

---

### Citation:

Tang, S. (2024) An Examination of Teacher Perspectives of Assessment and Testing in a Primary School and the Role of the Headteacher. [Online]. Leeds Beckett University. Available from:  
<[https://figshare.leedsbeckett.ac.uk/articles/thesis/An\\_Examination\\_of\\_Teacher\\_Perspectives\\_of\\_Assessment\\_and\\_Testing\\_in\\_a\\_Primary\\_School\\_and\\_The\\_Role\\_of\\_The\\_Headteacher\\_/27215439/1](https://figshare.leedsbeckett.ac.uk/articles/thesis/An_Examination_of_Teacher_Perspectives_of_Assessment_and_Testing_in_a_Primary_School_and_The_Role_of_The_Headteacher_/27215439/1)> [Accessed 6 November 2024]. - [Link](#)

Link to Leeds Beckett University Research Data and Thesis Repository record:

[10.25448/lbu.27215439.v1](https://figshare.leedsbeckett.ac.uk/articles/thesis/An_Examination_of_Teacher_Perspectives_of_Assessment_and_Testing_in_a_Primary_School_and_The_Role_of_The_Headteacher_/27215439/1)

### Item Type:

Thesis

---

### Licence:

CC BY-NC 4.0

The aim of the Leeds Beckett University Research Data and Thesis Repository is to provide open access to the outputs and data from our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett University Research Data and Thesis Repository holds a wide range of outputs and data, each of which has been checked for copyright, licenses and any relevant embargo periods have been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository or believe there to be any issues with copyright please [contact us](#) and we will investigate on a case-by-case basis.

An Examination of Teacher Perspectives of Assessment and Testing in a Primary School  
and The Role of The Headteacher.

Stacey Tang

A thesis submitted in partial fulfilment of the requirements of Leeds Beckett University for the  
degree of Doctorate in Education

July 2024

## Abstract

**Purpose:** The aim of this research is to highlight the role of the head teacher in relation to the impact of testing in primary schools. It aims to contribute and stimulate a discussion not only on the role of headteachers but also on the impact of testing in general terms. An extended aim of this research; is to support other school leaders understand the complexities of a new curriculum, assessment system and new tests and in turn allow them to understand the impact and their role, before responding to the ever-changing demands of assessment and curriculum design.

**Design and methodology:** This thesis is a longitudinal piece of research to critically review the policies that have led to the overhaul of the primary school assessment system, resulting in the removal of levels and the creation of a pass or fail testing system, (introduced in 2016). I review the impact of this new system on pupils from the perspective of their teachers, investigate the teacher perspective regarding whether formal assessment is working as intended and how the role of headteacher has led to the teacher perspectives analysed. To do this, I analyse, using a thematic approach and present data from primary school staff based within a school setting; using semi structured interviews and reflective diaries over an academic year. The findings then lead to a reflective analysis of school systems, based on Bronfenbrenner's eco systems and how headteacher decisions sit within these systems.

**Findings:** This research shows that, within the research setting, testing does not have a negative impact on pupils. These policies, which schools must legally work within, do have some degree of flexibility. This is largely due to the role of the headteacher and the decisions and culture created. School leaders are able to challenge, teach and raise standards and negate the negative impact that testing can have. They can make a difference and do not need to be ruled by fear of data or Ofsted. A school can be 'Good',

achieve floor targets and have pupils and staff that enjoy school without worrying about testing or even focussing on it!

Implications: This is an important aspect for school leaders and policy makers to consider when designing school curriculums and looking at the implications of published data and testing. It is also vital that school leaders and headteachers are aware that they do have the power to control and change the ecological systems within which they work. National policies are not as rigid as they appear and the decisions made by leaders and headteachers when creating their own policies linked to these, are vast. They do allow headteachers significant control over how they design their school ethos and curriculum.

Originality/value: This research provides an argument that headteachers do have power and control despite the current educational policy system being a top-down approach, coupled with more testing than ever before at primary school level. Alongside addressing the research questions, this research forms a narrative of the journey of a first time headteacher and novice researcher.

## Acknowledgements

There are many people that I need to express my gratitude towards. Firstly, the participants themselves for giving up their time and for seeing the value of what I hoped to achieve. The governors of the school for also supporting this research and believing in me as a headteacher.

My supervisors for the endless amounts of patience and support shown.

My husband and children for their continued support and belief. My staff not only for participating but believing in my leadership and vision. Finally, while they unfortunately are no longer here to see this research complete, my thanks to my dad and gran for making me believe I could do it in the first place.

## Table of Contents

<b>Chapter 1</b>	<b>Introduction</b>	<b>Page</b>
1.1	Introduction	8
1.2	Research Rationale and Context	8
1.3	A Brief History of Educational Assessment	25
1.4	The Introduction of SATS	36
<b>Chapter 2</b>	<b>Literature Review</b>	
2.1	Introduction	49
2.2	Aims of the Policy	52
2.3	Rationale for Change in Assessment	56
2.4	Impact of the White Paper	62
2.5	White Paper Conclusions	73
2.6	Report of the Assessment Review Group: Redressing the Balance January 2017	74
2.7	Conclusions of the Impact of the Assessment Review Group	78
2.8	Growth Mindset	84
2.9	Conclusion on the Impact of Dweck's Growth Mindset	89
2.10	Teacher Perspective Research	93
2.11	Conclusion of the Impact of Teacher Perspective Research	95
2.12	Literature Review of Resilience	97
2.13	Literature Review of Curriculum Design	100
2.14	Power in Education and the Role of the Headteacher	103
2.15	Research Questions Within the Literature Review	106
2.16	Analytical lens/Theoretical Framework Application of Literature Review	104
2.17	Literature Review Findings	109
<b>Chapter 3</b>	<b>Research Methodology</b>	
3.1	Introduction	110
3.2	Methodology and Design	110
3.3	Sampling and Selection	113
3.4	Research Methods- Case Study	117
3.4.1	Interviews	126
3.4.2	Reflective Diaries	130
3.5.	Analysis Procedure	132
3.6	Data Analysis	136
3.7	Conceptual Framework for Analysis	138
3..7.1	Application of Theoretical Framework	142
3.8	Positionality and Ethics	148
3.9	Reliability, Validity and Generalisations	156
3.10	Conclusions	161
<b>Chapter 4</b>	<b>Results and Analysis</b>	
4.1	Introduction	162
4.2	Main Findings	163
4.2.1	Microsystem Analysis of Testing and the Impact on Pupils	164

4.2.2	Microsystem Analysis of The Headteacher's Role In Relation to the Impact on Testing	174
4.2.3	Research Context	176
4.3	Microsystem Analysis of Attitudes and Behaviour	182
4.3.1	Microsystem Analysis of the Headteacher's Role in Shaping Attitudes and Behaviour	188
4.4	Mesosystem Analysis of The Learning Environment	199
4.4.1	Analysis of the Headteacher's Role in the Mesosystem	205
4.5	The Use of Labels	212
4.5.1	Analysis of the Headteacher's role in the Use of Labels Between the Micro and Mesosystems	216
4.6	Future Thoughts	220
<b>Chapter 5</b>	<b>Conclusion</b>	
5	Conclusions	222
	<b>Appendices</b>	
	Appendix 1 – Pilot Research Questions	238
	Appendix 2 – Final Agreed Areas of Discussion	239
	Appendix 3 – Example of Monthly Thoughts from Participant's Reflective Diaries	243
	Appendix 4 – Participant Information Sheet	253
	Appendix 5 – Example of Participant Clarification Sheet	256
	Bibliography	257
	Reference List	259

## Table Contents

Table 1	Research Participants	p113
---------	-----------------------	------

## Figure Contents

Figure 1	Flow Chart of Research Design	P134
Figure 2	Bronfenbrenner's Ecological Systems	P137
Figure 3	Education System Application of Bronfenbrenner's Ecological Systems	P138



## Chapter One: Introduction

### 1:1 Introduction

This chapter explains why this research was undertaken, the desired outcomes and the importance of the research findings. I explore the context of the research and provide a brief overview of the use, strengths, weaknesses and aims of assessment, and highlight my personal stance on education and testing.

### 1:2 Research Rationale and Context

In September 2014, I was appointed to my first headship. This was within a school where I had been an NQT and Deputy Head, and the position was accepted reluctantly. I had never thought of leadership and certainly did not see myself as 'headteacher material'. The school had recently been graded as requiring improvement by Ofsted and had not achieved nationally set floor targets in either foundation, key stage one (year 2), or key stage 2 (year 6) for quite some time. I wanted to change many things and indeed there was enormous pressure for rapid improvement. Ofsted were due back within 2 years, and an early decision made was that a vital change needed to be a shift away from data and targets at the heart of the school, to putting the pupils and a love of learning at the centre.

Pupils (in England) going through compulsory education at that time, were among the most assessed that the state education system has ever produced (Gipps and Stobart, 1993, p.1). In 2018, Nick Gibb claimed, during a joint hearing of the Education and Health Committees, that exam pressure is simply a part of school life and has always led to anxiety among young people. He suggested that children should therefore sit tests early to help them cope with exam related mental health problems in later life (Busby, 2018).

One year later in 2019 Damian Hinds, Education Secretary, stated that testing and the amount of testing was still a necessity and 'there are very few things that are agreed the

world over about education – the need to assess primary school attainment is one of them’

He also identified that the importance of testing and the variety of tests ‘has been one of the main things Labour and Conservative governments have agreed on educational policy over the past quarter of a century’. However, as a headteacher, I saw little improvement in relation to the data we achieved, considering the increasing number of tests pupils undertook. The school, therefore, created a whole new curriculum based on pupil and staff interests, with a clear focus on all subjects, not simply maths and English. We removed unnecessary paperwork and devised an assessment system that relied on professional teacher judgement, not tick boxes or a steady stream of testing. We had one simple strapline: ‘Everyone is important’, and we genuinely meant it. However, to get pupils to believe this, and that they matter; we had to work extremely hard to develop the resilience and mindset of both staff and pupils. We also had to put in the missing life experiences (such as visiting a beach, making a meal etc). Pupils created a list of 50 things to be completed before leaving primary school. We wanted to show pupils that education is something to love and be excited about, not something to despise and fear. Prior to this, attendance was well below national average, behaviour was not good, and observations highlighted pupils were not engaged with their learning.

James (2008, p.25) suggests that ‘learning is not simply absorbing information but an active process of meaning-making’. She identifies that cognitive constructivists view learning as people making sense of the world by building mental models of how things work and link, so they can interpret and make sense of new information. Although I was not aware of it at the time, this stance on learning and education, was the underlying belief that threaded through the changes I was making. I agreed with James’ (2008, p.90) statement that ‘today’s pupils are living in unpredictable and confusing times, in which they must have skills to think for themselves and be self-initiating and self-directed’. Pupils need the capacity to learn, change, and seek solutions for problems not yet in existence. Learning in the 21<sup>st</sup> century

involves far more than knowledge acquisition and responding to this was the vital shift I had to make.

Prior to the headship, I had long supported, the view of Davis (1998, p.168) that 'assessments for accountability purposes or its role in supporting teaching styles have severe flaws'. As a teacher, I witnessed little impact of the hours spent creating data reports and completing tests, so the headteacher could have folders of data and 'evidence'. In resonance with Davis (1998), I noticed during the informal interactions and assessments I had with pupils, the pupils grappling with a problem and my informal scribbles around planning had the most impact and yielded far better results than any of the formal testing mechanisms. This mirrored the original sentiment of the task group for assessment and testing and is the underlying premise of my first headship.

The relatively recent 'craze' for Shanghai maths in primary schools, provides one example of the extent some schools will go, to align their teaching and ethos to perform well in high-stakes testing. This was not, and is not, my stance. The school focusses very clearly on learning and not testing, as will become clear.

As a headteacher, I support Lucas, Claxton, and Spencer's (2013) stance, in that I view education as distinct from, and more valuable than league tables and testing. I believe it is the right of every child to have the right tools to enable and equip them for life, whichever path they may find themselves upon. In effect, "Education should be a great deal more than the servant of competitive economies" (Davis, 1998, p.28)

This is supported by Torrance (1998) and Williams (2008) who raise concerns relating to testing, and the distortion it provides. However, I acknowledge that other researchers such as Sumner (1987, p.29 ) suggest testing can be an enjoyable experience.

In my own experience a substantial proportion of children have said that tests were interesting and enjoyable to do. When the results are promised beforehand, the motivation of pupils can be very high, especially when the purpose is explained fully and the conditions under which the tests are done are understood.

As the daughter of a coal miner and factory worker, living in an area of high deprivation and low aspirations, the research setting reflected very much my own situation as a child. Due to this, it was important to me that I fully understood the impact of education and how it shapes school decisions and policies to allow all pupils to flourish. With the privilege of hindsight, perhaps a subconscious reason for wanting to undertake this research is to ensure that pupils have a meaningful education and to know that I did everything I could to improve their experience.

The idea of education that is more than traditional schooling is not new. Ever since statutory education for the majority of children became a reality, thoughtful educators, concerned parents, perspective employers and enlightened national policy makers have argued that schools should connect more effectively with the real world. (Lucas, Claxton, Spencer, 2013, p.1)

After 18 months of headship the school was officially ranked by Ofsted as good in every area, and floor targets were achieved (just) in key stage one and two. Perhaps the most obvious thing to focus on may have been testing (and teaching to the tests). However, this was the one thing I deliberately did not do. Instead, I made the tests the least important thing in school. Staff targets were not solely linked to standards, and pupils were beginning to believe that positive results would come if they just enjoyed school. Although the compulsory testing in year 2 and 6 remained, we removed the focus on it. At this point, both myself and the staff were positive concerning the changes made.

However, in the summer of 2016 formative assessment changed nationally following the final report of The Commission on Assessment Without Levels. The introduction of a scaled score system resulted in pupils (at the age of eleven) either passing or failing SATs (though this was labelled 'working below the standard' and 'working at the standard') this was coupled with a significant leap from what would have previously been expected to pass. The new pass mark was almost a whole level higher to achieve the new expected standard.

Ofsted also announced in June 2015, through the Education and Adoption Bill, that if a school's result be considered 'coasting' for two simultaneous years, or they achieved three 'requires improvement grades', then they would be forced into academisation.

Academisation was mentioned again in the Levelling Up White Paper (February 2022) when it stated that schools not meeting floor targets or achieving successive 'requires improvements' would face academy intervention.

I had spent nearly two years building an ethos of 'your best is always good enough, don't worry about tests', with both staff and pupils, and it felt as though Ofsted and government policies had suddenly pulled the rug from beneath us. Staff were confident in both the school's devised assessment system, the national tests and the new National Curriculum; the pupils knew where they were on the spectrum of achievement levels 3 to 6. Staff observations revealed that pupils nearly always worked out which table was the 'top table' despite teachers attempting to deflect from abilities. To have a brutal 'pass or fail' system was something that needed careful consideration to inform our next steps. We felt that the stakes could not be higher; we needed to address the new standards and expectations without risking a failure to not achieve floor targets. We needed to understand how pupils and parents might respond to a pass or fail and how staff may feel (given that their own data would be blunt, and national pay agreements were determined by pupil performance). It was this starting point that led to the research initially – what would be the impact of a pass or fail system: indeed, what is the impact of testing on our pupils? In hindsight it was here, that the role of the headteacher and their impact became vital.

Black (1998, p.123) suggests "The expectations that pupils have built up from their own experiences of assessment in school can constitute an obstacle to their taking a positive role in assessment" and as pupils were my focus, I decided to undertake a formal piece of research to investigate the impact of the new testing regime in primary schools on pupils.

Furthermore, there is a significant body of research regarding teachers teaching to the test and using transmission methods of teaching, resulting in reduced pupil motivation, anxiety, and shallow learning for the purpose of passing tests (Harlen, 1993, p.140)

Given that assessment results have an 'effect on both those who use the results and those that are assessed' (Black, 1998, p.37) and that 'test anxiety' is a recognised syndrome and has been extensively researched. It is probable that anxiety is increased when pupils realize the decisions and social kudos that are attached to the results (Sumner, 1987, p.155).

Alongside this, the Children's Society Report 2020 highlighted that UK pupils are most afraid of failure. It was the right time to examine the changes forced on schools, and this in-depth research would provide a solid understanding of the lived experiences, highlight issues, and allow us (as a school) to respond accordingly.

Whilst there were issues such as addressing staff morale, informing parents of the assessment changes, tweaking planning and adapting to the curriculum, my key focus was what I could do to help 11year old pupils come to terms with a pass or fail label. The potential impact on pupils seemed to be an appropriate area for deep investigation. As a school leader tasked with making critical decisions, it is vital these decisions have the desired impact. Undertaking this formal research on testing and its impact, will underpin my future decisions. Thus, it was both timely and important in terms of my own development, and the school's strategic direction. It was here that research took a slight change and moved from simply the impact of testing on pupils, but what could I, as the headteacher, do or had already done to limit the impact of the very system I planned on researching?

Eggleston (1991, p.58) suggests that "Perhaps due to their ambiguity, there has been little sociological examination of examinations. Like curricula it has been taken as a given to be reckoned with but not seen as a socially created phenomena", and with the recent 2021 and 2022 school closures and lack of formal testing, research is needed in these areas. While

test scores in order to pass have increased, the tests themselves and what they do test have remained very similar. As will be discussed in later chapters, the 11 + examination, reflects very much the current SATs tests of today. There is very little research on test content and the relevance of this on pupil's subject knowledge, the necessity for this particular area to be assessed and any future links between test content and application of knowledge and skills.

Kohn (2000) identifies that never before have tests been so frequent and never have they played such a prominent role in schools, with children of such a young age. He goes on to say that standardised testing is now threatening to swallow schools whole, under the guise of accountability. With this in mind it is important the impact of testing, especially with the youngest children involved in standardised testing, is investigated. Currently, primary school pupils are baselined in Foundation, at the start of their school journey and within their first 6 weeks of starting school. They then have a phonics test in Year 1, which, if not passed, is repeated in year 2. Year 2 also undertake key stage one SATs. A multiplication check occurs in Year 4 and finally KS2 SATs in Year 6. This leaves only years 3 and 5 with no testing in the education of a primary school pupil.

Gipps (1994, p.3) suggests that it is problematic that 'tests solely designed for purposes other than to support learning have unwanted and negative effects on teaching and the curriculum'. However, this does not consider the impact of testing on pupils and my research aims to address this. Stenhouse (1984) defined educational research as a 'systematic activity directed towards providing knowledge or adding to existing knowledge', and this is 'of relevance for improving the effectiveness of education' (Wellington, 2000, p.11). As practitioner/researcher, my research is central to improving my knowledge of the impact of testing and improving my professional practice.

The implementation of the government's new directives was far from smooth, and a general lack of information resulted in both local and national issues. I did not want the national 'pilot'

year to become the benchmark for the school's responses and committed to make changes to improve things for my pupils. However, it was apparent that an in-depth piece of formal research would provide a valuable tool to inform the next steps for staff, the school, and the leadership team.

My qualitative research began in the academic year 2017/18 and has shed light on the impact of the new testing regime (specifically the impact on the pupils). For the purposes of this research, I adopt Savin-Baden and Howell-Major's (2013, p.11) definition 'Qualitative research is simply social research aimed at investigating the way in which people make sense of their experiences and ideas'.

I decided in 2015, with governor approval, and therefore at the very beginning of the new assessment systems implementation not to make any changes to our systems for a full academic year. Instead, I undertook a piece of descriptive research, to inform my future decisions. This was preferential to basing actions upon a reaction to the political hype and sensationalised news articles that seemed to be everywhere. As will become clear, the headlines which had formed part of the concerns around this new system were not reflective of the research setting and, therefore, had we reacted to these, we may have caused unnecessary impact. Staff did not want to change what they were currently doing, and we were determined that there would be no negative consequences if they did not respond to the new test expectations (and teach to them). I was asking them to stand still, allowing me to step back and genuinely look at 'things', when the rest of the education sector were making significant changes to their curriculum and teaching. This, in hindsight, was also the start of the research, my action of not reacting and the impact this had. I was making a decision to shield staff and pupils, until I myself had the facts at hand with which to respond. Unknown at the time and not realised, until analysing the data, I was that shield between Bronfenbrenner's ecological systems.



It took a leap of faith for staff not to make significant changes and adaptations, to let things 'play out' to make longer term changes and hopefully gains. Initially I believed that staff would take some convincing to sit back and let things unfold so we could analyse the impact of the new assessment system, but I was wrong. Staff agreed with my request to make no changes to our current systems and were fully aware of the potential impact the new pass or fail system may have. The national training, which they had received made clear the coasting definitions. Alongside this, the Educational Excellence Everywhere document (2016) highlighted the new accountability measures and expectation of 'progress for all pupils at every level of attainment'. This document also stated that the assessment reforms would ensure 'they are as challenging as those in the highest performing countries in the world.'

Although there was general anxiety, which was reflected in the press and social media, the school agreed to a short-term plan of evaluator research. This would then be used as starting point on which to base future actions. Perhaps, I should have also released here the impact my actions have. Staff trusted in my request, despite other schools making changes and the consequences we would face if we reacted too late. The ethos I had created and the honest conversations we had as staff, did not form part of my thinking at that time.

I was aware that a consequence of committing to the research as opposed to implementing new practices may result in a decline in pupil results, or reduction in pupil engagement. Pupils may now revert to not liking school and any of the other millions of permeations that ran nightly through my head. It took until long after my research data collection period had ended, to realise that what I was doing all along was analysing my decisions in relations to external forces and not the external forces themselves. The external forces of policies, legislation and standards reflected in Bronfenbrenner's (1979) ecological systems theory. Whilst I confirmed with staff, that no significant changes would take place concerning

teaching and learning, there was still an expectation that staff would adapt practices to respond to day-to-day issues, and tweak plans to meet pupil needs. While the research had moved from the generic impact of testing, to how my role as headteacher could limit the impact of teaching, I still needed the school to 'run' as usual in order to analyse the data.

Davis (1998, p.1) suggests that educators are aware of the 'continuous growth of assessment within schools and the broader definition, purpose and high stakes now attached to the term'. Of the eight teacher standards, two link directly to promoting good progress and outcomes or using assessments accurately; they highlight the necessity of assessment in a teacher's day to day role. A further three standards link to teachers' high expectations, good subject knowledge and teaching structured lessons. All of which feed into outcomes.

Despite the many DfE and Ofsted documents, the term assessment to which these standards refer, has little or no clarity. Formative and summative assessments are common terms in educational settings, with a broad definition of formal external tests (summative) or a teacher's ongoing assessment (formative) based on the systems in place in a class, in a school. Given the lack of consensus as to what assessment should consist of, it is little wonder that summative assessment is the one published and that schools and teachers are judged on. Summative assessments are the one constant, the same test that every pupil will sit regardless of what they have been taught and are scored by independent markers. They are not low stake assessments designed by class teachers to suit their class and check the knowledge of their pupils. It is the national standards benchmark.

Pupils, teachers and headteachers increasingly face scrutiny of their achievements through standardised tests, 'the results of which can make or break reputations' (Davis, 1998, p.1). This creates many moral dilemmas for school leaders and teachers; in terms of teaching to pass tests in contrast to teaching what pupils personally need for the future. Guardian 9<sup>th</sup> January 2001 illustrates the struggles and dilemmas experienced by some schoolteachers,

when finding their professional commitments and personal wellbeing at odds with the demands of performance (Ball, 2013, p.60). Indeed Ball (2013) highlights many examples of how much the learning environment and the pressures from the different ecological systems can and have impacted on each other.

The pressures and constraints from the regional and national high-stakes systems, such as league tables and the need to pass tests, create pressures for teachers to work within the parameters of a school's policies and parental expectations. In addition, their own personal thoughts as to the interests of their pupils 'may mean pupils have to take second place' (Black, 1998, p.120). The introduction of the new testing regime meant that teachers who had previously managed to find the extremely delicate balance of testing and pupil need, had to refine their choices.

Within the school, back in 2014, there was little to lose in terms of making drastic changes. According to Ofsted, we were already 'requires improvement' and not achieving data targets. However, two years later, the school was 'good' and 'achieving floor targets'; it now had something to lose. Yet again, it should have been here that I realised the impact of the role of headteacher. Within 2 years, under my leadership we had become good and achieved the required standard, without testing. Though, as with many school leaders, credit is rarely accepted and any changes for the better are often dismissed and credited elsewhere.

Responsibility when things go wrong is always taken, but rarely credit.

