	Functional Skills
26	16-18	High/Above/Average/Increased	Equal/Some	GCSE E&M
27	16-19	High/Above/Average/Increased	Equal/Some	E&M
28	16-18	Expected Progress/In line	Most/The Majority	Foundation
29	16-18	Expected Progress/In line	Most/The Majority	Intermediate
30	16-18	Expected Progress/In line	Most/The Majority	Advanced
31	16-18	Expected Progress/In line	Most/The Majority	Apprenticeship
32	16-18	Expected Progress/In line	Most/The Majority	A-Levels
33	16-18	Expected Progress/In line	Most/The Majority	All Courses
34	16-18	Expected Progress/In line	Most/The Majority	Functional Skills
35	16-18	Expected Progress/In line	Most/The Majority	GCSE E&M
36	16-18	Expected Progress/In line	Most/The Majority	E&M
37	16-18	Expected Progress/In line	Few/The Minority	Foundation
38	16-18	Expected Progress/In line	Few/The Minority	Intermediate
39	16-18	Expected Progress/In line	Few/The Minority	Advanced

40	16-18	Expected Progress/In line	Few/The Minority	Apprenticeship
41	16-18	Expected Progress/In line	Few/The Minority	A-Levels
42	16-18	Expected Progress/In line	Few/The Minority	All Courses
43	16-18	Expected Progress/In line	Few/The Minority	Functional Skills
44	16-18	Expected Progress/In line	Few/The Minority	GCSE E&M
45	16-18	Expected Progress/In line	Few/The Minority	E&M
46	16-18	Expected Progress/In line	Equal/Some	Foundation
47	16-18	Expected Progress/In line	Equal/Some	Intermediate
48	16-18	Expected Progress/In line	Equal/Some	Advanced
49	16-18	Expected Progress/In line	Equal/Some	Apprenticeship
50	16-18	Expected Progress/In line	Equal/Some	A-Levels
51	16-18	Expected Progress/In line	Equal/Some	All Courses
52	16-18	Expected Progress/In line	Equal/Some	Functional Skills
53	16-18	Expected Progress/In line	Equal/Some	GCSE E&M
54	16-18	Expected Progress/In line	Equal/Some	E&M
55	16-18	Below/National Average/Do Not Complete	Most/The Majority	Foundation
56	16-18	Below/National Average/Do Not Complete	Most/The Majority	Intermediate
57	16-18	Below/National Average/Do Not Complete	Most/The Majority	Advanced
58	16-18	Below/National Average/Do Not Complete	Most/The Majority	Apprenticeship
59	16-18	Below/National Average/Do Not Complete	Most/The Majority	A-Levels
60	16-18	Below/National Average/Do Not Complete	Most/The Majority	All Courses
61	16-18	Below/National Average/Do Not Complete	Most/The Majority	Functional Skills
62	16-18	Below/National Average/Do Not Complete	Most/The Majority	GCSE E&M
63	16-18	Below/National Average/Do Not Complete	Most/The Majority	E&M
64	16-18	Below/National Average/Do Not Complete	Few/The Minority	Foundation
65	16-18	Below/National Average/Do Not Complete	Few/The Minority	Intermediate
66	16-18	Below/National Average/Do Not Complete	Few/The Minority	Advanced
67	16-18	Below/National Average/Do Not Complete	Few/The Minority	Apprenticeship
68	16-18	Below/National Average/Do Not Complete	Few/The Minority	A-Levels
69	16-18	Below/National Average/Do Not Complete	Few/The Minority	All Courses
70	16-18	Below/National Average/Do Not Complete	Few/The Minority	Functional Skills
71	16-18	Below/National Average/Do Not Complete	Few/The Minority	GCSE E&M
72	16-18	Below/National Average/Do Not Complete	Few/The Minority	E&M
73	16-18	Below/National Average/Do Not Complete	Equal/Some	Foundation
74	16-18	Below/National Average/Do Not Complete	Equal/Some	Intermediate
75	16-18	Below/National Average/Do Not Complete	Equal/Some	Advanced
76	16-18	Below/National Average/Do Not Complete	Equal/Some	Apprenticeship
77	16-18	Below/National Average/Do Not Complete	Equal/Some	A-Levels
78	16-18	Below/National Average/Do Not Complete	Equal/Some	All Courses
79	16-18	Below/National Average/Do Not Complete	Equal/Some	Functional Skills
80	16-18	Below/National Average/Do Not Complete	Equal/Some	GCSE E&M
81	16-18	Below/National Average/Do Not Complete	Equal/Some	E&M