Assessment outside of education, is not usually seen as such a burden or restraint but more a useful tool to look at ways forward and what we know or do not know, or to solve a problem. Running a health assessment, or car diagnostic for example is not seen as anything to worry about, but an essential tool to understand an issue. Sitting a driving test is also accepted as ensuring a set safety standard has been achieved. Yet in my experience, assessment within education is the ongoing debate and political objective never resolved. It is for this reason that my research focuses solely on the impact of standardised testing.

Gipps and Stobart (1995, p.36) argue that 'the idea tests can measure standards in education is one thing. The idea that testing can raise standards is another'. Yet, the implicit belief of government after government, is that introducing a testing programme will raise standards. As a headteacher, I am unclear how testing within my own setting could and has raised standards. Tests are not taught to; they are not completed every half term or even after the completion of a unit. The work in books is pitched at the same expected standard, using wording and layout similar to the tests pupils will sit and reflects any work that would be completed in a test. For this reason, my stance is that testing does not raise standards but can, in the right school context, support teacher judgements. Indeed, in recent years the examinations have been deemed in the national press to be 'easier', as the percentage of pass marks and higher grades increase. Headlines have arisen such as *Students are Being Spoon-Fed* (Daily Express 22.8.17) and *Are Exams Getting Easier: Are Schools Standards Going Up or Down?* (BBC News 30.4.20). The basis of this argument is not fully clear, but what is implied is that 'standards cannot be rising due to teacher capacity' (Davis, 1998, p7). Even when data improves, teaching is not credited with this. The tests are the driving factor; they were simply made easier. There is no reference to teaching or the effort of students.

Gipps and Stobart (1995, p.15) identify there is 'often concern with the standards achieved via testing in education, though it is rare for critics to say what they define as 'standards', and any evidence for this is anecdotal'. Testing and exams have therefore become the suggested way of 'raising the undefined standards' (Gipps and Stobart, 1995, p.26). Gipps and Stobart (1995, p.27) also suggest that 'standards' is a term used more loosely than any other education term. It is interesting to note that the term 'standard' is now used as the official level of the new testing system i.e expected standard and below the expected standard.

As headteacher, I agree with Gipps and Stobart (1995), that properly used assessment is a valuable and essential part of the educational processes. However, assessment can easily be misused and/or over-used, with both teaching and learning consequently suffering. Drummond (2008) highlights that there must be agreement on how assessments are used. She argues, that at best, assessments can be used to help teachers understand what is going on and plough this understanding back into their provision, curriculum, and interactions. At worst, testing approves the assumption that defining, or labelling what a child can do on that day (or what they cannot do) is a completely reliable predictor of future successes. One of the findings from Gipps and Stobart (1995, p.16) was that teachers viewed testing as 'useful for others.' The teachers themselves did little with the standardised test results but were happy to continue using them, as they may provide useful information to someone else. In the case of the KS2 exams, the results arrive in July just before pupils leave. There is little chance to address any issues or indeed improve the grade the pupil will move to high school with. This supports Gipps (1994) who refers to the work of Glaser (1980) that assessment must be used in support of learning rather than an indication of current or past achievement.

I am not opposed to testing and at the time of undertaking this research, simply wanted to understand any potential impact of the new pass or fail testing regime being implemented. At this stage, my role on any impact had not been realised. However, as the research evolved and the role of headteacher became clear, so too did the research questions and regardless of my reluctance to analyse my own decisions, I too became part of the research. How was I, as headteacher, impacting on testing?

What my research aims to uncover is not necessarily what assessment is used for (though I do discuss if assessment is used as intended) but the impact it has on those being assessed and how, as a headteacher, I can change any impact. Whilst I acknowledge, as a headteacher, there are many things that can be influenced; the results of testing (whether

we agree with it or not) are what pupils are judged on and this was now a pass or fail which determines staff's salaries and influenced communities. It was therefore essential that the impact of a new testing regime and my role in its implementation, be researched. The choice to implement government policy and requirements is not optional. How I implement them is solely my decision. As such, my research gives voice to its participants and highlights the impacts within their lived experience. It also aims to support headteachers in understanding their role within the ecosystem and the impact their decisions have.

It is necessary, for the purpose of this research, to have a definition of what a standards-based assessment should be. William's (2013, p.81) identifies several desirable features of standards-based assessment:

- Authentic tasks and activities: Real world problems and situations that allow the application of skills or knowledge in order to see the relevance in everyday life.
- Driver of the curriculum: Assessments reflect the body of knowledge and curriculum understanding that teachers need in order to be able to teach.
- Promoting progress of children: Pupils are aware of what is expected and what will be assessed.
- Increasing quality of teaching: Frequent assessments inform teachers what pupils have learnt, identify difficulties and in turn seeking out alternative approaches.
- Continued CPD: This is linked to increasing the quality of teaching and developing the staff knowledge to be able to have that understanding of alternative approaches.
- Increased accountability: Assessment ensures a greater accountability to ensure coverage of the national curriculum with the required breadth of learning. Teachers should relentlessly focus on moving children towards a set standard.

In relation to Williams' (2013) desirable features, I would argue the most important feature in any testing should be authentic tasks and activities, yet this is something I have not seen in the formal SAT tests pupils undertake. Reading texts used in SATS have little link to the life experiences my pupils have had. The previously mentioned list of activities pupils hoped to undertake by the end of their primary school education, involved things such as: visiting a beach, going to an airport to watch a plane, having a water fight or making a den. It does not reflect texts based on rowing boats and fishing (2017), playing in old farmhouses (2018) or understanding park wardens and how important parks are (2019). Standard English and

higher-level vocabulary are not something my pupils hear regularly beyond the school walls. Few people, even teachers, use the language of books when speaking on a day-to-day basis outside of lessons and parents are often heard when bringing and collecting their children, using sentence structure and language that does not always support school content. Writing is something they see little purpose in and rarely see people do. Baselines highlight many pupils, on entry to nursery, have never written or held a pen. Teachers need to teach wrist movement and the correct finger positions by daily exercises before starting with letter formation. Many parents today communicate with schools via text messages, social media, or telephone calls. Writing is not something often seen in homes, notes are made on phones not pen and paper. This is evident in the number of pupils that start nursery being able to use an iPad but unable to form any letters. When trying to solve a mathematical test question involving money and change, they are unsure what a papaya is and if it is a maths term, they need to know in order to solve the problem. The language and context of the questions is not simply related to maths but based on an assumption of shared experiences. The jobs pupils will need the skills for, in all likelihood, do not even exist yet. When faced with the crossroads of teaching to pass the test or give pupils life skills necessary to thrive and function as adults, teachers and school leaders are now facing a choice. This research would indirectly answer whether there needs to be a choice or whether both testing and meeting the needs of pupils can work together. Highlighting the impact of a new testing system and the role a headteacher can play will inform not only my own future decisions, but hopefully the decisions of other school leaders.

We live in an age of increasing accountability and education, more than ever before, is evaluated by its cost effectiveness and the extent to which targets and objectives are achieved. Conner (1991, p.30) identifies that the 'purpose of schools, especially primary schools, is more than preparation for the work force'. Rousseau (in Sutherland, 1988 p.14) denounced the education system of his day, believing it made children miserable on the

assumption that they were being prepared for a happy or successful life, that the skills on offer would be valuable. He argued that it is wrong and 'unnecessary to sacrifice a child's present happiness for the sake of a hypothetical future'. To some extent this is still relevant today, with the skills being tested and perhaps even the objectives being taught not relevant for the future. Tests are not really a true benchmark of preparation for the future.

This research therefore aims to understand the impact of formal assessments on pupils in primary school, from the perspective of the teacher, and how the role of headteacher influences this. The two key research questions are what are practitioners' perceptions regarding:

- How can headteachers limit the impact of testing on pupils?
- How do we understand the headteacher's role through the lens of Bronfenbrenner's ecological systems theory?

Alongside this, run three fundamental aims. Firstly, to gather and analyse narratives of teachers in order to understand the new testing regime. Then, to reflect on the school's policies and practices that impact on pupils: my role as headteacher. Finally, to have a positive impact on the experience pupils have within the school, based upon the research findings. Again, this would be the actions of the headteacher. In the background of these aims and research questions runs an undercurrent: Is assessment working as intended?

Cochran-Smith and Lytle (2009, p.118) identify inquiry as a stance, as a power notion, that recognises the 'collective capacity of practitioners to work with others to transform teaching, learning and leading schools.' They classify this stance as grounded theory of action positioning practitioners and their knowledge as central to the goal of transforming schooling. Within my research, a group of highly skilled practitioners work together to improve their own knowledge with the aim of informing whole school practice, through allowing me to be able to analyse my decisions. My goal, generally as a headteacher was to transform the school I was in.



Despite recent criticisms of educational research having no impact on practitioners, Wellington (2000, p.9) disputes this claiming that many teachers 'take on board ideas almost without notice or acknowledgement'. The research aims to inform practice at least within the context of the school. It will inform the learning environment and therefore to some extent, the work of Rosenthal and Jackson (1968), Ball (1981), Keddie (1973) and Brown (2001). It also touches on the notion of the self-fulfilling prophecy (Haralambos and Holborn 2013) and the learned helplessness identified by Seligman (1967). It will allow honest reflections of my decisions made prior to the research and would inform decisions long after the research has concluded. It would, if all went to plan, show the impact a headteacher has – good or bad. The very notion of undertaking research and wanting to do something, shows that a headteacher can have some degree of control. In hindsight, this is obvious, but at a time of a complete overhaul of not just assessment, but also testing, this was not a picture I could see. As many headteachers felt, something was being done to me and that (wrongly) I had little say or choice in this.

It will also aim to support the setting's current work on resilience, defined by Dweck (2012) alongside the works of Chandler (2014) and Lucas, Claxton (2013).

The research also aims to reflect Bronfenbrenner's (1979) theoretical framework within which it is set; in that it starts with the smallest ecological system, the pupil, and looks at the subsystems around it. It relates to how the many staff within a school fit within a set ecological system and how these various systems influence each other. The theoretical framework follows the very reasoning for the research. The policy decisions made by governments impacting on a new testing regime, impacting on the curriculum design, impacting on what and how pupils are taught before finally impacting on the pupils themselves. The interconnectivity of Bronfenbrenner's ecological systems mirrors that of a school and the education system.

Kvale (1996) highlights the purpose of qualitative research interviews as being to understand themes of the lived daily world from the perspective of the subjects, coming close to conversation, which was precisely the rationale for this research. How have, and how will my future decisions as a headteacher, genuinely impact on those living with the consequences of my actions. A headteacher is able to leave, but the pupils and community will still live under the actions they created.

A further coincidental element of the research would be to address the perceived parental concerns regarding SATs and testing.

Families are becoming increasingly worried about the number of practice papers and revision classes that y6 children face during the Easter holidays and that the whole of year 6 solely focusses on SATs (Bushby, 2018). Bradbury (2019) also highlights that 61% of parents believe there to be too much testing and that schools spend too much time focussing on maths and English

While findings would not be relayed directly to parents, they would hopefully see the implications.

### [1:3 A Brief History of Educational Assessment](#)

In order to look at the impact of testing in today's education system, it may be helpful to first briefly discuss the evolution of education and the education system alongside the chronosystem in which it is embedded. If we are truly to understand the impact of a new testing regime, some knowledge of how it evolved may support understanding.

"There is a relationship between how a country runs its education system and the way it runs its economy." (Winch and Gingell, 2004, p.11).

The new assessment regimes this research aims to understand are rooted within a need for the economy it serves. In 2015 (the year before a new testing regime and the year

education was a government focus) employment in England was at its highest since records began yet the countered deficit rose to a record high, the highest in the developed world (BBC News 31.3.16)

Winch and Gingell (2004) raise the issue that until 1999 the British education system had no explicitly formulated aims, perhaps this remains an issue for education today, as there are still many stances, viewpoints and opinions as to what the aims of education are.

Sutherland (1988) argues a part of basic education long taken for granted is the teaching of the three Rs and this is certainly the focus of SATs, published data and world league tables. These 'core subjects' are culturally embedded within the British education system.

My own viewpoint has been made clear; education is more than testing and should be, in primary schools at least, about developing a love of learning. I agree with Dweck that "An assessment at one point in time has little value for understanding someone's ability, let alone their potential to succeed in the future" (2012, p.29).

However, I acknowledge this is one opinion in a sea of many. Winch and Gingell (2004) identify the main reason, prior to the formulation of aims, and still the main purpose for education today is to provide a free, basic education, combining the core skills of literacy and numeracy for the working population, whilst supporting the existing social and political order. My research specifically highlights how testing impacts pupils from areas of high deprivations.

It should be acknowledged here that many researchers have long argued over the purpose of education and thus, in turn, the role of assessment and testing. This ranges from the transition of norms and values, to teaching skills specific for future occupations (whether that be in an industrial or cyber society), social stratification and the key to national economic progress.

Winch and Gingell (2014) suggest the education received today, through the implementation of the national curriculum, is still set with a need in mind. Sex education became compulsory in 1993 after several high-profile deaths linked to AIDS (acquired immunodeficiency syndrome) in the early 1990's raised the profile. This was also just one year after AIDS was the biggest cause of death in US males aged 25-44. British Values were introduced in 2014 following terrorist attacks in America and London and the recent academic focus has swiftly focussed on reading and phonics following a drop in the worldwide league tables. Following the release of MQ Transforming Mental Health 2018 data that three in four mental illness start in childhood and 10% of school children have a diagnosable mental illness, mental health/well-being lessons became compulsory in September 2021, though schools were recommended to teach this in 2020. Included in this was the also teaching of 'healthy relationships' following a national campaign on controlling or coercive behaviour becoming against the law in 2015. More recently, following a dental shortage, school are required to brush teeth daily. Interestingly, none of these are tested, yet are sold to schools as 'vital' for society. The needs may change, but the basic premise of a national specific need impacting on what is taught is still there. Durkheim saw the main function of education to be the transmission of norms and values (Haralambos and Holborn, 2013, p.663). The need today has simply now shifted to competition and global competition. Ball (2013, p.3) refers to David Cameron's speech in Norwich, 9<sup>th</sup> September 2011 where he stated:

We want to create an education system based on real world excellence, with a complete intolerance of failure. Yes, we're ambitious, but today you've got to be. We've got to be ambitious if we want to compete in the world. When China is going through an educational renaissance, when India is churning out science graduates. Any complacency now would be fatal for our prosperity. (Ball, 2013, p.3)

In this speech, failure is not to be tolerated. Yet the new primary school testing systems creates just that – pass or fail. This speech perhaps also alludes to the real reason for

testing - world league tables. Ball (2012) also argues that global organisations are involved in global educational policies.

To impact on the world league tables, any education system needs the best students. Within the current British education system, the best students are identified through a selection process and a conveyor belt system of testing from the beginning of their educational journey. This concept, as will become clear is not new and instead is rooted in the 19<sup>th</sup> century.

The first formal qualifying exams were designed to determine competence for the medical profession in 1815. This was followed in 1835 by written examinations for solicitors and in 1880 for accountants, (Gipps and Stobart, 1995). Whilst these are not educational assessments, they were introduced to address the growing and changing needs and social structure of society and arguably to protect the higher status professions. As a capitalist economy flourished, the need for middle class workers in managerial positions developed and therefore 'the need to control those deemed suitable and competent' (Gipps and Stobart, 1995, p.2). Ball (2013) develops this to argue that the state education of the 19<sup>th</sup> century was as a direct need to manage the new working classes and accommodate the social and political aspirations of the new middle class. It began the movement away from professions being determined by patronage and nomination, to ability and academic achievement. It signalled the development of examinations in response to 'particular needs and requirements of the time' (Gipps & Stobart, 1993 p.3). Again, the link between need and educational content is clear. A link that continues with a new testing regime being implemented, alongside a slip in world league tables.

Torrance (1995) credits the impetus for educational change in assessment with political interest, ensuring school accountability through testing to raise standards. Broadfoot (1995, p.9) identifies that educational assessment, first used in the nineteenth century, was in

response for a need to provide a 'ladder of opportunity' into the expanding industrial economies of the era. However, Bourdieu (1997, p.164 in Nice Richard), argues that a consequence of society's greater reliance on tested ability is a more 'rigid class structure that will make it more difficult for individuals to improve on their social position above that of their parents'.

Following these changes, the demand for university places from the middle classes soon increased, and in the 1850s Oxford and Cambridge introduced examination boards, with London and Durham introducing entry exams. By 1855 the Civil Service had also introduced entry exams, and in 1917 formal examinations at the end of secondary schooling were established. It also, through these changes and the creation of assessment, created the first market of educational skills and wealth. Whilst it did provide some movement, at this point education reflected the social class system of society at the time. According to Torrance (1995):

Educational change does not take place in a social or political vacuum of course and much of the impetus for change in assessment has come from political interest in rendering schools more accountable and using tests to raise standards. (Torrance 1995, p.3)

The 1926 Hadow Report's many recommendations concluded that almost all children were eligible for secondary school and discussed the need for differentiation and a joined-up thinking of teaching. It identified the need for primary school to end at 11 and that 'normal' children should proceed to 'some form of secondary education.' This new conflict of social class and cultural capital would now need to be addressed. This in turn led to the creation of the 11+ in 1938, where children would be classified by ability, based upon examinations. This new exam tested four areas: verbal reasoning, non-verbal reasoning, maths and English. It is also worth noting that the SAT papers pupils sit today are reading, grammar, punctuation and spelling, maths, and maths reasoning. Many questions within the maths

11+ paper is very closely reflected in the SAT papers still in use. Today's grammar test, a relatively new paper, also reflects the questions asked some seventy years earlier. While the accountability of testing has increased, what is tested has changed very little.

It would make sense, at this point, to raise that educational assessment today is a broader model and an assessment culture is embedded, with assessment in some form happening in most classroom lessons. However, it is still the summative assessments (SATs) that are valued more by governments and as already identified have been the preferred measure of 'suitability' for some time. The value attached to a formal pen and paper procedure has not really progressed since testing began many centuries earlier.

The Year 6 SAT results achieved in today's educational climate determine starting groups and subjects for the next 5 years of secondary education. This one-off test impacts on potential options and choices available for the next life stage of each individual, just as it did during the 11+, entry to the medical profession and even to Spartan children. As the pressure to get the grades increases, schools inevitably find ways to 'shoe horn students through tests, often at the expense of deeper development' (Lucas, Claxton and Spencer, 2013. p.12).

The impact of testing and the pathways it creates do not appear to have significantly moved. Despite being packaged as a new testing and assessment system, what is tested is not new. Simply the label given is the only new element. This label, and its impact, are the heart of my research.

The era of the eleven-plus was considered the heyday of standardised testing, by the conservative government of the time (Gipps and Stobart, 1995, p.64). It consisted of a battery of tests, which pupils were prepared for by taking a series of tests regularly. The percentage of pupils achieving a grammar school place was the criterion by which parents, and teachers judged the 'success of a school' (Gipps and Stobart, 1995, p.64). The percentage of pupils

achieving at SATs, GCSE, A level or degree is still the measurement of an educational establishment. The prestige attached to testing had continued.

The memorandum Education after the War (The Green Book) became the basis for the next White Paper. The White Paper of 1943 suggested that secondary school allocations be based on a teacher report with the aid of intelligence tests. The concerns of Ernest Bevin and Cyril Burt regarding intelligence tests meant that the passing of the 1944 Education Act increased this selection process and permitted each LEA to decide on their own assessment procedures, following the creation of the tripartite system of schooling: Grammar, Secondary Technical and Secondary Modern. This idealistic system, based on the assumption that all pupils (regardless of background) were entitled to an education matched to their specific needs, aimed to meet the needs of the economy through creating technical workers, intellectuals and general workers and focussed very clearly on the skills the country needed rather than the cost and aimed to address the demand for social reform started before the outbreak of the war. This is arguably the start of educational assessment now also being used for social need. I would also suggest that this system of allowing LEAs to create their own assessment system was simply regurgitated, again for social need, some 60 years later with the introduction of the new SATs pass or fail and each school being able to create its own assessment system. This will be explained in some depth in subsequent chapters.

By the 1960s it became clear that the tripartite system was not developing talent and the desired skills for economic growth were not being provided (Haralambos and Holborn, 2013, p.677). With grammar schools receiving most of the funding, it reinforced the assumption that this was the better system. With The House of Commons Library (January 2020) identifying 165 grammar schools still in existence today, this highlights the prestige attached to this one strand of the system. As many grammar schools continue with entrance exams, testing is still seen as something 'high status' for those wishing to enter the grammar school



system (Haralambos and Holborn, 2013, p.677). Alongside this, only one in ten pupils at a secondary modern sat a formal qualification, with many leaving school before sitting O Levels (Rosen and Williams, 2013). The equality and social mobility hoped for in the 1944 Act was far from achieved (Haralambos and Holborn (2013). The solution: comprehensive education for all, aimed to address this. However, following the introduction of comprehensive secondary schooling, the overall demise of the 11+ in 1965 became inevitable. Without the constraints of examinations and teaching to the tests, the curriculum widened and more discovery learning opportunities were now available. This is something I replicated within my setting. We removed the unnecessary testing and broadened and balanced the curriculum. Gipps & Stobart (1993) identify that there is no doubt the changes that took place to broaden and open up primary education would not have happened if the 11+ had still been in existence. It could be argued by educators today, that true freedom within classrooms is only possible without testing. Whilst my research is focussed on the impact of testing of pupils, it also looks at what classrooms are 'like' to understand impact. It does, by default, look at classrooms and teaching to potentially answer this question.

By the late 1970s, many local educational authorities had already reintroduced standardised testing, following the Black Papers of 1975 and 1977 and the concerns over teaching and learning made famous in James Gallagher's Ruskin college speech, highlighting the need for accountability in education and signalled a greater state involvement. However, as results were not required to be published; it had little impact on teaching. Unlike the 11+ which had been a series of tests with pupils prepared for it by completing various tests throughout the year. Teaching during the 11+ era, was solely preparation for the 11+ test. While testing was back on the educational agenda, the high stakes attached to it, at this particular time, had not yet been established. This would come in the 1980s and remain part of the educational environment.

By the 1980s, educational changes reflected the government concern for market forces and competition, reflecting the desire to use assessment to 'encourage lower ability and less motivated pupils' (Gipps & Stobart, 1993, p.9). With the 1988 Education Reform Act and the introduction of the National Curriculum, the government aimed to motivate students by engaging them in their own learning and used assessment to encourage this. I would suggest that assessment is not necessarily a motivation tool for pupil engagement and will address the strategies I implemented to allow pupil engagement. Macintosh and Hale (1976 cited in Horton, 1992, p.16) suggest that motivation may become a seventh possible purpose for assessment, following from their much quoted six possible purposes. With educational policies highlighting that education should mainly be concerned with promoting economic growth through improving skills, competition would be the best way to achieve this (Haralambos and Holborn, 2013, p.687). A way to ensure competition, would be through parental choice of schools, made through the publication of inspection reports and results.

The Education Reform Act of 1988 introduced open enrolment, giving parental choice of schools, formula funding, a national curriculum and the testing and assessment system still in place today. West (1994, p.109) identifies that few studies examining parental choice have been carried out. In the few studies analysed, happiness, proximity and other siblings in attendance are the common factors. Bradbury's 2019, *More Than a Score* research highlighted that only 12% of parents felt a school should be measured by tests, with only 25% stating that a schools' results influenced their preferences. The overwhelming majority, 77% of parents, chose a school place based on if teachers cared and inspired their child. While this research field is limited, I would suggest schools are led to believe that league tables are not just for governments but also a useful tool for parents choosing their child's educational future. Yet, the reality is that few parents choose state funded school for their child solely based on academic results.

The DES 1987 described the introduction of the national curriculum and assessment, as a proven and acceptable way of raising standards. With this introduction, teaching and learning became a 'topic, teach, assess system' (Gipps and Stobart, 1993, p.101) However, despite reading the consultation papers, I am still to understand the 'proven' method or how it is acceptable.

The national assessment programme outlined in the 1988 Report of the Task Group on Assessment and Testing (TGAT), that pupils be assessed against attainment targets by their teachers and at the ages of 7,11,14 and 16. Testing was now back, and not just starting at 11years old. The TGAT identified five reasons for the implementation of the statutory assessment tests:

- Formative. Providing information on pupil achievement allowing teachers to adapt and plan next steps.
- Summative. Providing overall achievement information.
- Evaluative. Providing aggregated information on classes, school and authorities.
- Informative. Providing information to parents.
- Professional development. Providing teachers with the ability to review their own work. (Hutchinson & Schagan, 1994, p.2)

As headteacher, all five reasons appear to be justifiable, yet how the information is used creates many issues. In line with the research on parental choice briefly mentioned earlier, I feel (as a parent) the school's performance as a whole has no impact on my education choices for my children. Whether they are happy, and enjoy school, are my main benchmarks. However, I fully accept that some parents may be influenced by league tables and the prestige of these schools. Producing league tables under the guise of supporting parents to make more informed choices can be counterproductive. Within this research setting, staff spend vast amounts of time trying to explain systems and results each time they are changed.

Her Majesty's Inspectorate of Education in their Her Majesty's Stationery Office (1988)

publication suggested that assessment of pupils' work have 4 main purposes:

- To provide pupils with an indication of their progress and achievements.
- To help the teacher identify strengths and weaknesses and therefore adjust teaching.
- Enable pupils to evaluate ways they can improve.
- To show others what standards of work have been achieved.

This HMI publication gives a more open account of why assessments were introduced, though noticeably, left the standards element until the end. None of the 4 purposes identify testing as the method of assessment yet testing appears to have become a synonym for assessment.

It is hard to break the hypnotic spell of 'standards', as defined by examination success. Despite an increasing barrage of fine words and good intentions, it is the examination results by which schools and students' performance will be ubiquitously judged; by politicians and by the media. (Claxton, Chambers, Powel, and Lucas, 2011, p.22)

William (2008, p.123) highlights assessment as a 'central process in education'. It allows teachers to assess what has been taught and adjust. Parents can make decisions on potential schools and policy makers use assessments to provide information about schools or curricula and adapt to changing needs of society. From this definition it would suggest that assessment has been taken to mean testing, as parents' choices would be based on league tables or published school data; all of which are based on statutory test data. However, William's ideals do reflect the original HMI's four reasons for assessment in their HMSO publication in 1988, some twenty years early.

Conner (1991) suggests that the proposals of the TGAT and HMI created a primary education with selective examinations, created the teaching professions desire to achieve results dictated their classroom behaviours. The government's response to TGAT gave far greater emphasis to the role of pencil and paper testing, significantly reducing the teacher

assessments based on classroom activity. In July 1989, £14 million of contracts were awarded to develop SATs in the core subjects (Ball, 2013, p.132). Formal testing therefore became the focus of the school assessment system.

Gibbs and Stobart (1995, p.14) take this further and suggest that the purpose of national assessment is to ensure compliance from teachers to teach the National Curriculum. They highlight it is more than simply testing and refer to the little impact the eleven plus had on teaching (1995, p.65).

In conclusion, the impact on testing spreads beyond the pupils themselves and can influence individual classrooms and whole school policies and ethos. Decisions about what schools should teach depends on the desired outcome of the educator, all of whom have differing opinions (Sutherland 1988, p.110). Torrance and Pryor's (1998, p.131) position that assessment does interact with and have an impact on learning and motivation (although the impact can be both negative and positive) is the starting point for this research.

It should also now become clear that policies, classroom ideals and beliefs (to name a few actions) are linked to testing and each of these sits within its own ecological system. The interaction between these ecological systems has to also be unpicked in order to genuinely understand the impact of testing on pupils.

#### 1.4 The Introduction of SATs

Formal assessments take the form of Standard Assessment Tests (SATs), with both assessment and test in its title. It is possibly the origin of the misunderstanding that the word assessment means test. As with the need to look at the history of assessment before looking at testing, it is necessary to now look at how testing and in particular SATs have been developed

In the early development of the KS1 SATs, it was suggested that as many as possible attainment targets were covered (Gibbs and Stobart, 1995, p.73). Given there were thirty-

two attainment targets, and the test was not pencil and paper, this soon proved problematic and caused significant time and organisational issues (Gibbs and Stobart, 1995, p.73). By 1991 the SATs were still largely active and infant based practice but targeted fewer attainment targets, currently 9. Though reduced, the very nature of group and practical aspects, meant the test still took a considerable length of time to complete (Davis, 1998, p.7). This prompted a threatened boycott by many teachers, which was narrowly avoided following a ballot (Sainsbury, 1994). In response, the government announced that by 1992 the test would be less active, shorter, and moved to a more pencil and paper structure. By 1993 further changes were made to reading and spelling testing, the SATs had become a mixture of tasks and more traditional pencil and paper testing. It should be made clear at this stage, that the boycott was not in the hope of a written test being created and something 'easier', simply a protestation that the system was not working (Sainsbury, 1994). This has not happened to the future testing systems, when staff have also raised concerns that it is not working. In these cases, testing has remained, and the pencil and paper testing has not been negotiable.

The assessments of KS2 testing were to be of three different levelled abilities, and tests not tasks, with pupils entered for the correct ability test. These KS2 tests followed later to allow pupils to have covered the full national curriculum before the SATs were administered. By 1993 the first national pilot was implemented. This then led to a NUT proposed teacher boycott of the assessment system due to claims of excessive workload and 'doubts about the nature of some of the tests' (Sainsbury, 1994, p.6). This was stopped by Sir Ron Dearing (now Lord Dearing) leading a review and recommending the National Curriculum be slimmed down. It should also be noted that the criticisms voiced by teachers were mainly concerned with the mode of implementation and not the tasks themselves Teachers were mainly concerned with the length of time the test would take and the impact on teaching time that would be lost (Broadfoot, 1995) testing was again, not opposed.

By 1995 less than 50% of schools were achieving the required target grade. Despite the lack of success, the newly appointed government of 1997 announced new higher targets of 80% within the next five years. While results did steadily increase, government targets in 2001 and 2002 were still missed, some 10 years after the original implementation. A further boycott was proposed in 2004 for very similar reasons to the previous proposals, but again insufficient numbers were gained to strike. Following this, Wales abolished SATs testing, and Scotland moved towards pupils sitting the tests when teachers felt them ready to do so.

This reflects Torrance and Pryor's (1998, p.10) stance that 'assessment is inevitable'. It changes form, assesses different things yet always remains. Whilst the tests themselves are continuously discussed and evolving, what is not discussed is removing them, simply the connotations associated with them. In contrast Holt (1969) states:

Let me not mince my words. Almost all educators feel that testing is necessary, or useful or even excusable. I disagree. At best, testing does more harm than good; at worst it hinders, distorts and corrupts the learning process. Holt, (1969 p.51)

Holt paints a very bleak picture of educational assessment yet offers no real alternative.

Perhaps Jeremy Corbyn's speech in April 2019 at the National Educational Union's Inaugural Conference, where it was announced SATs would be scrapped should Labour come to power, provides an alternative. Certainly, the recent examination debates forced to the forefront by Covid-19 Pandemic (202-2022) will at the very least mean testing should be reviewed.

Conner (1991, p.41) highlighted that the 'disadvantages of assessment system at primary school level outweighed the positive'. arguing that they would become more in line with secondary education under the influence of exam boards, ability banding, formal teaching, and stricter subject boundaries. Conner (1991) goes on to raise that primary education within a selective examination-orientated curriculum was influenced enormously by the fact

that teachers' intentions, teaching styles and even classroom organization were dictated by the need to get results.

98% of Headteachers in a YouGov (2019) poll believed SATS put unnecessary pressure on teachers and 96% believed it to have a negative impact on pupil well-being (Rigby, 2019). This will be explored in depth in later chapters.

Having focussed so far on the creation of SATs and testing, the following section will focus on the use and purpose of testing. The criterion of formal assessment superseding teacher assessment highlights the shift in the purpose of national curriculum assessments. There was a move away from a formative approach helping to provide useful feedback to an individual, to the current more summative means of evaluation of school and local authority performance by the means of aggregated data (Gipps and Stobart, 1995, p.91).

Filer (1994, p.48) highlights that it is generally recognised that Conservative governments since 1988, have 'little faith in the reliability of teacher assessment in comparisons to national standardised tests.' This is similarly supported by Hannon (1990) and the view that government policies appear to distrust teachers.

In addition, Broadfoot (1995) raises the concern that SATs do not provide very well for any of the purposes identified. They are not frequent enough and not sufficiently integrated into classroom routines to provide summative and evaluative information. Since they occur at the end of key stages and identify only broad levels, it would suggest their sole purpose is to compare one school with another. The changes made in this decade, moving away from the creative education post 11+, back to selection and teaching to tests supports this. Harlen (2008, p149) believes the current practice is often driven by the pressures of being 'responsible for ensuring that pupils make regular and visible progress'. This stems from the national curriculum levels at 7 and 11 being used to publish value added in 1993 and then secondary school value added in 1999.



The SATS administered today have moved away from the practical and activity focused origins to a more 'formal and time-consuming system' (Davis, 1998, p.7). Ironically, concerns raised over the time taken to implement hands on approaches were replaced by full weeks of pencil and paper tests. With testing today, year 6 have a week of tests; one per day.

Year 2 have a month in which to undertake their tests, Year 1 have a full week in which to complete their phonic test and Year 4 have 2 weeks in which to undertake their multiplication check. Reception children have to be 'baselined' within the first 6 weeks of starting school.

In 1997 a Labour government continued with previous Conservative plans (rather than the one size fits all school incorporating special schools) with the testing and assessment regimes remaining in place. As Johnston and Costello (2009, p.145) state, 'What gets assessed is what gets taught.' This highlights the stronger focus on teaching to the tests.

Gipps and Stobart (1995, p.910) identify the 1988 Education Reform Act as 'generating a new climate for assessment; one that reflected the governments concern with market forces'. They highlight that whilst the national curriculum assessments would be used to partly inform on an individual's progress, the main use was one in which to judge the performance of schools and local education authorities. League tables were introduced shortly after in 1992, followed by the 2010 White Paper: The importance of Teaching, to encourage competition (thus by extension a winner and loser system, as the words 'league table' and 'competition' imply) and therefore raise standards and provide incentives and rewards for success. Ball (1994) concluded that parental choice, published results and the implementation of new policies made education more market orientated and identified that in any market there are winners and losers. This is reflected later in the research of Harold, Aitken &, Shelton (2007) and later still with Ferretti, Ganley and Kofler (2019) all of whom make links between parents and their fear of academic failure. Understandably, parents

want the very best for their child and this will be in the form of success. Choosing a school that is deemed successful, is part of that choice and reduces the risk of failure.

It is worth noting that while school leaders have quite severe consequences to their placing within league tables, education secretaries seem except. The pressure created by a league system has been well documented with senior leaders even losing careers following the alteration of pupils' answers in order to pass the test. The STA reported in 2017 that 599 schools were investigated for maladministration during SATs, increasing from 526 in 2016. A spokesperson for More than a Score, which represents 18 education and parents' organisations, said:

If you create a toxic system where schools and teachers are judged on the results children get in SATs, then of course it is no surprise that some schools and staff will cheat to avoid being downgraded by Ofsted, academised, being paid less or even sacked. (Busby, 2018, para 8)

In addition, Bushby's (28<sup>th</sup> November 2018) Independent Newspaper article notes:

When primary school leaders are treated like football managers who can be disposed of if their children 'underperform' in the SATs, grade distortion is hardly surprising. Primary school teacher pay is also linked to their students results, and govern reports of teachers using food banks and sleeping in their cars because they can't afford rent, teachers are forced to teach to the test and 'game' the system just to survive. (Bushby, 2018, para 6)

It could be suggested that this period of educational change narrowed education to a purely economic function, solely focussed on the UK being able to compete effectively in a global market (Ball, 2013, p.1). Testing became the check mate in the global world chess game and headteachers, teachers and pupils the chess pieces, each with their own disposable value to win the game.

The coalition government of 2010 attempted to promote more independence and freedom for headteachers and to make schools more accountable to parents, rather than the local

authority. A new curriculum was introduced in 2014, with English and maths following in 2016. This replaced previous curricula and gave freedom on how to teach but not on what to teach. Despite this new freedom, the primary assessment committee found that teachers were focussing deliberately on English and maths to ensure pupils passed the SATs, while arts, science and humanities were neglected. Not only did this report identify that the current high stakes system can negatively impact on teaching and learning through narrowing the curriculum, it also identified that it negatively affects teacher and pupil well-being (Turner, 2017).

In 2014 the DfE acknowledged that the current system of levels used to report attainment would be removed and not replaced, with schools free to create their own assessment system. As already shown, authorities creating their own assessments systems had not been successful when implemented years earlier. Yet individual schools were to now make their own systems a success. This DfE acknowledgement did not however remove formal examinations but allowed schools to simply create their own internal assessment recording system for use in other year groups, with SATs data remaining at the end of Key Stage 1 and 2.

Black (1998, p.16) stipulates that the overall outcome from changes in assessment to 1994, have left us with '5 parallel systems of testing and assessment'; with these systems still reflecting the original justifications for assessment. It should now be apparent that there have been many changes to education, but the removal of formal assessment has not been one. Gipps and Stobart (1995, p.14) stipulate the 'purpose of a national assessment is to make sure the national curriculum is taught'. To some extent this is true, but the national curriculum also includes: history, art, Modern Foreign Languages etc, none of which are tested at primary school level. However, I would suggest not all subjects are taught with the same level of coverage and Ofsted (2020) have also highlighted a narrowing of the

curriculum. Davis (1988, p.167) implies that the UK is already showing signs that 'the teaching profession are losing their status as reflective practitioners'. Instead, they are shifting to the role of technicians simply implementing government policy not only on the standards to be achieved but also on the methods to be employed.

Swaffield (2008, p.xi) identifies assessment as an integral part of learning, teaching, schooling, and education, which, in recent years, has become 'more pervasive and prominent; becoming a focus of attention for pupils, parents, practitioners, researchers, policy makers and the media'.

While attempts are being made to use assessment data to show impact via the value added and contextual value added by schools and teachers. This data is not nationally published in league tables. As a headteacher, I would suggest this data is the truer reflection of the standards of teaching and impact. Yet it would appear that, it is not these standards that the league tables, policy makers or governments want. Progress would appear not to count, unlike attainment and what pupils can do at that fixed point. Perhaps progress would indicate the varied starting points pupils have.

Assessment now has several functions including: the diagnosis of causes of young people's success or failures, the motivation to learn, the provision of valid and meaningful accounts of achievements and the evaluation of teaching. The emphasis has shifted from 'assessment for summative purposes and much more kudos is now laid on formal assessment; that is the use of information gathered to improve the current education process' (Brown, 1992, p10).

Murphy and Torrance (1988, p.10) highlight the common origin and occasional meeting ground for educational and political debate is the 'never ending discussion on standards.'

In the White Paper: The Importance of Teaching (2010), the foreword written by the coalition government of David Cameron and Nick Clegg highlights the view that league tables are the country's ranking – they are the driving force of education change. It reads:

So much of the education debate in this country is backwards looking; have standards fallen? Have exams got easier? These debates will continue, but what really matters is how we're doing compared to our international competitors. That is what will define our economic growth and our countries future. (2010, p.3)

This may be the acknowledgement that the changes which were to come may not necessarily be for the benefit of pupils and schools, but to boost standing in national league tables. My research examines the voices of those impacted by the changes and offers a narrative of the tightrope headteachers must walk.

It would be unfair not to mention that there have been some attempts to move away from testing. Marlow, Norwich et al. (2014, p.413) discuss at some length, the comparisons between testing and teacher assessment. They highlight that between 2008 and 2011 the government invested £150 million to support schools in developing a strategic approach towards classroom assessment, in the aim of 'securing accurate good practice'. The Assessing Pupil Progress (APP) initiative was a formative part of this and having been developed over 5 years aimed to give a common language for talking about pupil attainment and create a common reference point regarding standards. The grand aim of APP was to replace the current assessment arrangements, using an evidence-based system over the course of the school year. Ofsted's 2011 review of APP concluded it had improved teacher accountability, increased consistency, secured accurate assessment and had strengthened the existing systems. It also suggested that moderation within schools was also improved. However, Marlow, Norwich et al. (2013) explain it is unclear how these findings were established. Despite initially appearing to offer a glimmer of hope towards the removal of testing, APP was never formally introduced and eventually removed from schools, remaining only ever a formative tool. It is also worth noting that Marlow, Norwich et al. (2013) identified the cost of external assessments in 2013 in primary school alone was £243.1 million. Given the significant sums involved with testing and assessment, it would be a fair conclusion to

imply that if teacher assessment was seen as valid and accurate by governments, they would not be spending such sums on testing. It may also have arguably reduced the number of U-turns and unrest during the 2020 school closure results.

Assessment in schools is not a new concept. What has changed is the role and style of assessments, and for the current educational assessment system to be researched, it is important that the context and previous changes were first highlighted. Filer (1994, p.51) raises that assessment results have 'increasingly been used to fulfil different purposes; from monitoring standards and progress, accountability and to make comparisons.' This is supported by Torrance (1995, p.5) who identifies that the 'National Curriculum assessment is hard to summarise since it has been constantly changing since its first introduction in 1988'. This is also supported by the work of Davis (1988), referred to by Winch and Gingell (2004) and his sustained attack on criterion-assessed assessment. Davis (1988) argues that this form of assessment can never be valid in that it can never accurately measure what it sets out to measure, namely the amount of knowledge acquired. Formalised systems are unable to capture knowledge effectively. The rich knowledge, which is the goal of education, is not assessed effectively due to narrow and restricted procedures only capable of measuring procedural knowledge (as has been the case for well over 30 years).

Having highlighted the origin of testing and the purpose of testing, I will now briefly focus on the impact. As already stated, this research aims to focus directly on the impact of a new testing regime and the role of a headteacher within this. It would therefore be apt to look at the current educational landscape regarding the impact of testing.

Eight out of ten primary school leaders (82%) who took part in The Guardian's May 2017 survey, reported an increase in mental health issues among primary school children around the time of their exams. This report also highlighted that two-thirds (68%) of primary school leaders believed changes to the curriculum and school performance pressures had a

negative impact on pupil's mental health and wellbeing. Perhaps the most hard-hitting aspect within this report is:

One child was said to have lost all their eye lashes due to stress, while others worried about academic failure and some had to be comforted after being found sobbing during their exams. (Weall, 2017, Guardian Education)

Further headlines, which will be discussed in later chapters, all highlight the impact of testing on mental health.

Shockingly some pupils (4%) were swapping their usual breakfast with energy drinks or coffee to get exam ready. (O'Grady, 2018, para.5)

22% of students saying that studying for their SATs exams felt it drove them to no longer enjoy learning (O'Grady, 2018, para. 6).

SATs risk giving children mental health issues, the education select committee has warned. (Turner, 2017, para 1)

Since 2014 there has been a 74% increase in the fear of academic failure and 55% increase in depression among pupils. One child was said to have lost all their eyelashes due to stress and others were found sobbing in the toilet. (Weall, 2017, para. 3)

Three in four mental illnesses start in childhood, with 10% of school children having a diagnosable mental illness. 10% of children aged 5-16 have been diagnosed with a mental health problem (Weall, 2017, para. 6)

Alongside media headlines, there were also union statements such as:

The Right will say pressure is good for children because it prepares you for life. But that is not true. Resilience isn't built up by failure. Too many of our children and young people at school are experiencing failure after failure after failure. This is just unethical and immoral. There has been an explosion in child and adolescent mental health and I think much of that is related back to the pressure we are putting them under when they can't stand it. Dr Bousted (Joint secretary of the NEU) (Busby 2018)

In contrast to this, Jerrim (2021) states that national tests do not impact enough on well-being to be abolished. Interestingly, these findings were in the middle of a Covid pandemic and during times when testing had been momentarily paused. What is missing for the vast research on testing is evidence that pupils benefit from testing in terms of their academic development and lifelong learning

Increasingly, assessment and testing will continue to be used to fulfil different purposes than has previously been the case. Alongside the national curriculum, assessment systems are being used to 'monitor and compare progress of individuals and accountability of teachers and schools' (Filler, 1994, p.51).

Broadfoot (1995, p.10) raises the notion that 'due to the high-stake environment, test results can determine life chances.' Torrance (1985) goes further and raises that the pressure on schools to maximise test results is clear, though at what cost to 'real' education is a matter of debate. Given the high stakes of assessments (not just on pupils, but also the teaching profession) local authorities and governments, the impact of the 'new' assessment system should at the very least be investigated, if only to inform systems within this research setting.

As early as 1984 Frederiksen had raised concerns that tests would greatly influence what is taught, (Gipps, 1994). 1993 Gipps and Stobart (1993 p.37) raised concerns regarding the introduction of testing:

- It can focus attention on the subject being tested and therefore increase the teaching time of these subjects.
- It may result in teaching to the test, which is likely to result in improved performance in tests.
- Curriculum backwash may occur, in that test content may also impact on teacher practice other than teaching to the test.

These concerns, especially a narrowing of the curriculum, are still prevalent today and indeed mirrored my own concerns when reading of the new pass or fail system. The 2017 House of Commons Education Committee identified itself that 'during 2016, primary schools



faced a challenging year as new national curriculum assessments were rolled out across all state schools and 'expected standards' were raised' (2017, p.4).

As Torrance (1995) highlights, there is a pressure to achieve results and testing will have some form of impact, McAllister (2018).

This chapter has outlined the context of my research and indicated my rationale for selecting the area of study. Although I present detailed information concerning the methodology in Chapter Three of the thesis, it is pertinent to inform the reader that the data presented in this qualitative research is drawn from various members of staff from the school which on which this case study is based. During the academic year 2017/2018, the Year 6 teachers were asked to keep a reflective diary, and in addition two participant group interviews were held: one at the start of the year and another at the end. By triangulating the content of the diaries, the interview transcripts and material uncovered in the literature review, it has been possible to systematically analyse the results and use the findings to create a valid contribution to research in this area. The implications of my findings are presented in the final chapter, the findings themselves are presented and analysed in Chapter Four, and Chapter Three presents details of the methodological considerations. However, the next chapter, Chapter Two of my thesis, provides the Literature Review which underpins the context and analysis of my research.

## Chapter Two: Literature Review

### 2.1 Introduction

This chapter is split into 2 sections; a literature review before the research and a further literature review following the research. In this first section I address four main areas: a review of the policies underpinning the new testing regime, a review of teacher as researcher, a review of the systems already in place within the research setting, and conclusions drawn from each area within the literature review. It also links the literature review back to the research questions.

This section critically explores the rationale and implementation of government policy and investigates the growth mindset culture introduced within the research setting. As this research is from a teacher perspective, a review of this area is also conducted. For this reason, the literature review is divided into four sections, part one focusses on government policy, the second on Dweck's 2000 and 2012 Growth Mindset theory before moving to teachers' perspectives within research. Finally, it then focuses on key definitions and the work underpinning the research setting.

It was difficult to conduct a substantial literature review on an assessment system that had only just been implemented. One of the main reasons for conducting this research on this assessment system, was indeed to explore the unknown impact. Since this research and this literature review were undertaken, some research on testing and its impact has been carried out. Jerrim (2021) provides a key research piece, however, even this research is focussing solely on mental health and not comparing the impact of the new testing regime, around which this research is focussed. There has also been a Covid pandemic and the closure of school and temporary pause of testing.

In order for there to be no grey areas within this review, it is also important that some key definitions are firstly identified. Assessment, as already alluded to and will become clear

later, I argue, is now a synonym for testing. However, within the context of this literature review and the policies it is central to, the DfE (2018) define assessment as a way to monitor national standards for literacy and numeracy in order to determine whether standards are declining or improving. Assessment can be both external and internal and exam or non-exam based, thus, creating a definition that can be interpreted in many permutations. However, it is also acknowledged that the DfE use certain terms in which that differ from their use in everyday language.

Testing within this review and within the review of the research refer solely to the formal SAT tests all pupils in England sit in Y6 and are those identified by the Standards and Testing Agency.

Remaining within the national policy arena and within Bronfenbrenner's macro system of policy makers, wellbeing is also defined and success criteria towards mental wellbeing within primary school are highlighted in the 2021 DfE Physical Health and Wellbeing Statutory Guidance p4 highlighting that all pupils should know:

- that mental wellbeing is a normal part of daily life, in the same way as physical health
- that there is a normal range of emotions (e.g. happiness, sadness, anger, fear, surprise, nervousness) and scale of emotions that all humans experience in relation to different experiences and situations
- how to recognise and talk about their emotions, including having a varied vocabulary of words to use when talking about their own and others' feelings
- how to judge whether what they are feeling and how they are behaving is appropriate and proportionate
- the benefits of physical exercise, time outdoors, community participation, voluntary and service-based activity on mental wellbeing and happiness
- simple self-care techniques, including the importance of rest, time spent with friends and family and the benefits of hobbies and interests
- isolation and loneliness can affect children and that it is very important for children to discuss their feelings with an adult and seek support
- that bullying (including cyberbullying) has a negative and often lasting impact on mental wellbeing
- where and how to seek support (including recognising the triggers for seeking support), including whom in school they should speak to if they are worried about their own or

someone else's mental wellbeing or ability to control their emotions (including issues arising online)

- it is common for people to experience mental ill health. For many people who do, the problems can be resolved if the right support is made available, especially if accessed early enough

This defines mental wellbeing for pupils aged 5 to 11. Ironically, the very definition is yet another check list that teachers much teach to and evidence, perhaps impacting on their own wellbeing.