82	19+	High/Above/Average/Increased	Most/The Majority	Foundation
83	19+	High/Above/Average/Increased	Most/The Majority	Intermediate
84	19+	High/Above/Average/Increased	Most/The Majority	Advanced
85	19+	High/Above/Average/Increased	Most/The Majority	Apprenticeship
86	19+	High/Above/Average/Increased	Most/The Majority	A-Levels
87	19+	High/Above/Average/Increased	Most/The Majority	All Courses
88	19+	High/Above/Average/Increased	Most/The Majority	Functional Skills
89	19+	High/Above/Average/Increased	Most/The Majority	GCSE E&M
90	19+	High/Above/Average/Increased	Most/The Majority	E&M
91	19+	High/Above/Average/Increased	Few/The Minority	Foundation
92	19+	High/Above/Average/Increased	Few/The Minority	Intermediate
93	19+	High/Above/Average/Increased	Few/The Minority	Advanced
94	19+	High/Above/Average/Increased	Few/The Minority	Apprenticeship
95	19+	High/Above/Average/Increased	Few/The Minority	A-Levels
96	19+	High/Above/Average/Increased	Few/The Minority	All Courses
97	19+	High/Above/Average/Increased	Few/The Minority	Functional Skills
98	19+	High/Above/Average/Increased	Few/The Minority	GCSE E&M
99	19+	High/Above/Average/Increased	Few/The Minority	E&M
100	19+	High/Above/Average/Increased	Equal/Some	Foundation
101	19+	High/Above/Average/Increased	Equal/Some	Intermediate
102	19+	High/Above/Average/Increased	Equal/Some	Advanced
103	19+	High/Above/Average/Increased	Equal/Some	Apprenticeship
104	19+	High/Above/Average/Increased	Equal/Some	A-Levels
105	19+	High/Above/Average/Increased	Equal/Some	All Courses
106	19+	High/Above/Average/Increased	Equal/Some	Functional Skills
107	19+	High/Above/Average/Increased	Equal/Some	GCSE E&M
108	19+	High/Above/Average/Increased	Equal/Some	E&M
109	19+	Expected Progress/In line	Most/The Majority	Foundation
110	19+	Expected Progress/In line	Most/The Majority	Intermediate
111	19+	Expected Progress/In line	Most/The Majority	Advanced
112	19+	Expected Progress/In line	Most/The Majority	Apprenticeship
113	19+	Expected Progress/In line	Most/The Majority	A-Levels
114	19+	Expected Progress/In line	Most/The Majority	All Courses
115	19+	Expected Progress/In line	Most/The Majority	Functional Skills
116	19+	Expected Progress/In line	Most/The Majority	GCSE E&M
117	19+	Expected Progress/In line	Most/The Majority	E&M
118	19+	Expected Progress/In line	Few/The Minority	Foundation
119	19+	Expected Progress/In line	Few/The Minority	Intermediate
120	19+	Expected Progress/In line	Few/The Minority	Advanced
121	19+	Expected Progress/In line	Few/The Minority	Apprenticeship
122	19+	Expected Progress/In line	Few/The Minority	A-Levels
123	19+	Expected Progress/In line	Few/The Minority	All Courses