If we just focus on simply the second and final criteria from this guidance, we can link this quite clearly to the vast amounts of previous research regarding mental health and the impact of testing. Given this guidance was published in 2021, it is still very much reflective of previous research and the impact of SATs on mental health from the 1990's, mainly Reay and Wiliam (1999) 'I'll be a nothing' research, highlighting that educational policy has not really progressed. Pupil identity nationally is still significantly impacted on by assessment and testing at primary school level. Indeed, Reid (2019) makes direct links between attendance and anxiety. Chandler (2014) also raises that resilience (all the emotions of point 2) are essential for educational policy.

As this research focusses very much on the lived experiences of staff and pupils, it is important that the term lived experience is also defined. While there are many definitions within research literature, I feel that the definition of O'Leary & Tsui (2022) aligns with the methodology of this research. Given this research is through the eyes of the staff living the new testing regime. O'Leary & Tsui's (2022) definition states that lived experience is a personal knowledge gained from direct experience that would not ordinarily be apparent through observation or via representations constructed by a third party who has not 'lived' it through the eyes of those who were in the situation.

The main starting point of this review was to unpick the policies that lay behind the many pieces of general educational assessment literature, to focussing upon Bronfenbrenner's more outer ecological systems first. The White Paper of 2016: *A Review of the Educational Excellence Everywhere*, is arguably the one that underpins this research following the changes and recommendations is suggested. This was also the starting point in understanding the ecological systems in which this research is framed.

## 2.2 Aims of the Policy

The standards of education in England have resulted in a move down the world league tables. The Government's 2010 response was therefore the biggest overhaul of education since the introduction of the first national curriculum and attainment levels in 1988. The White Paper of 2010 – *The Importance of Teaching* set out the initial proposals. This was followed by the Government's response to the consultation on primary school assessment and accountability in 2014, closely followed by a final report of the commission on assessment without levels in 2015. All of these have provided the basis for the White Paper (2016) – *Educational Excellence Everywhere* in March. The simple aim was to raise the standards and attainment of education.

In the 2016 *Educational Excellence Everywhere* White Paper, the Secretary of State for Education, Nicky Morgan, highlighted a need for the current education system to be reformed. The basis for the reform being, "In 2010, we inherited an education system where 1 in 3 young people left primary school unable to read, write and add up properly" (Department for Education, 2016, p3). This is a bold statement and perhaps at first glance would cause concern regarding the education system. However, it is possible to dissect the statement into 2 sections – the number of pupils, and then the issue of attainment. Firstly, 4B would have been the governments expectation at this time, with the floor target for schools to achieve a 4B set at 65% in reading, writing and maths combined. This target is

allowing for nearly 1/3 of pupils to not achieve in the three core subjects. Achieving Government targets is the very argument put forward by the Secretary of Education for the need for reform. It is also worth noting that 6 years later the Government's floor target was still 65% of pupils to achieve the expected floor standard. This would suggest that the education system is not expecting all pupils to 'achieve' and is aware of the winner/loser system and arguably has been aware for some time. The Government themselves set the target of 1 in 3 young people not achieving; schools simply achieved it.

In terms of the 'unable' Morgan refers to, unable is a significant label to use for all pupils below 4B, or indeed the new national expectation of at the expected standard. Some pupils may just be below but quite capable of reading and writing and adding up 'properly'. The definition of 'properly' I have taken to mean the government's set expected standard but given there is no clear definition within the White Paper, this is my opinion and no doubt the opinion of other researchers will vary. To this extent, the whole basis for the reform could be seen as misleading, if not unreliable, by the educational professionals who were following and achieving government targets and teaching the required National Curriculum. If all targets are achieved, and by this I relate to the new expected standard and floor target, a percentage of pupils will still leave school 'unable'. This is despite a new curriculum, raised expectations, new assessments system and 'greater autonomy to schools.' I would propose that the White Paper strap line of a third of pupils failed by education certainly hits home, but it is not the principle aim of this paper. Raising standards is not disputed by teachers, the means of achieving it as outlined in the White Paper is. Some six years later in 2022, it is unclear how this paper has improved the percentage of pupils unable to 'read, write and add up properly', given the national data for KS2 in 2019 (the last national data set available) has not improved significantly.

A further issue to acknowledge, in the aims of this policy, is the underlying current of discontent within the education system. While the thoughts of Hannon's (1990) government

policies distrust of teachers are subtly hidden, it does aim to address some of the professions current issues. The retention of teachers, due to work life balance and behaviour within schools; coupled with a teaching shortage, performance related pay and increased accountability, were identified within the reforms set out in the 2010 White Paper. However, several years later it is clear they were not the solution it was sold to be. The DfE Workforce Census (2019) identified an increase of those joining the teaching profession increased by only 0.9% and alongside this, the 5-year retention rate dropped to 67.4%. The National Education Union (NEU) survey of April 2022 identified 22% of teachers had signalled they would leave the profession in the next two years, and nearly half intended to leave within the next five years. 52% of those surveyed stipulated their workload was unmanageable. Despite the 1265 directed hours, as per pay and conditions, the reality of this is far greater. 1265 amounts simply to the hours spent in school, not the weekends, holidays and late evenings spent working. It also worth highlighting that 1265 is not applicable for those on any leadership scale. The sacrifices between family and friendships and the paid job are defined under 'workload' and I will address this also within further chapters.

While previous boycotts had been avoided when initially implementing SATs, a further White Paper, six years later, needed the support of the education profession and their unions. It was therefore acknowledged, within the foreword, the hard work of teachers and that schools today are better than ever before; though, still not good enough. Perhaps it was a genuine attempt to acknowledge the hard work of teachers and raise the regard for the profession publicly, or perhaps it was a way of appeasing the views of the very people having to enforce the White Paper despite potentially not agreeing with any of its content. Those now telling pupils and parents that their child has passed or failed, at national or below national, must not be insulted within the introduction.

It was also the aim of the *Educational Excellence Everywhere* (2016) paper to indicate that no further changes would be implemented and confirm that the teaching profession had been well and truly listened to. However, Torrance and Pryor (1998, p.43) identify that many teachers regard assessment as a distinct activity separate from teaching. It is seen as something conducted to gather information for third parties for purposes of accountability rather than for the benefit of pupils or teachers. This evidence is supported by a study of primary school students in Geneva. Here it emerged that pupils believed that the assessments of them were solely for the school's and parents' benefits, not themselves. The weaker pupils believed the purpose of the test was to make them work harder and as it offered no support, they saw it as a source of pressure (Black, 1998). Neither teachers or pupils are clear on the benefit to them of testing and this White Paper had not really addressed that.

In identifying no further changes would be made, this paper simply highlighted that the government had not understood the teaching profession's view not only of assessment, but also their view on education. The White Paper (2022) *Levelling Up* has continued the pledge, and other than raising the expected attainment levels, highlights no further changes within the primary sector. Almond (1994, p.73) raises that while there is an agreement that education is in the interests of children, there is a disagreement about what good education is, and what it should achieve. The recent parent campaign of *Let Kids be Kids* (2016) highlights frustrations of the teaching profession and the society which the White Paper proposals are meant to benefit. "Investing in our education system is an investment in the future of our nation" (Department for Education, 2016, p.3).

The White Paper (2016) identifies seven elements to achieving educational excellence everywhere:

- Great teachers everywhere they are needed.
- Great leaders running our schools and at the heart of our system.



- A school-led system with every school an academy, empowering pupils, parents and communities and a clearly defined role for the local government.
- Preventing underperformance and helping schools go from good to great: school led improvement, with scaffolding and support where it is needed.
- High expectations and a world-leading curriculum for all.
- Fair, stretching accountability, ambitious for every child.
- The right resources in the right hands: investing every penny where it can do the best.

Given that my research focuses on my own decisions as headteacher, based on teacher perspectives of the recent assessment system, with a particular focus on testing and the impact of pupils once results are given The areas of focus of the 2016 White Paper within this review will be the curriculum and assessment aspects, which are underpinned by the above aims.

### [2.3 Rationale for Change in Assessment](#)

The removal of curriculum levels in September 2015 was promoted to allow schools to develop an assessment system which would align with their own curriculum and would work for both staff and pupils. The basic rationale for removing levels (identified by John McIntosh CBE in his final report on *Assessment Without Levels*, September 2015) was that too often levels became viewed as thresholds and teaching became focussed on getting pupils through the next threshold. In short, pupils were taught to pass tests. It also highlighted that too much time was being taken up with unnecessary assessment and time could be better spent in the classroom:

This government will rarely dictate how these outcomes should be achieved. It will encourage and support teachers and leaders to develop the best possible solutions for their pupils. (Department for Education, March 2016, p.9)

On face value it would appear that the government have listened to teachers with Nicky Morgan's foreword in the *Educational Excellence Everywhere* (2016), the *Assessment After Levels* (2015) paper and the vast DfE publications in 2016. All of these documents appear to attempt to give teachers back some degree of professionalism, freedom, and more time in

an attempt to address their work life balance. However, testing remained, and a new floor standard of 85% was set in March 2016. Whilst this would address the 1/3 of pupils failing previously discussed. It was soon evident that this was unachievable, despite more tests than ever before and an increase in expectation for each pupil. This is evidenced with the first set of national data published by the government in 2016, which identified only 56% achieved the floor standard under this newly introduced system and therefore the government had no alternative but to leave the floor target at 65%, which remains to date.

Instead of conceding the proposals may need to be reviewed, or even that 85% was set too high, the White Paper simply acknowledges that the curriculum (especially at primary) has increased stretch and was an ambitious programme. The White Paper (2016) states that the government will continue 'to hold schools to account for unapologetically high standard for all pupils, measured rigorously and fairly' (p.9). The 2022 Levelling Up Paper has however, included a target of 90% by 2030. While this paper has now been removed, it is unclear if this target remains. Regardless, this is a bold target to set and will undoubtedly produce further pressures somewhere along the education chain of standards. This is also another vital reason why the decisions of headteachers, and impact of these decisions, should be researched. Not the policies that the governments of the day make from their own ecological system, not even the tests themselves, but the interpretation and actions of those within the micro ecological system. The headteachers making those daily decisions are the key to policy and student success. Both of which may be vastly different things.

These new system of assessment without levels, in many schools, simply created a different threshold through which pupils must 'jump'. Instead of pupils having to achieve 4B as was expected prior to the new assessment system, they now had to achieve a national standard. This was the 'expected standard' and criteria for achieving this would be released each year, published as *The National Curriculum Framework* document. There would be two documents: Working below the assessed standard and working at the assessed standard.

This was a challenge once it was identified that the new national standard was higher than the previous 4B in relation to what was expected and that 85% (though this was not achieved - it was the aim) had to achieve it. This potentially then led school leaders and teachers to choose between teaching to the test, alongside the risk of creating new assessments systems that did not focus on the criteria of 'below' and 'working at.' Despite this, the DfE were clear that school leaders had the autonomy to devise how they teach the national curriculum and how they assess pupils' attainment:

We start from the basis that the country's best leaders know what works, and that good, enthusiastic leaders should be able to use their creativity, innovation, professional expertise and up-to date evidence to drive up standards. Department of Education, (DfE, March 2016, p.9)

The headteacher's decision here, on what their new assessment system would be was pivotal for future school development. As soon as one attempts to break from the traditional view of educational achievement, one is confronted with the need to also make a similar break from the views on assessment. In the same way that intelligence tests restricted and dominated the views of intelligence, testing and prominent assessment methods tend to distort the concept of educational achievement (Torrance, 1988, p.10). Given the high stakes associated with standards, it is easy to see how leaders and in turn teachers focus clearly on the finish line of SATs rather than the journey there. Mainly because accountability is mentioned more than any other subject within the paper:

An effective accountability system ensures that professionals are held accountable for the outcomes of their decisions using fair, intelligent, reliable and carefully-balanced measures of success and failure. These will emphasise progress for all pupils and measure more ambitious outcomes. (Department of Education, March 2016, p.21)

We believe that outcomes matter more than methods, and that there is rarely one, standardised solution that will work in every classroom for government to impose. The elected government should set out the outcomes – what needs to be achieved for the public money invested in education. But we start from the basis that the

country's best school leaders know what works, and that good, enthusiastic leaders should be able to use their creativity, innovation, professional expertise and up-to-date evidence to drive up standards. This government will very rarely dictate how these outcomes should be achieved – it will encourage and support teachers and leaders to develop the best possible solutions for their pupils, and will hold them to account for rigorous, fairly measured outcomes. (DfE, 2019, p.9)

Everything from what is taught, how it is taught, and when it is to be taught was changed because of the White Paper. The one thing that remained was the high stakes testing; the 'outcomes' referred to within the paper. Until this is addressed, everything else is almost irrelevant. Primary school teachers are still measured, paid, and judged on the same results as before the implementation of the paper. The role of the headteacher is addressing this forms the basis of this research.

Despite attempting to uncover any national data to show the systems for assessment in place in schools, I found little published information, other than the ones schools themselves had created. This may be because testing remains and accountability is still attached to them, which has outweighed any attempt for schools to move away from teaching pupils to pass them. Given that from 2017, Ofsted have moved their focus to the curriculum and the necessity for a broad and balanced offer, this suggests their own concerns with a teaching to the test education system.

Perhaps schools could be excused for an initial teaching to the test approach, given this was a new system and the guidance on exactly what the national standard would look like was not released until February 2016, progress guidance was not released until after SATs, a whole year after the removal of levels. However, at the time of my research analysis, teachers and leaders had entered the fourth year of a new testing system and could reflect on whether assessment systems had really changed. There was no reference within the White Paper that testing would cease, leading some to question whether any of the aspects raised within it would really cause change.

It will be interesting to see how the report from the Assessment Review group on Redressing the Balance (January 2017) impacts on future White Paper proposals. They identify they are not opposed to testing but advocate a different approach to the administration and value placed upon them. This has not yet been raised in any further DfE proposals since its publication.

There is a clear assumption within this 2016 White Paper that the reforms on which it is based were working and believed in. Therefore, a brief review to understand how and why the concepts within this White Paper came into being would support fully understanding the White Paper's publication.

On the 17<sup>th</sup> July 2013 proposals were published to reform primary school assessment and accountability under the new curriculum. The consultation closed on 11<sup>th</sup> October and received 1187 responses, with a series of discussion then held at events and conferences.

Given the sheer volume of changes and successive White Papers built on these results and the high stakes for educators, the published data shows few respondents in comparison to the number of teachers nationally. That is not to suggest the figures are not true, more that they reflect the timings of when the consultation was released, which was the start of school summer holidays, when teachers are usually fraught and ready for the break. It could be suggested that this is not the best time to engage with the teaching profession, especially as teachers leaving the profession are at their highest towards the end of the school year (School Workforce Census, 2019). Closing in October, again at the start of a new term when routines are being established, class relationships created and so on may not also be the best time to consult with teachers. That said, regardless of the timing of the consultation; teachers would have responded if the stakes were clearly identified, and the events and conferences clearly published. Whilst this was a personal opinion, I actively searched for

these events in 2013 and believe them to be difficult to find and not easily accessible, in venues that would require some considerable travel due to so few planned.

The consultation posed a series of questions:

Q: Will these principles underpin an effective curriculum and assessment system?

No – 606 responses (57%)

Yes – 196 responses (18%)

The system was implemented.

Q: Does a scaled score and value-added measure provide useful information from national Curriculum test?

No – 726 responses (70%)

Yes – 163 responses (16%)

Scaled scores were given to support the result of either at national or below national.

Q: Should we introduce a baseline check at the start of reception?

No – 572 responses (51%)

Yes – 382 responses (34%)

A baseline was implemented (at a significant cost). Schools were given the option on which baseline, from the 5 options, to use. These took considerable time and are now no longer compulsory, as the results from different baselines are not comparable. A whole new system five years later was being designed solely to look at the rushed in baseline reception. It is only the last few years this new baseline system nationally has been implemented.

There was no 'choice' this time and the results of which are not shared with schools.

The concerns raised during consultation are now implemented and other than the baseline, continue. In their research evaluating educational assessment and accountability, Skedsmo and Huber (2019) identify that while bottom-up approaches demonstrate positive results in relation to school development and student outcomes, more top-down oriented models seem to have a less positive influence. Instead, they seem to produce a range of unintended

consequences, and perhaps questions could be raised as to whether they are in danger of contributing to the de-professionalisation of teachers. Given the new reforms have yet to impact despite many years of testing since its introduction and the current discontent within the teaching profession regarding assessment. It could be suggested that the very basis of the White Paper is a top-down approach to controlling education and is far from the giving autonomy and powers back to schools and school leaders as promised. A simple glance at the data, show the little teacher views that were given, were blatantly ignored.

#### 2.4 Impact of the White Paper

The Educational Excellence Everywhere (2016) paper reformed assessment and qualifications to be as 'challenging' as those in the highest performing countries. What it fails to acknowledge, are the differences between those countries' education systems and our own. It refers to following the example of other countries, regardless of context.

While this research is not looking at attitudes towards education, a link between these attitudes and Bronfenbrenner's (1979) ecological systems should be highlighted. With educational prestige linked to the wider culture of a country and the ecological systems in which they are embedded – wealth, poverty, ethnicity etc. It is unclear if the research regarding our own country would be able to emulate another by simply copying one aspect of a whole ecological system. One of the main criticisms of Durkheim was his assumption that societies had a shared culture, transmissible through education. It is debatable if multicultural countries, such as Britain, would have a single culture on which to base a curriculum (Haralambos and Holborn, 2013, p.664). The same criticism could be applied to this aspect of the White Paper.

Respect and prestige are not usually synonymous with the education system. The notion proposed by the *Educational Excellence Everywhere* (2016) paper in only recruiting the

highest graduates, will not necessarily address this. It is unclear how achieving a higher-class degree will improve the ability to work with children, especially those children that struggle with education and to them, a first-class degree seems far from possible. The National Audit Office suggests “degree class is a reasonable indicator of subject knowledge but a less clear predictor of other aspects of teacher quality” (Milne, 2017).

Furthermore, in 2016/17 only 18% of all trainee teachers had a first-class degree. While this is a slight increase on 16% in 2013/14 (DfE, press release November 2013), the official data still shows that less than a quarter of all teachers have a first-class degree. Given the national floor target data at 65%, those teachers without a first-class degree must also still be ‘performing’.

Hanushek, Piopiunik and Wiederhold’s (2019) research comparing top teaching graduates to their student scores, highlights correlations between teacher quality and student achievement. They also raise those external factors, such as pay, which also impact. Furthermore, they also make a valuable note in that many pieces of research have been concluded on higher graduate teachers and their results, but fewer research pieces on lower graduate teachers and their student attainment are available. To a small extent, this research will also touch upon this. If we relay these findings to the existing national standards data and the available graduate data, it can be seen that results are being achieved by teachers with less than ‘top grades.’ There are simply not enough teachers with top grades currently within the system to be achieving the yearly results and as the percentage of ‘top’ graduates entering the teaching profession rises, the same cannot be said for the national data.

My own personal journey is far from a top graduate. I had three jobs to be able to afford university, was a teenage mum and struggled in many areas of education due to who I was, not what I could do. If this system had been in place twenty years ago, I would not even



have been given the chance to teach. I would suggest that those who genuinely understand and have 'lived it' are those best placed to advise and raise the aspirations of pupils and in turn, the prestige of teaching. In this statement, I am by no means belittling those teachers who achieve top grades and I acknowledge that these grades will have taken significant effort, what I am highlighting is that a top grade alone is not going to change the education system in which it is set. Those that have lived the education system in which they work, are best placed to change and understand it. Few education secretaries have been educated within a state system, excluding grammar schools, and even fewer have been teachers themselves. This highlights yet another conflicting decision for headteachers to make. Employ those on paper that are top graduates, in the hope this is an indicator that they can raise standards; or employ those with lower degrees, but have a genuine understanding of the issues facing their pupils. Of course, there will be teachers who have high class degree and have a lived experience but these are not the majority. Given the workload, behavioural issues and complexities of education, judging potential educational employees based on a grade alone and not their character, may not be the best indicator. It is the very thing we are doing to our pupils from the age of 5 and given the world league tables, it does not seem to be greatly successful.

Review after review, White Paper after White Paper highlight shortfalls and failures of the profession and therefore implementation of a new fad or in this instance a completely new system is introduced. Yet this overhaul takes little account of the pupils at the heart of this system. It is easy to see why young people become disillusioned, even more so now with a pass or fail attitude. Teachers have very little power to take an interest and tailor support for each pupil, not because they do not want to but because they must get those pupils to the magic floor target by hook or by crook. Their career is based on it and therefore pupils could begin to feel part of a system rather than an individual – they are there to get results and make up the percentages. Gipps (1994) identifies the culture of British primary teachers

who maintain that assessment of pupils should mainly be concerned with diagnostic purposes and are clear on the dangers of the impact on testing. She also highlights that teachers are fully aware of the external factors impacting on their students such as family, social backgrounds, nursery provision, summer born and so on. The reconstructionism stance of teachers creates a conflict between their own ideologies and those they are employed by. Just as I had to make the choice regarding who I needed to be accountable to as a headteacher (my pupils or the local authority), so too do teachers. The decisions made by a headteacher impact on a whole school community and thus ecological system.

Following on from previous data already discussed, national statistics show that more and more teachers are leaving the profession, with 50% in the first 5 years. The Government's own data highlights that it also missed its own training target to recruit new teachers. A quarter of all staff surveyed (11,177 teachers) in the DfE (2022) Working Lives of Teachers and Leaders Report raised that they were considering leaving within the next 12 months. 76% of these were considering leaving due to government policy changes and initiatives, resulting in high workload. It is becoming more and more evident that education within England is at risk of not being regarded in the same esteem as it once was. Yet this is also not an area addressed in the White Paper – It assumes that pupils will be willing passengers on this White Paper journey. Their thoughts, feeling and anxieties are not discussed despite the vast research showing clear links between educational issues and individual thoughts and feelings. Sugrue (2008) raises how the social environment impacts on education, yet these policies are made with little regard for the ecological system in which schools set. He goes on to say that because education is so important and because our children are so important- educational issues will continue to engage people's feelings and commitments strongly. It is hoped, at the very least, that this research stimulates feelings and that my research can give a voice to these 'passengers', albeit passengers within my own setting.

However, in raising this issue, the White Paper does acknowledge to some extent that pupils can and will impact on implementation. The area the White Paper focusses on is that behaviour is a concern. This is only in so far as it is acknowledged as a reason for teachers leaving the profession. What the paper does not address is how this behaviour can be tackled considering all the changes it proposes. Perhaps a further reason why teachers have continued to teach to the test; is that the pupils are used to it. The paper also fails to identify that behaviour might be caused by the system of teaching to tests and the high stakes placed on results. Styron (2012) identifies that while vast amount of research shows both positives and negatives linked to teaching to the test, what is not there are deterrents for not teaching to the test. If behaviour while having a teaching to the test approach is acceptable, then there may seem to be no reason to change it. A further finding of Styron's (2012) research was that 'superficial' teaching of skills which are disconnected will continue until tests measure the actual depth of knowledge. Claxton (2002) and Claxton, Chambers, Powell and Lucas (2011) also raise this by advocating for real life learning. In relation to the White Paper (2016), there is no change to the testing and a depth of knowledge is not needed. In not challenging thinking or reasoning and develop a deeper understanding of skills, pupils will not fail and have relatively 'easy' success in rote learning. They will be excellent student in passing tests, but perhaps not so excellent with real world or even life problems. Behaviour will not be challenged as the White Paper, despite permitting it, has had little impact on daily classroom practices.

It is also fair to say that the White Paper also briefly discusses resilience:

A 21<sup>st</sup> Century education should prepare children for adult life by instilling the character traits and fundamental British values that will help them succeed: being resilient and knowing how to persevere, how to bounce back if faced with failure, and how to collaborate with others at work and in their private lives. (Department for Education, 2016, p.94)

With harder tests and a pass or fail outcome, it is apt that the White Paper should highlight resilience and managing failure and perhaps reflects Saleebey's (1996) view on resilience. This will be discussed in some depth within this research.

Whilst there are many issues the White Paper does not address, as the extract above already given highlights. It would be unfair to say it does not address some of the impact issues faced by teachers.

There is no doubt that the raised expectation in SATs tests is a direct result of the paper which continues to impact on the day-to-day teaching within the education system.

Teachers do not have concerns regarding being held to account or even that there are testing regimes. Their issue is the high stakes associated with these tests, the evidence gathered in a 40-minute test at the end of a year in comparison to a year of rigorous on-going assessments, observations, and discussions. Gipps and Stobart (1993) take the position that properly used assessment is not only valuable but essential. It is the misuse and overuse that impacts on both teaching and learning.

The accountability measure that is frequently mentioned throughout the paper is solely measuring what teachers must do, not what matters to them. It does not measure if a pupil finally made a sports team, mastered English when it is not their first language or how good a pupil is at art. It measures what is tested, creating a classroom culture of teaching only those things being tested, with other subjects being squeezed out. James (2008, p.21) identifies two sets of assumptions underpinning assessment, with both being preoccupied with what is taught and how knowledge is transmitted. She identified learning as the acquisition of knowledge and therefore learners are viewed as passive, with success judged by the extent of the knowledge acquired.

Other than a strong moral compass, why would teachers focus on these things if they go unnoticed? Lucas, Claxton and Spencer (2013, p.8) raise the notion of education being a

moral business and create pupils 'who know, think and behave one way, rather than the other'. Have teachers got time or perhaps more importantly support to say to parents, I understand your child may have struggled with core subjects, but they have tried so hard, are a credit to the school, have a fantastic work ethic and so on? Do the teachers have the support of a headteacher to openly value the things not tested? These types of statements may not keep their job safe and there is no guarantee that parents would understand or even value this. Creating a new curriculum will not stop this approach and a hierarchy of subjects, unless the accountability is made to be more than a test score in reading, writing and maths:

Our new accountability measures provide a fair reflection of progress and measure outcomes while giving teachers professional autonomy to decide how those outcomes should be achieved. (Department for Education, 2016, p.106)

I would therefore suggest that the notion of autonomy threaded throughout this paper is not worth a great deal, until accountability is addressed. Gipps and Stobart (1993 p.104) identified that external assessment systems do not offer the high-quality information they claim and due to the high stakes associated with them lead to a narrowing of the curriculum. What they identify as missing from education is a system that promotes thinking, creativity, reasoning, and the ability to cope with the challenges of the next century. Despite this being identified over twenty years before this White Paper, none of these areas have genuinely been addressed. Biesta (2013, p.61) discusses that students are currently living in a fast-changing and complex social, economic, and political world in which they need to adapt by acquiring new knowledge, skills and attributes within a range of contexts.

The World Economic Forum identified the top 10 skills for 2025, these being:

- Analytical thinking and innovation.
- Active learning and learning strategies.
- Complex problem solving.
- Critical thinking and analysis.

- Creativity, originality, and initiative
- Leadership and social influence
- Technology use, monitoring and control.
- Technology design and programming
- Resilience, stress tolerance and flexibility
- Reasoning, problem solving and ideation

This White Paper neither addresses or promotes many of the skills learners may need and instead repackages previous systems.

I have already addressed why school are not changing their decisions on how outcomes should be achieved. The notion of a fair reflection of progress has briefly been addressed, in that it is a measure of progress in the things that are tested. It is perhaps a touch ironic that so much has changed and so quickly following the White Paper, except the one thing that needed changing – testing. Some years after the White Paper, at the *Education Conference* January 2012, Nick Gibb (schools minister) boasted that the Government had cut over 6,000 pages of guidance. By July 2012 the Department for Education website had 4,238 publications related to education and cognate matters (Ball. 2013 p.4/5).

Despite the length of the White Paper, it gives no guidance to school leaders on what a new assessment system should look like and the impact of this would be fair to say was chaos. It is my decisions taken during this period that need reflecting on as part of this research.

The House of Commons Education Committee for Primary Assessment (April 2017) identified that The Standards and Testing Agency did not oversee the implementation of the new assessment system in 2016 effectively. It identified that guidance was delayed and test papers leaked online. The impact of this caused significant disruption in primary schools, with schools feeling there was too little time to implement effective new assessment systems and prepare teachers and pupils for SATs. The review also highlighted:

While the new assessments were being introduced there was little additional support offered to schools to implement new assessment systems to cope with 'life after levels'. Primary school teachers only receive limited assessment training during initial

teacher education and must have access to continuing professional development on assessment, as well as high quality advice and guidance on effective assessment systems. (House of Commons Education Committee, April 26<sup>th</sup>, 2017, p.3)

If a national benchmark is needed in the form of testing to create league tables, it is difficult to understand how a system created by individual schools can be any good outside that of school. How are pupils who change schools part way through a school year baselined – a test, as there is no commonality between schools. Pupils have no real grasp of where they are in terms of what they can and cannot do. The previous system of a numerical level and sub grade gave them something to compare, regardless of the school. They knew where they were in the class and what the next steps would be. To now be told they are at national or below national is not of much use to them. Questions such as: How far below? What do I need to improve on? Would not be as clear as achieving a 4C and knowing that was just below 4B as in the previous system. It also had the potential to impact on teacher pupil relationships, with teachers now having to inform a child that has moved so much and achieved so much, they are below national. Highlighting in the White Paper a need to 'bounce back' and develop 'resilience' does not take away the heart-breaking impact of teacher pupil relations this paper has created. This is reflected by secondary schools who now have cohorts of year 7 pupils to group, based on 2 groups: national and below national. The government did eventually create an answer to this – scaled scores. As previously discussed, during the consultation in 2013, teachers had responded to state that scaled scores were not useful. Despite this, the agreed scaled score to achieve national is 100. Therefore, a test score was given to pupils, but a scaled score then reflected multiple test score answers collectively and pupils just found it difficult to comprehend. Despite the white paper expecting;

An effective accountability system ensures all children receive the education they deserve and plays an important role in informing parents about school performance so they can make informed choices and scrutinise the performance of their child's school. (Educational Excellence Everywhere, 2016, p.105)

The impact of this statement is that parents will see that fewer schools are achieving the floor standard of 65% and that whilst they have a national or below national label for their child, it does not show them everything their child has achieved. This is clear in the sheer volume of SATs letters sent by teachers and headteachers over the internet each year, explaining that the test score does not show the true 'them'.

Given the areas highlighted within this review in relation to all the issues the White Paper ignores, it does have some helpful guidance or information. There are some very clear statements on what to expect. It is clearly stating that the tests would be harder, with more challenging questions. It highlights the importance of a knowledge-based curriculum and discusses at some length that: 'no matter how skilled they are as readers; children cannot fully understand what they are reading unless they know the meaning of words and references in texts'. (p.89) It also highlights a need for synthetic phonics. If leaders and teachers took these as warnings or advice, they would have focussed on phonics, understanding vocabulary, and addressing raising attainment through working at old school level 5.

This would have been a wise move. The tests produced and taken by children during the first year of the new proposals, were significantly harder than previous years. Some Y6 maths questions were extremely like the old KS3 maths paper. The phonics expectation was raised to create the national result at its highest to date. The reading paper for the first time, was written specifically for the test to have the required vocabulary pupils should know. Most of the questions were created to check understanding of the use of language not if the pupils could read the text. What was tested was if they could understand the language used and then write in a way which proved their understanding. This was made clear in section



6.4 of the White Paper (2016) and highlighted that pupil must know the meaning of words and references in texts.

The White Paper (2016) may not stop the teaching to the test or promote teachers creating their own way to achieve the outcomes. Gibbs and Stobart (2013) raise that tests are linked to the content taught and aims only to assess what students know, rather than what they do not know. If the National Curriculum is what needs to be taught and is what is assessed, teaching will change very little. Murphy and Torrance (1988) raise the notion that assessments should play a crucial role in education, and it is reasonable for teachers and other stakeholders to be curious about what has happened in terms of progress and outcomes. They identify that a good education by its very definition would encompass assessment but disassociated themselves from previous assessment systems run under the guise of good assessment.

Furthermore Broadfoot (1995) raised that the process of SATs themselves are far from normal classroom styles where teachers would engage with pupils to support misconceptions and errors, listen to pupils' reasoning before making choices and addressing behaviours. The White Paper (2016) still has an everyday classroom practice element running parallel to a testing regime. However, it still only tests what it tests and no more.

My own opinion is that while teachers may have different ways of delivering a curriculum, word choices and understanding vocabulary will be a significant feature in most year 6 classrooms. Phonics and, in particular – nonsense words will be the main staple of reading sessions in year 1, as they are the outcome tested. The White Paper itself states it is 'outcomes not method' which are the focus. The hoops to jump through to pass a test have not been removed, schools just get to label them themselves. The role of the headteacher in doing this is explored in depth in subsequent chapters.

## 2.5 White Paper Conclusion

Nearly thirty years ago Gipps and Stobart (1993, p.38) highlighted that assessment data would conclude in pupils and schools being evaluated based on results. If results are less than desirable, then schools may face closure staff may face redundancy or lack of promotion. Today teachers still have performance related pay, but it is just now linked to national expectations instead of levels.

The level of accountability remains the same and the high stakes attached to results remains unchanged. It appears to be a token gesture – teachers are free to teach how (not what) they like, and to assess how they like but the end product has not changed.

In conclusion, the impact of the White Paper is significant and will continue to impose ideologies at conflict with the school leaders and teachers at the heart of the policy. Whilst this is not a new concept, it is one that appears to be here to stay, despite the evidence highlighting the blatant issues. The recent report on primary assessment (May 2017) lead by Neil Carmichael called for an overhaul of the league table system to reflect a three-year average, stating that whilst schools need to be held accountable for attainment and progress, the stakes need to be lowered; “Many of the negative effects of an assessment in primary schools are caused by the use of the results in an accountability system rather than the system itself.” (Para,66) In addition, it states:

The high stakes system can negatively impact on teaching and learning, leading to a narrowing of the curriculum and teaching to the test, as well as affecting teacher and pupil well-being. (Para.58)

The issues with this new system were still being discussed in the Nery’s Roberts briefing paper of December 2019 for the House of Commons. It again highlighted issues with the original programmes of study drafted in June 2012. She raises original comments from Prof Andrew Pollard who, as a member of the expert panel, had described the proposals as ‘fatally flawed and overly prescriptive’. Professor Robin Alexander, director of the

Cambridge review, criticised the government as citing 'selectively and tendentiously' from the evidence and ignoring contextual and cultural factors in their case for reform. The briefing paper also created the press release from the NUT highlighting their concern that there was a risk in creating 'a task orientated curriculum which would stultify the learning process, due to its inherent inflexibility, making it impossible for many children to achieve in the time and space they need.'

All of these concerns were raised significantly before the introduction of the White Paper (2016) but would hopefully be addressed in the report of the Assessment Review group in 2017.

## [2.6 Report of the Assessment Review Group: Redressing the Balance January 2017](#)

Not long after the White Paper (10 months, and 2 years after the removal of levels) the

Review Group were already addressing future principles for assessment and in the foreword acknowledge that 'any attempt to design one is likely to prove controversial.' (p.2).

David Ellison's opening paragraph of the review states:

Assessment is at the heart of high-quality teaching and learning. It is the means to which teachers evaluate progress and diagnose pupils needs. It provides the opportunity for students to recognise their progress and help parents engage in their children's educational journey. To put it simply, assessment helps teachers to teach and pupils to learn (p.2).

For such a dramatic opening statement, there is one simple word missing – testing. As already alluded, while the word assessment is often used, what is meant is testing. Instead, the review identifies the usefulness of different types of assessment.

The assessment review identified six principles of assessment, two more than given by the HMO in 1988:

- 1) Assessment should be at the core of good teaching and learning.

It is here that the word 'ongoing' is first discussed and identified as at the 'heart of effective teaching and learning' (p.6). As a school leader and a teacher, the quick assessments (not tests) that formed the ongoing judgements and tweaks had the biggest impact in the classroom. These professional diagnostics are not necessarily measurable and uniformed , but certainly more helpful to the pupil and teacher than a 40-minute test. They are instant, in the moment analysis and answer situations. Errors are addressed immediately and via a non-threatening discussion, rather than sitting in silence and not observing the errors at the point of misconception or discussing the pupils thinking. Gipps (1994) highlights that it is test scores that rise when teaching to a test and not the skill itself. This first principle is concluded within the review with the suggestion that there should be a continued focus on improving the effectiveness of day-to-day assessments.

- 2) Statutory assessment should be separated from ongoing assessment that happens in the classroom.

This principle highlights the review's conclusion that the high stakes accountability can have negative consequences. Many of these have been and will continue to be discussed within this research; teaching to the test, narrowing of the curriculum, more emphasis on rote learning rather than acquisition of reasoning or problem solving. It concludes this second principle by stating that a balance between statutory assessment and every day ongoing assessments should be sought.

- 3) Data from statutory assessment will never tell you the whole story of school effectiveness.

This principle aims to address the presumptions that poor attainment is a reflection on the school's failure. It highlights the need to not draw conclusions about performance from raw data and in turn intervene. This principle clearly highlights that 'the misuse of data is at the heart of today's problems with assessment' (p.7). Willam (2008) highlights that a test tests

what a test tests, no more no less. There is no such thing as a biased test. Like validity, bias is not the property of the test, but rather the conclusions we draw from it.

- 4) Statutory assessment systems should be accessible to pupils of all abilities and recognise their progress.

The current labels of working below or expected standard are raised within this principle as unhelpful to schools and sends 'the wrong message to pupils, potentially having an impact on their future motivation.'<sup>p8</sup>. It also raises the language used in the reading tests as not being accessible. While the tests were specifically written for the sole purpose of the test, there was a significant expectation on the pupil's language skills and life experiences.

Evans' (2007) fieldwork highlights the social variation model which concludes working class pupils are less prepared for school and therefore are less able to compete on equal terms.

The pupils as alluded to earlier, with limited experiences and knowledge, could read the text quite fluently but had trouble with the use of language. Scold as though to tell off, is one example. This language is not one accessed by many within deprived areas where scold might be interpreted as 'scalding hot' and burning. 'Dawn breaking over the Savannah' are two girl's names, unless you know time and places. Pollitt (2014) identifies that each specialist discipline takes words from its own specialist 'natural' language. An example being 'power' in physics and the same word in mathematics. He suggests that the specialist in each area should make it clear words are not natural but linked within context and specialism and that each has a duty to acknowledge respective definitions. I would agree that the tests themselves were not helpful in breaking out of the working below mould and were far from accessible. Pupils could read them, but the test was not checking if they could read, it was their language and the understanding of it being tested, let alone their ability to write their understanding of what they had read. If the ability to simply read was being tested, pupils could physically just read an unknown passage from a text, without the need to sit any formal written test. The reading test measures language understanding, not the

ability to decode and read new words, as explained and highlighted in the Educational Excellence Everywhere 2016 paper. Gipps and Stobart (1995) raise many examples of testing and uncertainty of what is being tested and query if a single test score on a reading test genuinely represents a pupil's attainment in all skills in reading. Testing if pupils can simply read words is the purpose of the phonic test in year 1, with pupils expected to pass these 5 years earlier. Hart and Risley's (1995) 30-million-word gap highlight the deficits of working class families with their language, showing further links between reality and the national expectations of the reading test.

It is also worth noting that the reading test requires an ability to write an answer. It has not been unknown for a pupil to be able to read and indeed understand well what they read but lack the written skills to formulate or articulate their answers. While verbally they would be able to have a very good in-depth discussion on the text read, they are not able to write at the same in-depth level.

5) Progress should be valued over attainment in statutory assessment.

It is this principle that schools in deprived or tougher areas value and welcome. The DfE Primary School Accountability Guide (2018, p.8) raises that progress from Key Stage 1 (year 2) to the end of Key Stage 2 (year 6) 'rewards' schools for making progress with their pupils, regardless of ability. Pupils with extremely low starting points can make exceptional progress, but still fail to achieve the new expected standard or level 4 previously. In a similar way, pupils in more affluent areas can achieve very well but have made no real progress. This principle highlights it to be 'grossly unfair to base comparisons on attainment when children's starting points can be so different' (p.9).

This principle does attempt to address some of the concerns raised with the implementation of the new testing system and the raised expectation, that some pupils will simply never get there and their efforts would go unrewarded. It also, in some respects, addresses the issues

raised within its own third principle; data does not show everything and certainly not starting points.

6) The number of statutory assessments in the primary phase should be minimised. Here the review group highlight a lack of research that supports testing frequency and pupil outcomes and recommends a reduction in the amount of statutory testing. It highlights societal, economic, and cultural factors as having more impact on attainment than the amount of testing.

## 2.7 Conclusion of the Impact of the Assessment Review group

At first glance all six principles are exactly what many school leaders and teachers had asked for, yet the reality is that this review gave very little. Within a short space of time, areas from the White Paper were already dismissed. Previous chapters have already highlighted the floor target itself and forced academisation being changed within the first year of the paper's publication. From the six guiding principles it identifies, it could be suggested only two (principles five and six) have been addressed and even then, it is more lip service than an overhaul.

### 1. Principle One (assessment should be at the core of good teaching)

This principle acknowledges that day to day assessment is what good teaching is all about, but it has done little to change the notion that statutory tests are still more important, as any results are still linked to pay and published. This principle has to some extent 'justified' to teachers their own belief, that they are the ones that make the difference and have the most impact, not testing. It has not changed any aspect of the impact statutory testing has on the teaching and learning; it simply reassures teachers that they matter to their pupils. The House of Commons Review (2017) recognised the importance of holding schools to account but that the high-stakes system does not improve teaching and learning at primary school.

In conclusion, testing to the test will continue, as that is what is measured. This then leads to the second principle.

1. Principle Two (Statutory assessment should be separated from ongoing assessment that happens in the classroom)

This principle reinforces to educators the vital role in their ongoing assessments, yet it has not changed the perception of those beyond the school walls. The value given to a teacher's assessment has very little value to parents or pupils if a label of below expected is still given. While teacher assessments were still published alongside statutory assessment, it was usually used as a weapon to show that the teacher did not really know the class and therefore that must be why the results were low. Surely the teacher should have known they were going to fail and done something about it. No wonder he/she did not pass. It was rarely linked to the pressure of the test and that the test was indeed a snapshot of that day and time. Any difference in 'levels' given was simply teacher error or poor teacher knowledge/understanding of their pupils. The House of Commons Review (April 2017), only one year later; highlighted that in giving a pass mark of a 100, teachers aimed for this and adapted their teaching towards it. While they identify that this was not the original aim of the curriculum review, which was to encourage more 'mastery' of concepts at primary school. The review concluded that the expected threshold should be removed as it "encourages excessive focus on students at the margin of meeting the standard" (p.22). This enforces that more weight is given to the statutory tests' pupils take and not the day-to-day teacher assessment.

- Principle Three (Data from statutory assessment will never tell you the whole story of school effectiveness)

This principle does try to address the issues already highlighted from the previous 2 principles, but again does very little to change perception. Schools are still blamed for



results and while ever this is still the case, the tests will continue to be at the heart of what a school does to survive and avoid the blame for a broken system. However, it would be unfair to not raise the review groups recommendation that coasting, and floor targets carry such heavy interventions, to improve the whole picture being visible. Yes, schools will no longer be forced into academisation if they are coasting or fail to achieve floor targets, but they are still judged and placed in an Ofsted window. Parents see the league tables and children have a label of pass or fail, all of which are attributed to a school. Pay of teachers and headteachers jobs being at risk are not addressed, they are simply airbrushed away and not raised in the hope of implying that results no longer carry such high stakes. Professor Harvey Goldstein identified in the House of Common Review (2017) that publishing data in the current format has drawbacks and fails to show true comparisons. He develops this to raise the concern that currently the accountability component dominates everything else, and it distorts the curriculum, it distorts learning, it distorts children's behaviour (p.18). Therefore, the assessment system itself, is impacting across schools in a variety of forms. However, even with the acknowledgement that the core subject data may not be fully comparable, there are more aspects to primary education than reading, writing and maths. The National Curriculum has 13 subjects alongside other statutory guidance of British Values, Social Skills, Relationship and Sex Education to name just three. These are not reflected in any published data or as stated earlier even tested. The DfE (2021) guidance for report writing only highlights that: test results, attendance, general progress and strengths and areas to develop are the statutory requirements of annual school reports. Therefore, other subjects outside of the ones tested may never be reported on.

- Principle Four (Statutory assessment systems should be accessible to pupils of all abilities and recognise their progress)

This principle is the one that initiated my interest in this research, the label for pupils of pass and fail and what I could do, as headteacher, to limit this will be addressed later this

research. The other aspect of principle four and the use of language within the reading test I also feel has not been addressed. Reading scores continue to remain steady and national data year on year shows they are usually the lowest result in most schools within deprived areas. The language has not improved, but teachers are having to build it in to their everyday teaching. They are to some extent 'teaching to the test'. They are teaching word meanings and instead of reading texts the children are excited by, they are having to choose more traditional texts to expose children to the language they need. It is interesting to note here that the September Ofsted framework (2019) focusses highly on the teaching of reading and developing a love of reading. It is the one compulsory deep dive primary schools have from the four they will be inspected on. As this language is only acquired through reading in many schools, leaders may now have to decide if they want a love of reading or to give pupils the language acquisition, they need by year 6. Yet another headteacher decision with significant consequences.

In relation to 'accessible to all abilities' this was also identified in the House of Commons (2017) review where it highlighted that many of the criticisms of the new assessment system was its inaccessibility for pupils with SEND. The focus on spelling and handwriting could disproportionately affect pupils with dyslexia or dyspraxia, and there has been criticism of the level of difficulty of the tests. A survey conducted by NAHT found that "an overwhelming majority of respondents (98%) reported that tests at KS2 were not appropriate for children with SEND, with 82% reporting the same issue at KS1". (p.12)

It should also be acknowledged that accessible has only been taken to mean in terms of academia and not from an emotional or mental health stance. Despite pupil mental health and wellbeing becoming a statutory aspect of schools from 2021, this is not accounted for in a testing process. Pupils must still sit the one-hour test without help and in silence regardless of their home or personal circumstances. Pupils living with domestic violence or who are looked after (i.e in care) or who have anxiety can, at best, have additional marks

applied for following the test. However, while this may increase a test score for them. It may not have supported the teacher/pupil relationship, helped their mental health or even given them time to discuss their issues before a test. The House of Commons Review (2017) identified that the new testing system itself had impacted on wellbeing of both staff and pupils.

Progress of pupils is now recognised more, and progress data of the cohort is published. Individual progress of pupils is also shared with schools. While this is may be seen as welcome, as discussed earlier, it also has the potential to create issues. All pupils are given a converted score based on their KS1 data. From this schools can identify what individual pupils will need to secure the score needed to achieve progress at the end of KS2. This means targeted support in tested areas and as each pupil's progress is grouped together for the published data, only one pupil not achieving 'progress' will impact negatively on the published data. This can mean pupils are significantly challenged to achieve higher to allow for 'more than' progress to be made or to compensate for pupils who may not get there.

- Principles Five (Progress should be valued over attainment in statutory assessment)

It is this principle and principal 6, that I feel are the only ones that have had yielded any results. Progress is now published and is beginning to hold as much weight outside of the school environment as attainment. Parents are looking at both and indeed parents within my own children's school where attainment is always high, are challenging the senior leadership team on why progress is so poor. It has also given Ofsted another way to assess the school's impact and schools that historically would have achieved requires improvement for data, can now show they are making a difference. The starting points of pupils have finally been acknowledged, albeit to a small degree.

- Principle Six (to reduce the number of statutory assessments)

This principle has recommended the removal of KS1 SATs. However, this will be replaced by a baseline in foundation and so it would be unfair to completely imply the tests have been removed. They have been moved and given a different guise. The main point missed by the recommendations is that KS1 statutory assessments were never the issue. They were the baseline for the progress measure, but they were not the high stakes which KS2 statutory tests are. In leaving the very tests that carry the highest impact, nothing has really changed. The newly planned times table test, introduced in year 4 is not reducing the number of tests pupils sit but increasing them. The House of Commons Review (2017) supports this with their findings that:

plans to remove statutory assessment at Key Stage 1, which it hopes will help ease teacher workload and reduce “the overall burden of statutory assessment”. However, the accountability system relies mainly on data from Key Stage 2 tests so this will not have the desired effect across the whole of primary school teaching. (p.16)

School leaders are therefore in the same position they were at the start of the new assessment system some 6 years ago, despite reviews and recommendations to the contrary.