124	19+	Expected Progress/In line	Few/The Minority	Functional Skills
125	19+	Expected Progress/In line	Few/The Minority	GCSE E&M
126	19+	Expected Progress/In line	Few/The Minority	E&M
127	19+	Expected Progress/In line	Equal/Some	Foundation
128	19+	Expected Progress/In line	Equal/Some	Intermediate
129	19+	Expected Progress/In line	Equal/Some	Advanced
130	19+	Expected Progress/In line	Equal/Some	Apprenticeship
131	19+	Expected Progress/In line	Equal/Some	A-Levels
132	19+	Expected Progress/In line	Equal/Some	All Courses
133	19+	Expected Progress/In line	Equal/Some	Functional Skills
134	19+	Expected Progress/In line	Equal/Some	GCSE E&M
135	19+	Expected Progress/In line	Equal/Some	E&M
136	19+	Below/National Average/Do Not Complete	Most/The Majority	Foundation
137	19+	Below/National Average/Do Not Complete	Most/The Majority	Intermediate
138	19+	Below/National Average/Do Not Complete	Most/The Majority	Advanced
139	19+	Below/National Average/Do Not Complete	Most/The Majority	Apprenticeship
140	19+	Below/National Average/Do Not Complete	Most/The Majority	A-Levels
141	19+	Below/National Average/Do Not Complete	Most/The Majority	All Courses
142	19+	Below/National Average/Do Not Complete	Most/The Majority	Functional Skills
143	19+	Below/National Average/Do Not Complete	Most/The Majority	GCSE E&M
144	19+	Below/National Average/Do Not Complete	Most/The Majority	E&M
145	19+	Below/National Average/Do Not Complete	Few/The Minority	Foundation
146	19+	Below/National Average/Do Not Complete	Few/The Minority	Intermediate
147	19+	Below/National Average/Do Not Complete	Few/The Minority	Advanced
148	19+	Below/National Average/Do Not Complete	Few/The Minority	Apprenticeship
149	19+	Below/National Average/Do Not Complete	Few/The Minority	A-Levels
150	19+	Below/National Average/Do Not Complete	Few/The Minority	All Courses
151	19+	Below/National Average/Do Not Complete	Few/The Minority	Functional Skills
152	19+	Below/National Average/Do Not Complete	Few/The Minority	GCSE E&M
153	19+	Below/National Average/Do Not Complete	Few/The Minority	E&M
154	19+	Below/National Average/Do Not Complete	Equal/Some	Foundation
155	19+	Below/National Average/Do Not Complete	Equal/Some	Intermediate
156	19+	Below/National Average/Do Not Complete	Equal/Some	Advanced
157	19+	Below/National Average/Do Not Complete	Equal/Some	Apprenticeship
158	19+	Below/National Average/Do Not Complete	Equal/Some	A-Levels
159	19+	Below/National Average/Do Not Complete	Equal/Some	All Courses
160	19+	Below/National Average/Do Not Complete	Equal/Some	Functional Skills
161	19+	Below/National Average/Do Not Complete	Equal/Some	GCSE E&M
162	19+	Below/National Average/Do Not Complete	Equal/Some	E&M
163	WBL	High/Above/Average/Increased	Most/The Majority	Intermediate
164	WBL	High/Above/Average/Increased	Most/The Majority	Advanced
165	WBL	High/Above/Average/Increased	Most/The Majority	All Courses