Holt (1969, p.52) identified two reasons why children are tested. The first is to ‘threaten’ them into doing as we want and the second is to give us a ‘basis for rewards and sanctions’. He identifies the threat of a test as making students work and the outcome of those tests enable us to reward them. The economy of a school, like that of most societies, operates with greed and fear and the links between the economic mechanisms of the neoliberalism and macrosystems will be discussed further in this research. While this does link to the previously discussed government stance, that a way to motivate pupils and raise standards is via testing. I would disagree with this stance as a blanket statement and instead suggest that those in more deprived areas, who may struggle with education, see the tests as a reason to not try and turn off from education. If they do not engage, they will not fail. The

ecological systems use fear to a high extent. This will also be addressed in some depth in later chapters.

I thought it would be interesting, at this point, to look at future educational proposals following Amanda Spielman's decision that Ofsted's primary focus from September 2019 will be the curriculum and not data. She has made it abundantly clear that a data driven education system has created a narrowing of the curriculum and thus teaching to the tests. It was two years late, but Ofsted appeared to have listened to the review outcomes and the voices of educational leaders. Cultural capital became the new buzz word and a tight focus on how schools prepare and offer pupils what they need is a significant thread within the new Ofsted Framework. This is in stark contrast to the DfE and their promotion of league tables and floor targets. It appeared that educational leaders now need to make a further choice, this time between Ofsted and the DfE. Leadership teams and classroom teachers will now need to decide between results and teaching in order to achieve these to appease the DfE or teach a broad and balanced curriculum to achieve a 'good' Ofsted. As a headteacher, the 'side' chosen of either: DfE or Ofsted is analysed in depth in later chapters.

## [2.8 Growth Mindset, Carol Dweck](#)

Carol Dweck's (2012) Growth Mindset is something often quoted in relation to developing resilience. Resilience itself as a piece of research is vast from Hart, Blincow and Thomas (2007) identifying that resilience can be simply maintaining something, to Daniel and Wassell's (2002) 6 domains of resilience. It should be acknowledged here that the very definition of resilience is not agreed, with the various concepts of Hunnybun (2012), Prilleltensky (2005), Garcia-Crespo (2021), Howard et al (1999) and Mohaupt (2009) being discussed later. Within Dweck's (2012) Growth Mindset theory she identifies two beliefs, which she names as mindsets: fixed and growth. Individuals move along the spectrum between fixed and growth but usually tend to be nearer one mindset than another. A fixed

mindset holds a belief that qualities are 'carved in stone'. Therefore, creating a reason to prove oneself over and over, as individuals cannot improve or to not try at all and protect what they have. Not the type of pupil to respond to testing in order to raise engagement and attainment. A growth mindset believes that their basic qualities can be developed through effort and that they will 'get there' eventually, in short, pupils that would welcome testing as a means of raising attainment and engagement. It was these attitudes that had initially led me, as a school leader, to look at this method to improve pupil resilience. The culture of the school had been for many years one that bred a 'can't achieve' attitude. Pupils did not believe in themselves and certainly did not want to do anything that removed the status they had fought for. This also reflects the previous self-handicapping research of Rhodewalt (1994) and Robinson's (2014) research on attitudes and resilience are linked to how a person feels about the activity at hand. If a pupil feels they cannot achieve and are treated as though they can't, their outcome seems quite conclusive.

Dweck's (2006; 2012) studies illustrated that people find it difficult to estimate their own abilities. However, most of the inaccuracies were from people with a fixed mindset and my experience would support that. Dweck explains that many fixed mindset pupils, have a negative opinion of their abilities. Those with a growth mindset are extremely accurate and were fully aware of what they could do and the things they needed to work on. Dweck (2012) explains this as logical. If a person has a growth mindset and believe they can develop, then they need accurate information on their current abilities to improve, even if this is unflattering. This then leads Dweck (2012, p.22) to claim that those with a growth mindset thrive on challenges, while those with a fixed mindset thrive when things are 'firmly in their grasp.' While she believes it is common for students to turn off school and even adopt an air of indifference, we make a mistake if we think any pupils stops caring (Dweck, 2012). That the attitude of not caring or self-sabotage, is not fixed. Dweck (2012) thus argues that because a growth mindset is constantly growing, then an individuals' true potential can never

be known. This aligned with the newly created assessment system within my setting and that time and effort are different for each person and in turn potential is not something that could therefore ever be measured. Dweck's (2012) research puts forward the argument that praise for effort and not ability is the key to a successful mindset. This is in contrast to decisions being made based on test results alone, as policies had stipulated.

Dweck's (2006) initial research, based in America, focussed on a set on students completing 10 problems from a nonverbal IQ test. Following the test completion, some students were praised on ability and the others on effort. The research team identified this as the turning point for creating a sample of now two camps – fixed and growth mindsets. Dweck (2006) discusses that those praised for ability demonstrated a fixed mindset and rejected the option of a new task that challenged them, believing they did not want to risk exposing any potential flaws of their newly acclaimed ability. If they had been told they were clever, they would not risk losing that. In contrast those praised for effort who Dweck (2006) identified as displaying a growth mindset, had 90% of the group attempt a new challenge. They had nothing to lose trying it and would be praised again for 'just trying.'

At the next series of research test, both groups were given harder questions than the first. Those originally praised for effort and who Dweck (2006) had identified as a growth mindset, simply tried harder and did not see any errors as a failure or reflection on their abilities. The students praised for ability and who had been deemed as having a fixed mindset, now displayed doubt that perhaps they were no longer clever or intelligent and saw the test as a failure reflecting their own abilities.

In her 2000 papers, initially over two thirds of the students in both groups mastered the materials given, (76.6% of the helpless response group and 68.4% of the mastery orientated group). Dweck (2012) raises that these results are in line with previous findings before any failings occur. However, once tested with material that posed a 'threat. The mastery

students retained similar grades with 71.9%. The helpless group dropped significantly to 34.6%.

Dweck (2012) drew 3 findings from her later research:

- 1 Those with a growth mindset find success in doing their best; through learning and improving.
- 2 Those with a growth mindset find setbacks motivating and see them as informative.
- 3 Those with a growth mindset in sport took charge of the process that brings success and can maintain it.

Dweck (2012) develops this research to identify that every word or action can send hidden messages relating to the different types of mindsets. She raised that the messages given by teachers will either demonstrate that they are interested in the person's development or that they are judging based on their ability. This reflects the issues previously discussed relating to the DfE and Ofsted decision and the moral decisions needed to be made by teachers and school leaders. It also links to Rosenthal and Jackson's 1968 research on the self-fulfilling prophecy claiming that teachers' expectations can significantly affect student performance. They speculated that the teachers' manner, facial expression, encouragement and degree of support produced a self-fulfilling prophecy (Haralambos and Holborn, 2013).

Perhaps Dweck's (2012) biggest statement from her research is the most hard hitting, leading it to be used in many schools and the creation of a new buzz word. "After 7 experiments with hundreds of children, we had some of the clearest findings I've ever seen. Praising children's intelligence harms their motivation and performance." (Dweck, 2012, p.175). This runs against the new governments system in giving success to those that perform well.

Dweck's (2000, p.13) research also had begun to raise alarm bells, "All long-term goals involve obstacles. If obstacles are seen as posing a real threat and if they prompt self-doubt and withdrawal, then pursuit of these goals will surely be compromised."



If, 21 years later, research is still saying that focussing upon ability is harmful and limits life goals, then surely educational leaders, at the very least, should respond. Had I responded?

In 2016, Dweck's research became known in many educational authorities and schools suddenly started using the word 'yet', under a misconception that adding this one word would help change years of feeling a failure. Don't worry you can't do it yet, but you'll get there! Whole CPD courses on mindset were rolled out to schools. At a time of falling educational standards, Dweck (2012) offered (at first glance) a relatively quick and easy fix, with very little financial implication on already strained budgets. Dweck's research was aimed at teachers and how they personally could singlehandedly have a lasting impact. Her explanation of simply lowering standards to boost self-esteem and allow success, 'creating poorly educated student with an entitlement to praise and easy work would help no one' (Dweck, 2012, p.193). Yet raising standards, without giving students the necessary skills to achieve them would also fail. Dweck's (2012) theories hit those wanting to raise standards and those wanting to nurture and create a more child friendly education system. It led staff to believe they could finally stop choosing between their conscience and the needs of pupils and the requirements of governments to raise standards and do the job they are paid to do. Growth Mindset offered both; a way to raise standards but also offer a supportive classroom, interested in the development of each pupil and not test scores. Growth mindset appeared to offer a win/win situation. As a headteacher, it appeared to be the answer.

In her research Dweck (2012) uses the term 'great teachers' in several chapters. Educators reading this, or simply those sent on the CPD course, would certainly want to be a great teacher and so this must be what they need to do:

"Great teachers believe in the growth of the intellect and talent." (Dweck, 2012 p.194).

"Great teachers set high standards for all their students, not just the ones who are already achieving." (2012 p.97)

To support this, Dweck (2012) gives many examples of sporting stars and aspirational teachers that have made a 'difference' by using mindsets and their own attitudes towards learning.

This is supported by Gipps and Stobart (1995) who raise labelling as a danger within assessment: test scores or exam grades can determine ways of thinking about children and raise two issues with this. Firstly, the test scores may not be accurate and secondly that they can affect teachers' views about what children are and are not capable of.

Care has to be taken when categories of pupils are identified as misclassification can be damaging and the labelling syndrome should be avoided. Children who are labelled as belonging to one category or another will probably receive stereotypical treatment at the expense of their own learning needs (Sumner, 1987). In using Dweck's (2012) 'labels' the impact, at face value, is reduced. Pupils are either trying or not, but their attainment is not mentioned and that is critical.

## 2.9 Conclusion on the Impact of Dweck's Growth Mindset

Dweck (2019) herself has recently acknowledged that growth mindset is more 'complex' than she had originally thought (Severs, 2019). Dweck (2019) is on the verge of publishing her most recent research on growth mindset, stating:

There was evidence that growth mindset had been tested and shown to work in ways that were meaningful to students – Ok that was the evidence, but the evidence did not speak to how to implement in the classroom. That part was not yet evidenced based. Research takes place over many years. We continue to probe and validate and extend it. (Severs, 2019, para.5)

Dweck (2019) was clear that whilst she has produced a piece of evidence that identifies under certain conditions this happens, she has not yet looked at all the other conditions for learning and needs the feedback of teachers to do this. She identified that although her

research has been deeply gratifying, it has highlighted the fact that we need much new research to find out how to put theory into practice more systematically and effectively. Dweck (2019) explains three findings when reviewing the body of evidence using Growth mindset:

- 1) All research involves struggle and challenges, and when you take your research and its implications into the real world, the challenges multiply.
- 2) There is a misunderstanding in the field about what counts as a noteworthy effect size for real-world outcomes, such as grades and test scores.
- 3) The third, and quite unexpected, challenge came from practitioners' misinterpretations of growth mindset and how to foster it.

To date, Dweck has not identified steps forward with implementing her original theories into the practical classrooms of any education system. The criticisms of the mindset theory continue and cross-referencing local authority CPD on mindset, show that perhaps educationalists are also taking a step back in the promotion of this practice.

The Building Learning Power (BLP) principles, Claxton et al. (2011) discuss Cantonian High School, Cardiff who repeated Dweck's (2012) research over a three-week period, praising group A for effort and group B for ability. At the end of the three-week period, the findings suggested that group B were worried about grades, no longer enjoyed the tasks, were preoccupied with comparing results, would choose easier tasks and 30% did worse on questions completed originally. In contrast, group A became more confident in their skills, focussed on problem solving, were interested in learning and 50% did better than they had done originally. Whilst this does show some support towards changing mindsets, the results are arguably not as conclusive as the ones put forward in Dweck's (2006 and 2012) research.

Dylan Wiliam at his Bryanston Education Summit, (2017) raised the question of growth mindset and the impact on student achievements as not significant, noting that despite three attempts, Dweck's (2012) results have never been replicated. He also raised that there are few, if any, examples of schools that have successfully managed to change their own students' mindsets.

It is worth noting that prior to Dweck's (2019) announcement of potential issues within her growth mindset, Michigan State and Case Western Reserve University identified that a growth mindset intervention do not work for most students, in most circumstances. The universities research highlights that the academic benefits of growth mindset interventions have been overstated and are not supported by the evidence (Sisk et al., 2018). The team found that while there was a correlation between someone having a growth mindset and doing well academically, the correlation is extremely small. Stuart Ritchie, University of Edinburgh raises a valuable point for the many teachers that implemented Dweck's research:

The results of this study should make teachers – many of whom are very interested in the topic of mindset, and have changed their teaching practice because of it – seriously reconsider the amount of time, effort and resources their schools invest in promoting 'growth mindsets' in their students. (Beall, 2018).

While the notion of mindsets appears to produce quick gains, the effort taken has had little impact on changing mindsets. Therefore, it would be unfair to conclude that Dweck's (2006: 2012) research has made little positive contribution to education. It may not have had the impact sold; raising attainment, changing attitudes, developing resilience and so forth but it has made teachers think and review their own practice. For that alone, it is a valuable tool. Long before Dweck (2012) hijacked educational thinking, research in this area was vast. Drummond (2008, p.17) discusses research carried out in 2004 with her colleague Susan Hart. In this research they propose an 'alternative model of anti-determinist pedagogy and

the concept of ability with the central skills of learning capacity'. The research cites the work of Chitty (2001) who highlighted the failing of the education system to challenge the notion of fixed ability.

Earl and Katz (2008, p.94) discuss the 'habits of mind' that create critical, creative and self-regulated learners. Costa (2008) and Bourdieu (1990) also raise habitus and habits of mind as ways of understanding. Turner's (1982) theories of social influence on behaviour and attitudes suggest that this can occur even to the extent that individuals are apparently willing to deny the evidence of their own sense to go along with a majority view. Weiner (2000 p95) discusses the theory of attribution, whereby successes and failures are explained (attributed to) controllable factors and from this will follow 'adaptive tendencies'. Dweck (1989 p.85) argues that Weiner's (2000) attributes themselves are 'not so important'. She highlights motivation as being directly linked to achievement goals, whether these be learning goals or performance goals.

Everyone knows negative labels are bad but in a fixed mindset, both positive and negative labels can mess with your mind. When you're given a positive label, you're afraid of losing it and when you're given a negative label, you're afraid of deserving it. (Dweck, 2012, p.75)

The very notion of labels Dweck (2012) refers to, in its simplest terms, are used with the use of fixed and growth mindset labels. Despite this, the impact of the use of labels has been a main takeaway for educators.

Perhaps the adding motivational 'yet' to every sentence, which was the simple answer to increasing growth mindset, has in itself masked the complexities of the theory behind Dweck's (2012) work and the notion of a constructivist classroom. Something Dweck (2019) herself has identified. Richardson (2002) clearly highlights 'constructivism as a theory of learning and not teaching'. Biesta (2013) states that one idea that has significantly changed classroom practice around the world is constructivism. This classroom takes its inspiration

from a range of theories from Ernst von Glasersfeld's (1989) radical constructivism, to Jean Piaget's (1936) cognitive constructivism, to Dewey's (1938) transactional constructivism. While Dweck's (2012) growth mindset has not had the success initially thought, it has succeeded in making schools think about learning and therefore in turn what and how they teach. It brought back to the forefront, Williams and Thompson's (2007) idea that we all play the role of both student and learner. At best, growth mindset gave a glimmer of hope to teachers and at worst it made professionals think about their practice. Therefore, whilst not the 'answer' it was perhaps packaged to be, it is still rooted in some of the leading research and practices with the current worldwide education system. As a headteacher, this is the beginning of my own questioning of the ecological system in which I was placed. It stirred the personal debate on why can't my student achieve the same as those in other schools. Is it us stopping them?

Within the context of my research, I will revisit resilience and mindset when discussing the findings. The overarching aim of understanding the impact of a testing regime and the role of a headteacher in limiting impact, will ensure that as a school we must look at what we can do to address the research findings. The new system will either show no, some or significant impact on pupils and this will need unpicking and explaining if the aims of this research are to be addressed.

## 2.10 Teacher Perspective Research

Burkhardt and Schoenfeld (2003) suggest that despite over one million articles being held in the Education Resources Information Centre, educational research is often criticised as neither useful or influential. As a headteacher, one of the reasons Dweck became known to me was the ease with how it was explained and the simplicity of implementation.

Williams & Coles' (2007) research identifies that while teachers are positively motivated towards the use of research, their actual use of research information was limited.

McAllister (2018) raises that member of many professional groups have opined upon educational policies for many years, while teachers have been silent. Yet teachers have the direct experience of the consequences of the educational policies and are in a valuable position of being able to identify problems within school that otherwise would not be raised in a public forum.

As this research is based on a teacher perspective, perhaps firstly looking at teachers' views on educational research would be the apt starting point. Drill, Miller and Behrstock-Sharret (2012) highlight five key findings from their research on teachers' perspectives on educational research.

1. Teachers are not opposed to accessing and using research.
2. Scepticism can be reduced when research comes from a source that teachers trust and if the findings work in their classroom.
3. Teachers turn to research after consulting other resources.
4. Teachers give seeking and reading research low priority.
5. Teachers are less likely to use research if they do not see a connection between the study and their own classroom.

I address these points within the context of this research, and it is possible to see some correlations. As already evident from the review of a small number of government policies, research and decisions by those not at the 'chalk face' have been the basis for many changes in education and these changes have not always been welcomed. It would suggest that many teachers are indeed sceptical of research based on their perceived misuse of research for political gain.

Whilst not known at the original time of the literature review, pracademia and thus pracademics are something that has since become very much part of this research.

Pracademics are considered to belong in both the world of research and practice, yet belong to neither (Panda 2014). Hollweck, Netolicky & Campbell (2021) identify pracademia as a bridge, bridging or crossing the gap between practical and academic fields. They argue that pracademia as a concept has a valid place in education, with pracademics connecting the dots between research and the classroom. They argue that the concept of pracademia suggests that there is a possibility to reimagine boundaries and the roles in education alongside other fields.

Perhaps this is the link Dweck (2019) now needs, for those living within the education system to be the bridge between her original research and the educational systems today. In relation to this research, the very basis of this research to understand the impact of a testing regime and my role within it. In remaining with the bridge metaphor; to develop that bridge of understanding from a loose wooden structure to a robust solid bridge, all while controlling the traffic crossing. It will span across educational policy and the ecological system it sits within, to the day to day running of a school and the impact that has on the pupils.

### 2.11 Teacher Perspective Research Conclusion

Drill, Miller and Behrstock-Sharret (2012) also highlight a mismatch between educational research and the demand of teachers. They suggest that to balance the supply and demand equilibrium that researchers should:

- Get to the point
- Provide examples of the research in real class situations
- Write in an accessible manner
- Emphasize how findings can help teachers.
- Highlight the attributes associated with the research
- Be clear on the context of the study in relation to other settings
- Be proactive with engaging teachers to disseminate the research.



This is further supported by Williams and Cole's (2007) research on teachers' approaches to finding and using research evidence, which concluded that while teachers are positively motivated towards the use of research evidence, their actual use was limited. The main barrier being a lack of time and access to sources.

Schoenfeld (2009) notes that many educational researchers do not disseminate their research in ways which are useful to practitioners. This research, as is clear, is aimed at the layman. Those in the classroom and written in narrative that the teacher profession will resonate with. It will also be shared within the setting, in the hope of being useful. The basis of undertaking the research is to understand the implications of a new testing regime, the participants of the research are clear in that a shared understanding in order to respond is needed.

A further issue with teacher perspective research is the tendency for these to be small scale and for the purpose of the researcher themselves. They are rarely undertaken on a large scale to influence outside the research setting. This research is no exception. However, I would argue this makes it no less valid and due the researcher having a genuine interest in the subject matter, it is research for a real purpose.

In relation to Pracademia, many of the issues already identified would also fall into this area. In this research, as a pracademic, the research is small, based within my own setting and limited very much by time and the pressures of the 'paid job'. However, that sits comfortably within the reason for researching, to investigate the impact of testing and what I, as headteacher can do. That the pressures of the ecological systems in which I work do not stop and that prioritising and addressing these pressures is fundamental to this research.

The initial literature review was an in-depth analysis of the policies which had brought about the research and the things already in place within the setting, such as mindset. As a

pracademic, I did not stop reading and researching once the data collection period had concluded. Therefore, further research came to light following the conclusion of the literature review and this section will focus on resilience, curriculum design and power within education. This second section of the literature review is relevant to the research findings but did not interfere with the research questions, design or methodology. This literature became a useful analysis and discussion tool, when looking at the research findings.

Following on from Dweck's (2012) research and the links already made to existing research on resilience within this literature review, it became apt to look at resilience and curriculum design in a little more depth. To focus on what we mean by resilience.

Given this research is underpinned by policies which focus on resilience and that the research aims to look at the role of a headteacher, then curriculum design (which headteacher's ultimately create) are also discussed within this section of the literature review.

Finally, within this section of the literature review, I draw upon the research in relation to power within education and in particular the power of headteachers.

## 2.12 Literature Review of Resilience

The Oxford dictionary defines resilience as a noun meaning: the ability of a substance or object to spring back into shape or the capacity to recover quickly from difficulties, toughness. As previously highlighted resilience is a wide subject matter. Indeed, Hunnybun (2012) identifies that there is no single definition of resilience in the academic literature, that there 'appears to be as many definitions as there are writers in the field'. Resilience within education is defined widely, from a student's ability to achieve academic results significantly higher than would be expected according to their socioeconomic level (Garcia-Crespo 2021) to a positive adaptation despite adversity (Fleming and Legogar 2008).

Chandler (2014, p.1) highlights that 'resilience has increasingly become central to international and domestic policy-making' over the last decade. He argues that resilience is the guiding principle of policy governance and one of the key political categories of our time. Indeed, he raises that an all-party group from the House of Lords and House of Commons have called for resilience to be taught in all mainstream schools. Interestingly, this focus was not to improve success but how pupils could change their approach to failure (Chandler, 2014). This fits within Mohaupt's (2009) research that resilience is a process and not a personality trait. Hunnybun (2012, p.19) supports this and highlights 4 characteristics of resilience:

1. It is a process not a trait.
2. Resilience is not stable and will vary at different times of a person's life.
3. Protective factors can be located within families and the community.
4. Resilience is multi-dimensional and should not be implied in all domains.

Real resilience does not simply alert us to improving things, it can also reflect situations not becoming any worse (Hart, Blincow and Thomas, 2007). Prilleltensky (2005) discusses that resilience must go beyond a phrase limited to how individuals cope with adversity.

The international Resilience Project cited in Daniel and Wassel (2002 p12) identifies 15 check points to measure the differing levels of resilience:

1. The child has someone to love them unconditionally.
2. The child has an older person outside the home they can tell problems/feelings to.
3. The child is praised for doing things on their own.
4. The child can count on their family to be there when needed.
5. The child knows someone they want to be like.
6. The child believes things will turn out alright.
7. The child does endearing things that make people like them.
8. The child believes in a power greater than seen.
9. The child is willing to try new things.
10. The child likes to achieve in what they do.
11. The child believes that what they do makes a difference in the outcome.
12. The child likes themselves.
13. The child can focus on a task and stay with it.
14. The child has a sense of humour.
15. The child makes plans to do things.

These are similar to Art Costa's 1980s attempt to discover what 'intelligent' people do when they are confronted with problems. From this he formulated 16 habits of mind:

- Persisting.
- Thinking and communicating with clarity and precision.
- Managing impulsivity.
- Gathering data through all senses.
- Listening with understanding and empathy
- Creating, imagining and innovating.
- Thinking flexibly.
- Responding with wonderment and awe.
- Thinking about thinking.
- Taking responsible risks.
- Striving for accuracy.
- Finding humor.
- Questioning and posing problems.
- Thinking independently.
- Applying past knowledge to new situations.
- Remaining open to continuous learning. (Cited in Lucas, Claxton and Spencer, 2013, p.57).

Whilst, I acknowledge there are many resilience check lists, this is one used within the research setting. It was not a deliberate choice, more given as part of CPD, but nonetheless it is used. It gives a starting point for action alongside informing future curriculum design and lesson plans. There is a danger that tools, such as these, become simply another assessment tick list. However, if used as simply information gathering and another piece of information, it is useful and not assessment for the sake of assessment. As already highlighted by Gibbs and Stobart (1995) a test, only useful for others!

Daniel and Wassell (2002, p.13) also discuss that although there are many factors associated with resilience, there appear to be three fundamental building blocks which underpin it.

- 1- A secure base, where the child feels a sense of belonging and security.
- 2- Good self-esteem and internal sense of self-worth and competence.
- 3- A sense of self-efficacy, an understanding of strengths and weaknesses alongside a sense of control and mastery

Chandler (2014, p.2) raises a valid point in that although resilience seems to be ubiquitous, how the concept operates and the uses to which it is put are not always clear. The notion of resilience theories also has its criticisms, mainly that it is a 'poorly defined construct, laden with white, middle-class values' (Hunnybun, 2012, p.38). Saleebey (1996) accused resilience theories as simply a way to make misery more palatable to society. Howard et al. (1999) also raise the issue that the perspective of children in terms of risk and resilience should also be considered. They argue that children and adults may see the world differently and that adults should not assume there is one agreed definition of the context within which they find themselves.

Given, the conflicting and vast theories on resilience it is perhaps easy to see how Dweck (2012) made things appear so easy to teachers and something they could easily achieve. However, to make the definition explicit within the research setting we define resilience as the ability to bounce back, which fits simply with the Oxford definition. A dictionary is a pupil's main source of information, not research and so using similar wording aided understanding.

### 2.13 Literature Review of Curriculum Design

As will become clear in subsequent chapters and has already been alluded too, attainment and data are not a main driver of the settings or its curriculum design. Keddie's (1973) research of streaming in secondary comprehensives identified a relationship between perceived ability and social class. She concluded that classifications and evaluations of students and knowledge are socially constructed via interaction situations, resulting in those students perceived as having a low ability actually being denied essential knowledge for success. Keddie (1973) also highlighted quite ironically that higher ability students' failure to question what they were being taught contributed to their educational achievement. In Ball's

(1981) three-year study of Beachside Comprehensive school, he identified that most students were eager and conformist when first entering school but this attitude changed due to teachers' stereotypes. Ball claimed the teachers' expectations of the different attainment bands led to strong correlation between banding and performance.

Both of these pieces of research aided what we didn't want in our curriculum design, as neither aligned with Dweck's mindset, which had already been implemented – albeit not fully or correctly.

While not something I had consciously chosen to do, the research setting was following the principles of expansive educators in designing its curriculum. This in itself aims to develop resilience and touches upon Dweck's own research. The setting had aimed to mould pupils' mindsets to the three attitudes of expansive educators:

1. Communal virtues: Honesty, kindness, tolerance and empathy.
2. Self-regulation: Patience, self-discipline and the ability to cope with frustration without 'kicking off.'
3. Learning virtues: Able to deal with challenges and uncertainties through determination, curiosity, creativity and collaboration. Lucas, Claxton and Spencer, (2013, p.17)

The key premise of expansive educators is to help learners to do things for themselves, not being thrown when problems are phrased differently or set out alternatively are key life skills and crucially, good exam passing techniques, Lucas, Claxton and Spencer (2013).

This aligned very much to the ethos of the school and pupils being more than just data. As is already clear, our role is to prepare students for a world and technologies that do not yet exist. Expansive education raises the notion that building real world links, and focussing on these will boost exam performance, not risk it. Claxton and Chambers (2011) when discussing their growth point for the future, identify that it is the attempt to connect what

happens in school with the wider world in which our pupils will need to make their way as adults. Wherever their path may lead, their habits of determination, concentration, imagination and collaboration will be of use. It is the job of a twenty-first century school not just to build these habits in school, but to ensure they are transferred outside the school gates.

As already highlighted, testing within this setting was not high stakes and my attitude as headteacher was that if pupils enjoyed school, data would follow. Pupils have to be in the right place to learn and more importantly want to learn. In order to do this, there needs to be a real purpose.

Learning to learn is centrally important for all students and teachers. But it must not be seen as an alternative to teaching subject knowledge well and with passion, or as a modular add-on, not seen as centrally embedded in all teaching and learning. (Lucas, Claxton and Spencer, (2013, p.44).

As already clear, the setting's curriculum was one where pupils take the lead and are not 'spoon fed' answers. The pupils need to understand what learning is, as do staff. Imparting knowledge and simply getting everything correct, is not learning. Teachers needed to realise it was OK not to know the answers, to learn alongside pupils. That allowing them to be 'stuck' was actually helping. Schools should treat teachers as learners and teachers should also see themselves as students (Williams and Thompson, 2007).

The expansive classroom feeds very well into Dweck's notion of fixed ability when labels are used. Lucas, Claxton and Spencer (2019) raise that ability labels can too easily become a synonym for intelligence or talent. In a similar way to assessment becoming a synonym for testing, the curriculum in place needs to make clear its definitions.

I should raise here that while titled 'curriculum design', as a headteacher, I did not specifically design a curriculum (other than the covering the statutory National Curriculum). As already confessed, expansive educators were not known at the time but we were doing it.

As will be discussed in later chapters, I created an ethos and environment, which then itself created a curriculum. I have already highlighted that this stemmed from the pupils initially, how did I want them to feel, not what did I want them know. This then blossomed into our curriculum. I have given a brief explanation of the development of our curriculum, within the introduction chapter of this research and mention aspects throughout various other chapters. Having reflected upon the role of the headteacher, however, and having completed this research, this remains a consequence of my actions that I am not yet fully able to explain and will be something that continues beyond this research.

#### 2.14 Power in Education and the Role of the Headteacher

The Oxford Dictionary defines power as ‘the ability or capacity to do something or act in a particular way’ or ‘the capacity or ability to direct or influence the behaviour of others or the course of events.’ Christensen (2023) raises that the common definition of power identifies a single connotation of power being held by the powerful and exercised over the powerless and is thus, repressive. He argues that power can be non-possessed and productive and cannot simply be conceptualized, but must be understood as a complex phenomenon. A phenomenon which can be questioned, distributed and come from below. He identifies power as not just one, but several significantly different phenomenon. In relation to power within education, Tahira, Yousuf and Saboor (2021) raise that the role of a headteacher has evolved from a lone top-down authority to leadership dispersed among a variety of people and situations. For the purpose of clarity, it is Christensen’s (2023) general definition of power and Tahira, Yousef and Saboor’s definition of leadership in relation to headteachers, which is used within this research.

These build upon the work of Bourdieu (1990) with his identification of material and symbolic power, with symbolic structures appearing in the form of cultural capital. He conceptualizes the way in which social structures are integrated. This in turn mirrors the many ecological systems this research focuses on.



Within education itself and in particular the power of headteachers, Cossa's (2016) five qualities of power reflect the various types of power that headteachers' hold and reflect that the values they hold will determine how they exercise that power.

The first of the 5 areas are: regulatory and how rules, regulations and laws are enforced. This is at the discretion of headteachers and their interpretation. In essence the value they place upon rules, laws and regulations. This is a significant part of this research and is discussed in depth in later chapters.

Cossa's (2016) second area of financial power is acknowledged but is not a significant area of this research.

The third type of power, Cossa (2016) identifies, is manipulative power. The ability to persuade people's beliefs and behaviours. Headteachers' have significant power to do this for both staff and pupils. Again, this will be dependent upon the personal values and qualities of individual head teachers and will, therefore, be discussed in depth in later chapters.

Cossa's (2016) final 2 areas of power: informational (the ability to control and gatekeep information) and interpretative power (the ability to understand and interpret the spoken word) are a significant element of this research. The decision on what to share with staff is a key thread of a headteacher's role and sits alongside the pressures associated with testing. Both of these areas will be discussed in later chapters. It will show how power by the headteacher has been exercised and the impact this has had on testing.

Winthrop, Morris and Qargha (2023) have recently added a fourth 'P' to their work on transforming educational settings. Alongside purpose, pedagogy and position they also feel power is a significant factor in educational systems. The role of the headteacher and the power they have and that which they exercise form the underpinning of this research.

Zuccollo, Dias, Jimenez and Braakmann (2023) identify the influence headteachers have on their school. Their research shows that an effective headteacher within a primary school can add two months additional progress and reduce teacher turnover. Both of these findings are supported within this research setting and have already been discussed.

### 2.15 Research Questions Within the Literature Review

The two research questions of:

- How can headteachers limit the impact of testing on pupils?
- How do we understand the headteacher's role through the lens of Bronfenbrenner's ecological systems?

Need to be rooted within literature. As already identified, at the start of this research, limited research was available on the impact of the new assessment system and therefore the impact of actions by anyone, not just headteachers was not available.

In analysing government policy, this allows for the links to be made between the various ecological systems and the links between the ecological systems in which those policies are made and the ecological system of where they would be implemented. It will allow the interactions of ecological systems to be made explicit. In turn, it would then allow the role of the headteacher as the manager of one ecological system to be analysed.

As such, this research aims to understand how headteachers can limit the impact of testing. As already stated, it was firstly important it was clear where the policies had come from and what they were asking of headteachers. This would then allow actions to be analysed and any impact to be made explicit.

In reviewing literature on mindset, resilience and curriculum design. It allows both research questions to be answered. It will permit the ecological system of the research setting to be further explained and allow this to be linked to other ecological systems. It will also support

in explaining the actions and thus impact of the headteacher. As already addressed, these were deliberate actions taken by the headteacher, prior to this research.

Whilst the practitioner research review only really focusses on the first research question and the actions of headteachers. It does allow some insight into the views of the headteacher in relation to research, which then naturally lead to the actions they take.

In relation to power within education, specifically. This is very much dependent on the headteacher as an individual, their values and qualities. How they view power and how they allow it be distributed are also key factors. My own stance on education and testing has already been made clear, but my qualities and values, even though uncomfortable, will be analysed within the final chapters.

#### 2.16 Analytical Lens/Theoretical Framework Application of Literature Review

As will be discussed in subsequent chapters Bronfenbrenner's ecological systems are used as the theoretical framework in which this research sits. It is therefore apt, that the policies and research reviewed are also analysed within this framework.

We start top down, as the policies reviewed were written, and begin in the macrosystem and the 'blueprint' for the education system and society in place today. This chapter has identified that educational achievement is a powerful focus of policy makers. It is at the heart of these policies, with the words 'standards' and 'achievement' threaded throughout the majority of pages. In broad terms, it is here that the macrosystem divides and creates the policy makers view of 'achievement' against others working within education. However, I do acknowledge there are educators who will agree with policies, mainly due to schools being linked to the economic market at macro level. The macrosystem reflects the political system it is 'controlled' by, with the broader sociocultural contexts of capitalism and neoliberalism shaping the educational landscape of competitive testing and arguably future

compliance in the workforce, as already raised in previous chapters. This is also reflected in the exosystem's policy language and ideals of educational competition and market place. Testing by the very nature of design is a competition and as such will always create winners and losers. 'Despite rhetorical attempts to fudge the issues, students' good grades only have use in further education and employment as other students simply didn't get them' (Lucas, Claxton and Spencer, 2013, p.12).

The leaders who place banners outside their school gates, highlighting their Ofsted grade or recent test results, are using a market strategy to compete. I would argue these educators are those that believe in the policies or the 'greed and fear' Holt (1969) referred to. These banners mean nothing outside of education.

It is this personal stance that feeds within this ecological system and the two broad definitions of achievement; one linked to achievement within a testing regime, as reported by the policy makers and the other being achievement for each individual pupil. These two beliefs feed through to continue within the exosystem and it is here the impact of this view of achievement has the opportunity to impact. The formation of school curriculums, policies and even daily timetables will have been impacted on by the view of the leadership team within the macrosystem. Within this research, as will become clear, I hold the view that achievement is more than a performance in a test and share the view of Kilpatrick (1925) that education should be more than testing. Perhaps, on reflection, a form of resistance for those 'working in fear' or the 'I'll be nothing' pupils who are not heard. The community I serve, with significant deprivation and lack of cultural capital, are already disadvantaged before their educational journey starts. The neoliberal policies and marketplace culture it creates do not lessen this disadvantage, but exploit it. If I can in any way, lessen the impact and narrow that gap for my pupils, then I will of course do so. I will not define my pupils as the simple test they are forced to take.

These views then feed into the mesosystem, where the daily school interactions occur between staff, pupils and parents. I would argue this is the most influential level of the ecological system and it is here that the larger government policies created are implemented, though this will form many guises dependent upon the view and impact of the macro and exosystem. It is here that the literature review reveals the beginning of educational dilemmas between those interacting daily with pupils and those removed and rooted firmly within the macrosystem. Within the mesosystem as the educators enforcing policies, which is now clear, they have had little input nor necessarily agree. Here, within the mesosystem can be seen the first glimmers of impact and conflict resulting from the outer systems.

This conflict within the mesosystem can impact on the microsystem, as will be discussed in some detail in later chapters. Here individual educators follow through into their own classrooms and practice, the policies at both national and school level. The culture of the school and leadership team impacts greatly on each sub system and the impact it may have. The more conflict within each system, the more variety of impact seeps through to the pupil. In this research the focus is the impact of testing and thus test anxiety, successes and failure are created at each level but the impact on each level is dependent on the conflict and disagreements within each ecological system. In short, the decisions made by me as headteacher.

Within this research setting, a collective understanding of achievement is acknowledged and while conflict does arise on occasion, it is not around the importance of testing. As will become clear, the impact on policies and educational research is dependent on how direct a route it has taken through the many ecological systems.

As this research aims to understand the impact of testing and the role of the headteacher in this, then it would be apt to discuss my own stance in relation to Bronfenbrenner's Ecological

systems. As already highlighted, my own view on achievement is not necessarily the same as those who lead the educational macro system. It is acknowledged, but not taken into the mesosystem in which I control, as the headteacher of the school. If I am able to refer back to the bridge metaphor, then as the lead within this system the flow of traffic over the bridge is controlled, both in terms of information and also people (staff). It is this 'control' and the basis for these decisions that will be analysed and is a main aim of this research.

However, I also acknowledge that Bronfenbrenner's Ecological systems can and will be interpreted differently within that vast array of research in which it is framed.

### 2.17 Literature Review Findings

While there is limited research on the current testing regime and the role headteachers can play regarding impact, there are branches of individual factors: general testing implications, the role of headteachers, the impact of policies to name a few. This research aims to bring together two aspects – the impact of testing and the headteacher role in this.

As may now be clear, educational policy is created in isolation within one ecological system. The impact of these policies is also not known by those based within this particular ecological system, due to being created in isolation, or as shown within this review, with consultation but then disregarded. Or the impact is known but not considered greater than the policy itself and so again disregarded. Given the other ecological systems are situated within this policy making system, this research is not able to answer the conscience of the mesosystem and what it knows or does not know. It is the impact between the two ecological systems, between those creating policies and those living them, that is parallel to this research. This research reflects very much the creation of a new testing regime created in a mesosystem and the impact on pupils in the microsystem, with my role as headteacher becoming the bridge between the two.

Alongside this run conflicts between research itself and the day-to-day case of education, as highlighted with the work of Dweck (2012) and indeed the scepticism of practitioner research. Again, this in itself is bigger than this research and warrants its own research. However, what can be applied and is part of this research, is the role of headteacher. Their views on research and, their actions towards research are very much reflected within the microsystem they run. All of these will impact in some guise, either knowingly or not.

## Chapter Three: Research Methodology

### 3:1 Introduction

This chapter discusses the methodology, research methods of semi structured interviews and reflective diaries and the factors influencing the research. This chapter gives an overview of the chosen methodology and design and presents Bronfenbrenner's (1979) ecological systems as the theoretical framework underpinning the research. I also make clear my positionality, and the ethical considerations undertaken, before discussing at length the sampling and chosen research methods.

The purpose of this research was to investigate the impact of formal assessments (SATs) on primary school pupils, from the perspective of teachers, while addressing the research questions:

- How can headteachers limit the impact of testing on pupils?
- How do we understand the headteacher's role through the lens of Bronfenbrenner's ecological systems theory?

### 3:2 Methodology and Design

Stake (1995, p.43) suggests "The functioning of research is not necessarily to map and conquer the world but to sophisticate the beholding of it", and this research is to aid improvements through effective policy and practice within a specific school. Qualitative research was the obvious choice, given that Kvale (1996) raises "If you want to know how people understand their world and life, talk to them" (1996, p.1).

Education involves a wide range of variables: health, home, teacher, expectations, materials. These cannot all be effectively controlled and assessed, and 'theories of education cannot be based on scientific experiments of the kind used in some physical sciences' (Sutherland, 1988, p.3)



Pupil and staff voices need to be heard, if a genuine understanding is to occur, which is why qualitative research was chosen. In line with Kvale's (1996) 12 aspects of the mode of understanding qualitative research (life world, meaning, qualitative, descriptive, specificity, deliberate naivete, focussed, ambiguity, change, sensitivity, interpersonal situation, and positive experience) my research falls within life world and meaning. The topic of the research is the lived world of staff and their relationship to it. Its purpose is to describe and understand the subjects experience through discussions of mutual interest. It is also the 'meaning' that I seek, to understand the central themes in the life world with a priority on understanding what the participants say (Kvale, 1996, p29/31).

A literature review was conducted, on both academic and non-academic databases, and general internet searches were used to identify relevant literature. The review was not a systematic review, as there was no strict requirement on the areas to focus on. Initially, 'testing' and 'impact' in general became the main key words, yet this yielded results not linked to the new system and became too broad to narrow down, given the data was to run over a full school year and therefore had not yet been concluded to give any specific areas. In educational research, it can sometimes be 'difficult to find significant bodies of empirical evidence on relevant topics of interest, meaning the relevant research literature is quite small' (Punch, 2011, p.96). Many of today's interview studies often start with little or no theory of the themes investigated, and without a review of the literature (Kvale, 1996).

Therefore, the decision was made to review the policies that underpinned the implementation of the system being researched. This was in the hope that the policies were based in research literature. After initially reviewing the policies, it became apparent that other areas were then possible to review such as curriculum and resilience.

From this, targeted journals and websites could be identified, focussing on primary education and psychology (as impact is a key aim of this research). Initial starting points were *The*

*Review of Educational Research* published by the *American Educational Research Association* and *The British Educational Research Journals*.

Given the many references to resilience in the policies reviewed, this also became an area to review. As Dweck's (2012) mindset was already something in place within the research setting at the time of the research, it was a logical decision to also review the literature on this. This would allow the results of the data to be understood with a clearer understanding of the research setting.

As the research also relies solely on teacher perspective, a review of research in this field was also conducted in the aim of underpinning the validity of this chosen method.

The literature review, therefore, focussed on secondary sources ranging from academic journals, to books, reports and 'grey literature', including UK government policy documents. Due simply to ease of access, these were journals accessible within the universities' databases. When deciding on journals available through the search engine, date of publication, relevance to primary school, those with key words linked to systems within the research setting such as deprivation, resilience, creative curriculum etc were chosen. This ensured that any findings could be linked back to the literature.

The initial literature search was carried out between September 2017 and July 2018, with additional literature identified through reviewing bibliographic references and changes in educational policy and inspection until September 2019. Relevance to the subject and methodological rigour were considered when attributing weight to the different sources. Policies relevant to the research questions; the value of teacher assessment in contrast to the impact on teachers and pupils and is assessment working as intended, were logical starting points as they underpin the implementation of the tests being research. A review of the strategies used within the research setting were also to be a key review, to understand the school environment and the attitudes of staff and pupils on which the research is based.

Where necessary, short caveats have been added to the description of the research to highlight limitations of the studies. The literature review mainly focused on literature published between 2012 and 2019, as this is the start of the system being researched, and the period following the research during the data analysis.

### 3:3 Sampling and Selection

Initially the research design was to compare the teacher perspectives across a range of primary schools: size, location, Ofsted grade and so on. However, this soon became problematic in terms of time, comparability etc and in honesty became unmanageable. Each school has a unique context and sets of factors influencing the teachers' perspectives. The more a case study is instrumental, the more certain contexts may be of importance (Stake, 1995). "Case study research is not sampling research. We do not study a case to understand other cases. Our first priority is to understand this case" (Stake, 1994, p4)

The sample choice within this research was limited to the staff within the school of which I was headteacher. I acknowledge this group does not represent the wider population, but the purpose of the research is to investigate this setting in depth and analyse my role as headteacher, not create generalisations. The real value of a case study is not generalizability but particularisation, to take a particular case and come to know it well, not how it is different from others but what it is and what it does. Therefore, the emphasis is on uniqueness, and this in turn implies knowledge with regard to the differences of others, with the 'emphasis on understanding the case itself' (Stake, 1994, p.8).

I do not hide that this research is to benefit the school which I lead and therefore, in turn the experience of the pupils within it. Punch (2011, p44) gives four advantages to teacher-own-classroom research. I take each advantage in turn, presenting my reasons for undertaking the research, and the research design.

1) Convenience. This research would involve little travel and had few logistical problems. It would not necessarily add any more time to my working day and if things needed to be clarified or repeated, then the participants would be to hand.

2) Access and consent. Given participants were seen daily, there was time to explain the research benefits, design, and elicit consent.

3) Relevance. The research was relevant not just to me as headteacher and the staff participating, but to all staff. If the school could improve things, it genuinely wanted to and so the relevance of the research within our own setting was never in doubt.

4) Insider knowledge. This research would access the participants existing understanding of the school. They, and the researcher, were aware of the social, cultural, and micro-political aspects Punch (2011) refers to.

To counterbalance the four advantages Punch (2011, p.45) highlights four disadvantages; bias and subjectivity, vested interest, generalizability, and ethics. These are discussed in more depth later.

All researchers have some form of 'position'. The researcher position has both strengths and weaknesses, advantages, and disadvantages. While insider researcher may bring greater understanding and less objectivity, outsider researcher brings more objectivity but less understanding (Punch, 2011). If one had to be chosen, then understanding over objectivity would have been the choice as both researcher and Headteacher. The fundamental aim of undertaking the research was to understand a particular issue – the impact of the new testing regime, within this particular setting.

Kvale (1996) highlights that contributing knowledge to ameliorate the human condition and to enhance human dignity is one aim of social science, and in this case, it is for the pupils.

## Participants

All staff within the two formative testing classes of year two and year six were asked to participate. This would give the viewpoint of 3 teachers and 3 support staff. This was important as it gave wide viewpoints and permitted more than one staff member per class to participate. I am acutely aware that people within the same class, can see very different things. It also gave me, as researcher, a layer of added security in terms of staff absence or withdrawal from the research. Using this strategy also increased the participant pool. It is not 'uncommon for case studies to make assertions on a relatively small database, invoking the privilege and responsibility of interpretation' (Stake, 1994, p.12). I wanted to widen the data base as much as possible, for the reasons highlighted above, and because I felt that the more opinions, and people, I could check my interpretations with the more truthful the research would be. The final benefit of increased participation was increased anonymity for the participants.

Table 1 shows the research participants. Staff within both year groups (three teachers and 3 support staff who also cover/teach in the year groups) confirmed participation and I believed it wise to use my year two staff as a pilot sample. This involved one teacher and 2 support staff (who also covered lessons). In order to aid confidentiality of participants and make identification more complex. They were given two sets of participant codes, one for their diaries and one for interviews and check-ins. Their exact length of service is also not identified, to aid anonymity.

Participant	Length of service	Role
P/A	Over 15 years	Cover supervisor/teacher in Y6
C/M	Over 10 years	Teacher, Y6 for over 6 years.

W/J	Over 10 years	Teacher, Y6 for over 6 years.
E	Over 10 years	Classroom support in y5 and Y6.
K	Over 10 years	Classroom support In Y5 and Y6.

Table 1. Research participants.

The biggest stakes from a headteacher's stance are at year 6, with published data and league tables. Year two would, therefore, be used as a sample set to aid the final research design. This was a deliberate attempt to identify any design issues or research constraints with another statutory year group before undertaking the main piece of research. As the focus of the research was statutory assessment, year 2 are the only other year group to be able to offer an insight on the impact of testing.

You are looking at your selection of a subject – a marriage, country or whatever, with one, two or a few being focussed on without any expectation that it represents a wider population. So, it's not a sample, It's a choice, a selection. It is this selection that is vitally important for your study. (Thomas, 2016, p.63).

### 3:4 Research Methods - Case Study

It is important that the methodology chosen reflects that of the most appropriate and least invasive method, as well as supporting the philosophical assumptions of the researcher. It is important that the research questions are put before the method and let this 'determine the method you use to answer it' (Thomas, 2019, p.27).

Yin (2009, p.4) identifies the case study as having the ability to allow investigators to retain the 'holistic and meaningful characteristics of real events' and therefore a case study approach was adopted to answer the question: 'How are formal assessments in year six impacting on pupils?'. Stake (1994) distinguishes three main types of case study: Intrinsic,

Instrumental, and collective. I would suggest this research is both intrinsic and instrumental as defined by Stake (1994) and later further developed by Thomas (2016).