166	WBL	High/Above/Average/Increased	Few/The Minority	Intermediate
167	WBL	High/Above/Average/Increased	Few/The Minority	Advanced
168	WBL	High/Above/Average/Increased	Few/The Minority	All Courses
169	WBL	High/Above/Average/Increased	Equal/Some	Intermediate
170	WBL	High/Above/Average/Increased	Equal/Some	Advanced
171	WBL	High/Above/Average/Increased	Equal/Some	All Courses
172	WBL	Expected Progress/In line	Most/The Majority	Intermediate
173	WBL	Expected Progress/In line	Most/The Majority	Advanced
174	WBL	Expected Progress/In line	Most/The Majority	All Courses
175	WBL	Expected Progress/In line	Few/The Minority	Intermediate
176	WBL	Expected Progress/In line	Few/The Minority	Advanced
177	WBL	Expected Progress/In line	Few/The Minority	All Courses
178	WBL	Expected Progress/In line	Equal/Some	Intermediate
179	WBL	Expected Progress/In line	Equal/Some	Advanced
180	WBL	Expected Progress/In line	Equal/Some	All Courses
181	WBL	Below/National Average/Do Not Complete	Most/The Majority	Intermediate
182	WBL	Below/National Average/Do Not Complete	Most/The Majority	Advanced
183	WBL	Below/National Average/Do Not Complete	Most/The Majority	All Courses
184	WBL	Below/National Average/Do Not Complete	Few/The Minority	Intermediate
185	WBL	Below/National Average/Do Not Complete	Few/The Minority	Advanced
186	WBL	Below/National Average/Do Not Complete	Few/The Minority	All Courses
187	WBL	Below/National Average/Do Not Complete	Equal/Some	Intermediate
188	WBL	Below/National Average/Do Not Complete	Equal/Some	Advanced
189	WBL	Below/National Average/Do Not Complete	Equal/Some	All Courses

Appendix 2: Freedom of Inspection Request, letter



Date

[Name]
Principal & Chief Executive
[Name of College]

By email

Dear [first name]


Re: Freedom of Information Request

I am writing to you under the Freedom of Information Act 2000 with regards to the Ofsted inspection(s) which took place at your College in [month & year].

I am currently a doctoral student at the University of Winchester. My thesis concerns itself with the use of data and how it informs the 'outcomes for learners' judgement in inspections of further education colleges that took place between 2012 and 2015. Whilst success rate information and inspection reports for this period are publically available, there are other areas of data and information which are not, such as attendance, value added and destination/progression (although this might appear in governing body papers and minutes). You would have had to produce and offer this evidence to the inspection team as part of your inspection.

Therefore, using the template below, could you please send me the data and information requested. I understand that under the Act I am entitled to a response within 20 working days of receipt of this request. However, as an experienced senior manager within the further education sector, I am aware of the pressures that face colleges at the beginning of the academic year. As a result, I am content to receive this information at your convenience, but no later than **Friday 25 November 2016**.

Please note that all information provided will be anonymised and there will be no reference to your college by name in my thesis. If my request is denied in whole or in part, I ask that you justify all deletions by reference to specific exemptions of the Act.

I would prefer to receive the information electronically, which should be sent to: . If you require any clarification, I expect you to contact me under your section 16 duty to provide advice and assistance if you find any aspect of this FOI request problematic.

Please acknowledge receipt of this request, and I look forward to receiving the information in due course.

Yours sincerely

Robert Rees

Doctoral candidate at the University of Winchester (student number: 1013535)

Appendix 3: Freedom of Inspection Request, data table

Freedom of Information Request – to be completed

Attendance

Ref	Question	Answer			
A1	What was the College's overall reported attendance (cumulative) at the point of inspection?	%			
A2	What was the College's overall reported attendance (cumulative) at the year end prior to inspection?	%			
A3	What was the College's reported attendance (cumulative) at the point of inspection for English and Mathematics (Functional Skills and GCSE)?	%			
A4	Was there any significant difference between the reported attendance (cumulative) for English and mathematics and between Functional Skills and GCSE?	No	Yes		
A4Y	If the answer is yes to A4, please provide the detail				
	<table border="1"> <tr> <td>FS English = %</td> <td>FS Mathematics = %</td> <td>GCSE English = %</td> <td>GCSE Mathematics = %</td> </tr> </table>			FS English = %	FS Mathematics = %
FS English = %	FS Mathematics = %	GCSE English = %	GCSE Mathematics = %		
A5	What was the College's overall reported attendance (cumulative) at the year end prior to inspection for English and Mathematics (Functional Skills and GCSE)?	%			
A6	Was there any significant difference between the reported attendance (cumulative) for English and mathematics and between Functional Skills and GCSE at the year end prior to inspection?	No	Yes		
A6Y	If the answer is yes to A6, please provide the detail				
	<table border="1"> <tr> <td>FS English = %</td> <td>FS Mathematics = %</td> <td>GCSE English = %</td> <td>GCSE Mathematics = %</td> </tr> </table>			FS English = %	FS Mathematics = %
FS English = %	FS Mathematics = %	GCSE English = %	GCSE Mathematics = %		
A7	Did the inspection team share with the Principal or Nominee that there was a difference between the reported attendance (used by the college at the point of inspection) and the reported attendance by the inspection team?	No	Yes		
A7Y	If the answer is yes to A7, please identify how this was reported				
	<table border="1"> <tr> <td>They didn't give detail</td> <td>Inspection team attendance (%)</td> <td>They said it was higher</td> <td>They said it was lower</td> </tr> </table>			They didn't give detail	Inspection team attendance (%)
They didn't give detail	Inspection team attendance (%)	They said it was higher	They said it was lower		

Value Added

Ref	Question	Answer	
V1	At the point of inspection, did the College formally use and report on value added as a KPI?	Yes	No
V1Y	If the answer is yes to V1, what value added tool is used at the College i.e. ALPS, LAT, college devised? Please specify		
V2	At inspection, did the inspection team explore value added data and information?	Yes	No
V2Y	If the answer is yes to V2, how much weight was placed on it as a tool for measuring student outcomes?	Significant	Hardly any

Progression (progressing within the College)

Ref	Question	Answer	
P1	At inspection, what type of progression information was used? Intended progression (I) (normally collected by the course tutor or teacher prior to the student leaving the college in the summer) or Actual progression (A) (provided by MIS/central data taken directly from the ILR)?	I	A
P2	At inspection, what figure was presented to the inspection team that indicated positive progression?	%	
P3	What % was 'unknown'? Please indicate % by age-based cohort i.e. 16-18, 19+.	16-18 %	19+ %
P4	At inspection (or in the SAR used by the inspection team) were there any gaps where progression was identified i.e. between levels or within curriculum/subject areas?	No	Yes
P4Y	If the answer is yes to P4, please provide the detail.		
	Levels: • Curriculum Areas: •		

Destination (out from the College)

Ref	Question	Answer	
D1	At inspection, what type of destination information was used? Intended destination (I) (normally collected by the course tutor or teacher prior to the student leaving the college in the summer) or Actual destination (A) (provided by MIS/central data taken directly from the ILR, UCAS return, sub-contracted company)?	I	A
D2	At inspection, what figure was presented to the inspection team that indicated positive destination?	%	
D3	What % was 'unknown'? Please indicate % by age-based cohort i.e. 16-18, 19+.	16-18	19+
D4	At inspection (or in the SAR used by the inspection team) were there any gaps where destination was identified i.e. between levels or within curriculum/subject areas?	No	Yes
D4Y	If the answer is yes to D4, please provide the detail.		
	Levels: • Curriculum Areas: •		

This Agreement should be completed, signed and bound with the hard copy of the thesis and also included in the e-copy. (see Thesis Presentation Guidelines for details).

Access Permissions and Transfer of Non-Exclusive Rights

By giving permission you understand that your thesis will be accessible to a wide variety of people and institutions – including automated agents – via the World Wide Web and that an electronic copy of your thesis may also be included in the British Library Electronic Theses On-line System (EThOS). Once the Work is deposited, a citation to the Work will always remain visible. Removal of the Work can be made after discussion with the University of Winchester’s Research Repository, who shall make best efforts to ensure removal of the Work from any third party with whom the University of Winchester’s Research Repository has an agreement.

Agreement:

I understand that the thesis listed on this form will be deposited in the University of Winchester’s Research Repository, and by giving permission to the University of Winchester to make my thesis publically available I agree that the:

- University of Winchester’s Research Repository administrators or any third party with whom the University of Winchester’s Research Repository has an agreement to do so may, without changing content, translate the Work to any medium or format for the purpose of future preservation and accessibility.
- University of Winchester’s Research Repository reserves the right to remove the Work for any professional, administrative or legal reason. Equally, I may request that the Work is removed at any point in the future. I understand that once the Work is deposited, a citation to the Work will always remain visible, although the author retains the right to update the Work.
- rights granted to the University of Winchester’ Research Repository through this agreement are entirely non-exclusive and royalty free; that I am free to publish the Work in its present version or future versions elsewhere; and that no ownership is assumed by the repository when storing the Work.

Please select **ONE** of the following six statements: *(by placing an X in the box)*

1. The thesis is to be made available from the moment the deposit has been approved by the University of Winchester.
2. In order to facilitate commercial publication I request an embargo period of 1 year*
3. In order to facilitate commercial publication I request an embargo period of 2 years*
4. In order to facilitate commercial publication I request an exceptional embargo period of 3 years*

Approval from the Director of Research and Knowledge Exchange (DRKE) is required for a 3 year embargo.

**Signature of
DRKE:**

N.B. under Research Council UK (RCUK) rules and regulations, students who have received RCUK funding may apply for access to their thesis to be withheld for no more than 12 months.

**Please provide an explanation for the length of the embargo period and information about the proposed publication arrangements. The date of the embargo period will commence from the date of the Award Letter sent from the Director of Postgraduate Research Students.*

5. The title, abstract and keywords may be published, but the thesis may not be published in the data repository because it contains commercially sensitive data.^

^Please identify the commercially sensitive data.

6. The title, abstract and keywords may be published, but the thesis may not be published because my sponsors/funders own the rights to it. #

Please identify the funder/sponsor and the contract information.

Please tick if you agree to the following TWO statements : (by placing an X in the box)

- 1. I agree that my thesis may be copied on demand by individuals or libraries and that the copy of my thesis contains the following statement:**

Copyright © Author (year of submission) 'Full thesis title', University of Winchester, PhD Thesis, pagination, ORCID [number].

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the author. Details may be obtained from the RKE Centre, University of Winchester.

This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the author.

No profit may be made from selling, copying or licensing the author's work without further agreement.

- 2. I agree that my obligation is to notify the University of Winchester of any change of address in order to facilitate any future requests for licensing.**

Author's Name:

Dr Robert David Rees

Author's Signature:



Date:

11 September 2018

Metadata to be included with the thesis:

Title:

How are outcomes for learners judged in the further education and skills sector?

Author:

Dr Robert David Rees

Supervisors:

Dr Bridget Egan & Dr Kerry Ball

Key words (*up to 10 keywords – ask your supervisory team for advice*):

Inspection, Further Education, Ofsted, Accountability, Outcomes, Accuracy, Consistency, Success Rates

Abstract:

My study analyses how inspectors use success rate, value added, attendance, progression and destination data to evaluate student performance and to make judgements on outcomes for learners in further education colleges. Ofsted (2017) acknowledges that there is little empirical evidence about the validity of inspection judgements. My previous experience suggests that inspectors approach and interpret data inconsistently. To date, hardly any research into this aspect of inspection has been published. I analysed 165 further education college inspection reports from inspections that took place between September 2012 and July 2015, along with the publicly available success rate data the inspectors would have used to evaluate student performance. For aspects where data is not collected centrally, this was obtained directly from colleges within the scope of this study. Data are analysed using approaches that reflect the ways in which Ofsted uses numerical information. My findings indicate that there is (often considerable) disparity between the success rate data and the judgements made by inspectors, and that the use of the national success rate average as a foundation from which student performance is assessed is inconsistent. I further suggest that, because most of the data (other than success rates) is generated by the individual colleges in the absence of any national guidance, inspectors are not able to compare one college with any other objectively. Moreover, not all inspection reports make judgements to the same criteria, with too many reports missing key judgements. This evidence indicates that there is a lack of transparency about which student groups are included or excluded from the data used by inspectors to inform their judgements. I therefore conclude that the consistency with which inspectors use data and information to evaluate student performance and make judgements on outcomes for learners, 'requires improvement', and suggest how this situation might be improved.