A case study is on a broad spectrum, and so for the purposes of clarity Thomas' (2016) definition of an intrinsic case study/blue sky research/curiosity driven research was initially adopted. The subject was being studied out of interest, pure and simple. This research is interested in the impact of the new assessment and testing regime. The research became an instrumental case study evaluating the impact of the new system, with the 'purpose of understanding the impact in order to improve things' (Thomas, 2016, p.120). Instrumental case studies have the issue as central; this dominates the start and end. These issues are not simple and clean but 'intricately linked to a variety of contexts' (Stake, 1995, p.16). Within this research, these contexts would be the setting, staff, and pupils and reflect the theoretical framework within which it is based.

A case study is an in-depth exploration from multiple perspectives of the complexity and uniqueness of a particular project, policy, institution, programme or system in a real-life context. The primary purpose is to generate an in-depth understanding of the topic, programme, policy, institution or system to generate knowledge and or inform policy development, professional practice and civil or community action. (Simmons, 2009, p.21).

It is also worth noting, given the research is an in-depth study on one case, that Thomas (2016) accredits Darwin's theory of evolution, Einstein's annus mirabilis and Jocelyn Bell Burnell's radio pulsar as all stemming from the study, albeit meticulously, of a case.

Punch (2011, p.120) discusses four characteristics of a case study. The first being boundaries. This research is set within the boundaries of a year 6 cohort, within a set school and timeframe. Secondly that the case is a case of something. Within this research the case is simply testing and the impact this has on pupils. Thirdly, the case study should explicitly attempt to preserve the wholeness and integrity of the case. This research aims to have a holistic view by gathering information over the full school year and with participants

aware of the context and reasons for the research. Punch's (2011) fourth characteristics – multiple sources of data of the participants and their views of the given subject.

Wellington (2000, p.97) highlights the strengths of case study research to be: accessible, illustrative, attention holding and allows for reality to be identified. As this research is being completed alongside full time employment, accessibility in relation to access to participants had to be considered when undertaking the research design. The research is also from a teacher perspective and therefore their reality was vital and in having planned discussions and a year of participant reflections, attention would be focussed on research aims.

Hamilton and Corbett (2013, p.16) also identify case study strengths in allowing the practitioners to be central and draw on reflective data, over a set time. It can also build different kinds of evidence. In relation to this research, the research was designed to be time limited to the first year of the new assessment system. It was also planned to include evidence from group discussions and individual reflections and observations, based on the participants at the centre.

Alongside the strengths, Wellington (2000, p.97) also raises weaknesses with case study use. These are: the generalisability and how easy the studies would be to replicate or repeat. Hamilton and Corbett (2013, p.16) also highlight weaknesses such as personal bias, ethical issues, the time impact of using different data forms and the need to include additional perspectives to focus the researcher.

Whilst I acknowledge the weaknesses identified in using a case study approach, the reasons for undertaking the research, already made clear, outweigh the concerns highlighted. Each area of weakness highlighted had been taken into consideration at the design phase. The areas of case study weakness, such as bias, ethics and additional perspectives were addressed within the research design, as will be discussed later within this chapter. Punch (2011, p.123) takes the stance that 'properly conducted case studies, especially in situations



where our knowledge is fragmented, incomplete, shallow or non-existent, have a valuable contribution to education research'. He raises three strengths of case studies. The first being that we can learn from the study of a particular case. Within this research this will be the impact of testing within a specific year 6 class over a set period. Secondly, only in-depth case studies can provide an understanding of new problems, especially when complex social behaviours are involved. Given this research involves pupil reactions, then behaviours will be noted. Finally, Punch (2011) identifies case studies as an important contribution for other research approaches. The findings of this case could be used as a basis for further research within this research setting. It has the potential to be reviewed and once findings have been identified and acted upon, further research could be undertaken to analyse the action taken.

Donmoyer (2000, p.48) acknowledges three advantages of using a case study, which are fitting within my own context as a working headteacher. The first of accessibility is self-explanatory; I had very quick and daily access to a year 6 team. The second of seeing through the researcher's eyes is the biggest gain within this research. It allows the research to reflect the view in this school, with these pupils at this particular time. His final reason decreased defensiveness is also applicable within this research. He identifies that vicarious experience is less likely to produce defensiveness and resistance to learning. In this case, one of the main reasons for undertaking the research is to learn. There were no preconceived ideas, just a simple desire to understand. The research is not a reflection of the situation but a means to understand it and learn.

Undertaking a case study presents a unique opportunity to focus on social interactions and the meanings that participants attach and use to interpret events. It also offers 'multiple realities; the different and sometimes contrasting views participants have' (Swanborn, 2010, p.16). Given this research includes the views and interpretations of different participants

(teachers, support staff, intervention staff, cover staff) within the same class, it allows the opportunity for a variety of perception angles to be discussed and analysed. As pleasant as interview subjects may appear, it is not a given that they will provide the most valuable knowledge (Kvale, 1996). Therefore, in having many different types of participants in relation to their educational role, it allowed for some degree of 'truth' to be sought. It was hoped a wider collective of data would permit 'the bigger picture' and not the view of just teachers.

While case study was the final chosen method, I had considered a variety of alternative methods and methodology. Phenomenological research was the first path explored given it can use unstructured methods of data collection and its ability to allow data to emerge without looking for set answers. At first this seemed ideal, as it is often 'based on small case studies and actively seeks the views, opinions, subjective accounts, and interpretations of the participants' (Gray, 2014, p.30). Whilst phenomenologists align to my some of my own beliefs in terms of human construction on the world. I would agree that people construct and perceive but I may have been unable to completely 'bracket out my own preconceptions' (Gray, 2014, p.165). I would need to check my understanding with the participants, but I did not feel I could completely take away any bias and would need to ensure that any findings I may conclude, would need to be verified and whilst I could not avoid initially relaying on my own interpretations. I therefore concluded this path should not be taken due to concerns that without the participant verification of my perspectives and knowing myself, that the research would become what I wanted it to be. The research may have aspects of this approach but it not a solely phenomenological piece of research.

Having loosely used action research in previous small-scale research, I again initially thought of using it again. The planned research was within the setting in which I worked, and the data analysis would be collaboration between researcher and participants, but I felt

the cycle of planning, reviewing, action and reflecting would potentially lead to too much being changed before we actually had the bigger picture. While to some extent, changes would be made following the research findings. The 'cycle' of planning, action and reflecting would be concluded after the research had ended and perhaps some years later to review the choices made following the research. It would not be something undertaken until at least a further academic year following the research conclusions. My final path has in some respect followed an action research cycle: a plan of not changing anything for the year, reviewing the impact of the new testing system and then reviewing our practice, but has not been a true piece of action research. Thomas (2016, p.37) identifies a distinction between action research and case study, in so far as a case study is to understand the details of what is happening in comparison to action research, where the purpose is to develop practice. Within this research it was the understanding which was central and the changes to practice would come later.

A further area initially investigated when looking at data collection were focus groups. These are used in case study research and allow for the thoughts and feelings of participants to be collected. As the researcher, I did not want control or even hold influence (though as already stated, this may have happened) over the participants, which has been highlighted as a limitation to this approach. A focus group unlike interviews would allow for the 'synergistic building of data as respondents add to the views expressed by others' (Gray, 2014, p.469). However, a deciding factor in not using this method was a potential breach in confidentiality and that conflict may arise between participants, leading to potentially difficult working relationships later, Kaiser (2009). However, I did consider trying to implement some form of group discussion, which finally became group interviews, a focus group was not used. A further consideration of focus groups is that they need to have a clear focus and thus an agenda linked to the specific questions it hoped to find, Gray (2014). This may be problematic as specific questions were not necessarily known and the research focus was

on the wide area of 'the impact of testing'. A final reason for not using a focus group is that a strength of such method is they 'have a focus with an agenda and seek to answer a specific problem' (Gray, 2014, p.472). I was not sure I had a problem to solve, the basis of this research was to look at the impact of testing and not simply assume we would have a problem with it.

### Pilot Research

I planned the year two sample pilot as individual semi structured interviews which were recorded and are included in appendix 1, as a list of pilot questions. I had a list of broad topics to be addressed and would use these as the basis for the year six interviews, which were to be the main research sample. Following the pilot interview, I sought feedback from the staff regarding the process and was given honest but soul-destroying replies. The questions were 'adequate' but as it was one-to-one and recorded, staff felt it more 'an interrogation' than the professional discussion I had hoped. A participant, using the word interrogation hit home, and doubts on whether the research would ever start began to sink in. The intended questions were around things that were discussed frequently in staff meetings, and CPD sessions, but in my office with a recorder it was seen as head and staff, not researcher and participant. As one participant advised, it was "recorded for all eternity." The existing relationships impacted positively on disclosure, and it was an important finding as it changed the methodological approach.

Had the participants not expressed their concerns, the final research year may not have drawn the same conclusions. The initial pilot study allowed me to demonstrate that I was there to listen and had no ulterior motive or agenda. Perhaps because of this, the final research included comments from participants that potentially otherwise I would not have had.

Several discussions with the year two pilot participants highlighted that they would have preferred the interview to be as a group, rather than individually. This is something I had planned to purposefully avoid, as personal experience has led me to believe that some staff tend to take over, others do not speak, and things soon go off track. The interactions between interviewees can lead to spontaneous and emotional statements about the topic. However, group interaction also reduces the interviewer's 'control and the price may be a chaotic data collection leading to difficulties with systematic analysis' (Kvale, 1996, p.101). However, whilst being time consuming, the pilot was already yielding some benefits.

A further comment from the pilot study alluded to the venue being my office. The warnings of Kvale (1996) had not been known at this point, and as interviewer, I had not established a safe atmosphere allowing the subjects to freely express their thoughts and feelings. The venue had been chosen due to it being quiet, with a door and ease in terms of locking equipment away. In hindsight, this was my comfort zone and not that of the participants. Future discussions were held in the participants rooms or the staff room, with timing used to ensure privacy. Kvale (1996) also raises the valid point that personal interaction during an interview affects the interviewee and therefore the knowledge produced, which in turn affects our understanding of the situation. If I genuinely wanted to understand the impact of testing within the school context, then the interviewees needed to feel as safe and secure as possible.

The final pilot study comment referred to the actual recording of the discussion. The fact participants did not like how they sounded was swiftly insignificant and a discussion on the information being recorded took place. Participants always knew the interview would be recorded but some stated that it had impacted on what they replied, though they acknowledged it was 'stupid' they couldn't help but link it to a police interrogation and evidence. Naturalistic methods appeal as they are readily accessible without specialist

facilities or equipment. It is primarily based in participant observation and interviewing, though these are conceptually complex as interviewing it is not the same as a conversation (Walker, cited in Coe & Hedges, 2012).

Initially, I had assumed (quite wrongly) that any interviews had to be recorded, otherwise it was just a conversation. As a first-time researcher I felt research should be recorded but have no basis for this thought, other than being a novice researcher. However, Kvale (1996, p161) makes it clear that interviews can be recorded through the researcher's 'subjectivity and remembering'. Opdenakker (2006) and later Rutakumwa, Okello et al (2019) also highlight the case for not recording. The researcher can rely on their own memory and empathy and then write down the main aspects. While this may cause limitations in some respects, such as reliance on memory and forgetting key information. It also has strengths in that it can note social interactions and atmosphere, which recording cannot detect.

An agreement on how the interviews could be improved was agreed between myself as researcher and the participants, it was decided that the final research in year six, would be a semi structured group interview, scribed as in staff meetings, which would make it 'less threatening'. This would be typed up after the event and checked for accuracy with participants. Rather than tape record or write furiously, it is better to listen, to take a few notes and to clarify Stake (1995). The questions to be asked were also tweaked slightly for clarity and became a list of topics for discussion, rather than a list of questions. A research interview should proceed more like a 'normal conversation' but have a specific purpose and structure (Kvale ,1996, p.131). In having a schedule of discussion which allowed for the same topics for discussion at the start of the year and end, the research would be longitudinal and would show any shifts in teacher perspectives throughout the research year.

This was then discussed and agreed with the participants in year six and the issues raised with year two were also made clear to them, to understand where the proposals had been

derived from. Perhaps the biggest lesson from this pilot study was that I had tackled the research design as the easiest way for me, not the convenience or comfort of the participants. Once I realised the research needed to be about the participants and their thoughts and feelings, the process did become easier.

Gray (2014) suggests that 'trustworthiness is more important than concerns over validity or reliability' (Gray, 2014, p.185). It was therefore important that the participants had felt involved in the design, had been listened too and trusted that the research had benefits for those participating, it was not something being 'done to them'.

#### 3.4.1 Interviews

If Kvale's (1996) metaphors on interviewing are used, I would see my research as that of a traveller and not a miner. Whilst to some extent, knowledge on testing is 'buried metal waiting to be unearthed' (1996, p.3) as suggested by the miner metaphor, I do not believe the representation that knowledge is a given. I would place this research within the traveller metaphor on which there is a journey with a tale to be told and the knowledge is therefore a constructive understanding. It was hoped that "The journey may not only lead to new knowledge, but the traveller may change as well" (Kvale, 1996, p.4).

Indeed, one of the biggest factors influencing the undertaking of this research is to make the necessary changes within the context of my working life. I hoped to tell the tale and change the school.

Punch (2011, p.144) states that interviews are one of the 'most powerful ways we have of understanding others.' There are ways of accessing people's perceptions, meaning, definitions of situations and constructions of reality.

Semi structured group interviews were undertaken. The basis of using semi structured interviews was that it gives some standardisation of questions but still retains flexibility,

which may support data analysis (Hamilton & Corbett-Whitter, 2013, p.107). It also aligns to the reason of not using focus groups, in that only a general focus on testing impact was known. The research aimed for the participants themselves to highlight areas they felt were significant, within the broad spectrum of the areas or interview prompts and questions.

The semi structured interviews were not recorded electronically for the reasons highlighted within the pilot group. These concerns are supported by Wellington (2000), alongside the areas raised during the pilot, in relation to tape recordings, other factors identified by Wellington (2000, p.86) were also considered. He notes that recordings can preserve language, is objective and allows for the interviewer to maintain eye contact and concentrate. Recording can, however, also produce vast amounts of data to transcribe, and the context may not be recorded.

Each of the issues raised were taken into consideration during the research design and, where possible, addressed. As the participants would also keep a yearlong reflective diary, it was hoped the 'language' could be gained there alongside the interviews. Conclusions drawn would be checked with participants to try and reduce researcher objectivity and eye contact was something consciously attempted while making short notes.

During the design phase, interview strengths and weaknesses identified by Greenfield (2002, p.209) were also considered. He identifies interview strengths as: allowing the identification of body language, permits large amounts of data to be collected quickly and with access to immediate follow up questions. The weaknesses he identifies that were taken into consideration were: interviews are reliant on the cooperation and honesty of participants, they are difficult to replicate and that interviews rely heavily on the researcher to be honest, resourceful and control bias.

As much as possible had been done during the pilot research to hone the skills needed for research and the regular meetings to review researcher findings would aid any researcher



bias. The interviews were not repeated to replicate each other, but to show any shifts in participant thinking.

The idea is really for the interviewees to set the agenda. 'They should set the agenda and determine the direction and topics that emerge, for them to inform you of the issue' (Thomas, 2016, p.189). This would fit within the interpretive paradigm and would feel more like the conversation missing from the pilot group. Kvale's (1996) knowledge as conversation from his 5 key features of his post-modern construction site of knowledge (knowledge as conversation, narrative, linguistic, contextual and interrelation) also supports the tone and setting of the interviews, as my desired conversations. Kvale (1996, p.42) identifies knowledge as 'conversation or a dialogue' regarding a topic of mutual interest, with a move toward 'discourse and negotiation about the meaning of the lived world'. The final research group interviews would aim to be very much a dialogue with a view to understanding the impact on both participants and pupils.

Appendix 2 shows the final agreed list of topics to be covered, instead of questions, and allowed the freedom for the participants to raise their perspectives and not be led by a series of questions. It would also reduce the possibility that there would be a perception of a socially desired response. It also addressed the concerns I had in terms of focus groups and conflicts arising. Having a list of carefully agreed topics, would help steer but not control. While not a focus group technically, the 9 behaviours of focus groups, as identified by Gray (2014, p.481) were considered when planning the topics to discuss and his appropriate responses adhered to.

Given the perceived lack of success in the pilot interviews, which were one to one with a teacher and two support staff. I was, apprehensive about a group interview of two teachers and four support staff. To have some form of strategy, I discretely had the 9 behaviours of focus groups copied in my notebook to refer to. While not really needed, it gave me some

form of security that I would be able to deal with any issues that may arise. Time was valuable and I could not afford to reschedule, or do the interview again. While initially worried about the group interviews, I had forgotten the benefits of the research to the participants and so the group interview I had feared was a very professional and honest conversation. The behaviours I had been preoccupied with did not surface. Perhaps beginners' luck, but I would prefer to believe it is due to a group of like-minded professionals striving to understand the educational system they find themselves working within.

The same participants were interviewed at the start of the school year in September and again in July, once SATs were complete and results shared with pupils and parents. Visiting the same topics through group interviews at the start and end of the research also allowed for the development of the participants thoughts, and this then gave the value of some hindsight. After all, their opinions before really being in class and working with the children, may greatly change by the end of the year and having sat a week of tests. Yin (2009, p.109) suggests interviews should be classed as 'verbal reports only and are subject to problems such as bias, poor recall or inaccurate articulation'. This had also been considered and plans such as: recording, moving from individual interviews to group and location, were adapted following the pilot. Dawson (2009, p.27) refers to in depth interviews as 'life history interviews', as the researcher is attempting to achieve a holistic understanding of the participants point of view. The research itself is based on the perspectives of primary school teachers and support staff and therefore it was hugely important that they felt free to discuss their views and what they deemed to be important. To enable this, I had to be able to create rapport and trust with participants and it is for this reason that I feel basing the research within my own setting, was more a strength than weakness. At all stages of the interview, any conclusions drawn and recorded were verified with participants to reduce any unintentional bias on my part. Issues such as honesty of the participants and the cooperation of a small group, had to be the main area of weakness in need of addressing

and the pilot and reassurances given to staff prior to the research starting were a means to address these. While I appreciate this research is solely not interview based, interviews both formal (the start and end of year interviews) and informal (monthly check in interviews) are threads running throughout it.

### 3.4.2 Reflective Diaries

The semi structured discussions would give the start (taken in September) and end thoughts (taken in July) of the participants, but the thoughts along the way also needed noting and recording. A reflective diary was identified as a way of doing this and would be kept from September to July. Participants were instructed to note 'anything of interest' to them as class teachers and to provide the researcher access to them monthly. This allowed me to type up the notes and create a simple summary of thoughts to check with each participant. It also meant that should the diary be lost or damaged, all the data would be preserved.

It was a deliberate choice to not specify what or when to write to the participants. In a similar stance to asking questions in discussions, asking participants to comment about a particular subject, group or time may influence what the participants saw as important to me as the researcher. I would argue that this is still the most appropriate way of gaining a genuine teacher perspective, but it also created some issues in that the participants responses varied from a day-by-day account to weekly. Some also did individual lessons while others commented when things 'were only worth noting'. Future research would benefit from stipulating daily or weekly at the very least.

To not make the research onerous, the participants were asked if they would like to write a reflective diary or have one electronically. All but one of the seven participants (a teacher) requested a paper diary. Whilst there were very little guidelines on what would go in the reflective diaries, it was explained that pupils should not be identifiable by name and instead

by first initial. Diaries were not named but numbered and were written in exercise books the pupils were used to seeing staff write in. The aim of this was, that it would reduce pupil suspicion and therefore potentially give a more truthful observation of events and feeling. It would also be something participants are used to doing. The diaries mainly worked well but the participant that opted to complete a diary electronically struggled to find the time to get on a computer while working with the pupils and after the first month opted to also keep a paper diary.

Another issue that should be noted as something that should have been anticipated, but was not, was simply time with the participants. The group interviews were scheduled in at the very start and end of the school year, and the timings of these created no real issues. However, checking in with participants monthly in terms of my understanding of their diary was an issue. While it was important that clarification of my interpretation was sought and that it was done frequently, while still active in our thoughts, this did seem to create an added pressure, especially on the time of the participants. Some months were shorter than others, sometimes we had more things to discuss, some months were very busy, staff had after school clubs etc. Therefore, whilst the diaries were initially analysed each month and notes made; it was not always within the same month that I held the discussion with the participants. They were always discussed a month at a time, but often in later months and a few at a time. It was also not always an in-depth conversation situated in a relaxing environment, but sometimes quick chats in classrooms or corridors. This was not ideal as sometimes the comment/s I had planned to discuss were no longer fresh and time meant that perhaps the depth of clarification sought was not always available. However, the positive of this is that it gave a longer time to link ideas by the participants. Initially they may have thought one thing but when looking at my thoughts in several blocks i.e., October and November they were now able to say:

“Oh yes. I remember. I had just thought they were struggling but actually x was annoying them, and they just needed to be moved groups.” (Participant C).

A major challenge for interpretivist research approaches centres on the question of how the researcher can be sure that they are simply not ‘inventing data or misrepresenting the research participants perspectives’ (Mason, 2009, p74). While not as detailed as I had hoped and despite the challenges around ensuring the perspectives of the participants were accurate, every opportunity to ensure the research conclusions were accurate was undertaken.

### 3:5 Analysis Procedure

Reflective diaries were collected each month and I recorded my initial thoughts. Throughout the monthly collection of reflective diaries, Kvale’s (1996, p.189) six step analysis was used:

1. Subjects describe their lived world.

Within this research this was the participants’ reflective diaries. Participants were free to describe their own understanding of their lived world. Things they felt important or worth investigating.

2. Subjects themselves discover new relationships and connections.

This strand developed as the research progressed. As already mentioned, by default and time restraints the participants began making their own links and noting these within the diaries.

3. Interviewer condenses and interprets the meaning of what the interviewee describes.

This aspect was my end of month reflections. It would form the basis of a more detailed theme analysis later. It would also give starting points for future conversations with the participants.

4. Transcribed interview is interpreted by the interviewer.

This formed questions/clarification I then sought from the interviewee. Within this research, this particular aspect came before condensing, as the reflective diaries needed typing up. The information sought was not gained via interviews completely and so the diaries did not allow for interpretation as would have been possible had interviews been used.

5. Re-Interview.

This aspect was not such a full re interview but a quick clarification/checking my interpretations were accurate. These were the quick chats in the corridor or few minutes after school. While I appreciate it was not a full re-interview, it was the opportunity to seek further understanding.

6. Extend the continuum of description and interpretation to include action.

This was the main aim of the research yet it created many issues during the monthly mini analysis. This is mainly due to many thoughts each month, that in a usual classroom scenario you would act upon. Things such as let's try this and change X, Y and Z. I had requested staff as a whole school community to not change anything significant (large tweaks or changes) to create the bigger picture, so having thoughts such as boys appear to be struggling or this particular table of pupils need something, went against your instincts to investigate options and then make tweaks. While no pupils would have been 'left' we had decided not to intervene as swiftly as we would usually. A previous course had suggested that on average school staff only leave pupils for seconds before intervening. Staff were aware of this from our training and this was indeed one of the reasons we moved to the

philosophy of staff allowing pupils to try and fail, before stepping in. This research simply extended this philosophy.

At the end of the school year, I had my monthly thoughts, two lots of interviews and the actual reflective diaries themselves. The final data, following this, was then analysed using the six steps of a thematic approach, as described by Braun and Clarke (2016).

- Familiarisation of the data.
- Assign initial codes to preliminary thoughts.
- Search for patterns or themes within the codes.
- Review the themes.
- Define and name themes.
- Produce a report.

The data had been briefly analysed at the end of each month, to aid ensuring the correct interpretation of the reflective diaries with staff. The order in which I approached this task, therefore, changed a little. I had preliminary thoughts/ initial codes each month and noted these at the bottom of each grid, for each month and highlight an example of these in appendix 3. At the end of the school year, I then familiarised myself with the grid overviews, to refresh myself on thoughts I had had at each stage of the analysis.

These 'thoughts' then became possible thematic analysis within the reflective diaries and this was done by colour coding phrases and examples given by participants (Appendix 3).

Initially there were many possible themes: boys, girls, staff attitude, higher ability, lower ability, SEN, parent concerns, student concerns, staff pressure, pupil pressure, pupil confidence, reading, testing, resilience, new curriculum topics, perceived pressures, teaching strategies. This had to be narrowed down and reviewed, if any conclusions were to be drawn, never mind trying to understand why something had or had not happened. Where possible overarching themes covered smaller subsets until possible umbrella themes come to fruition. These became attitudes and behaviour, the learning environment and labelling. Each of these themes is intricately and delicately linked. The behaviour and

attitudes of pupils identified within the reflective diaries, is a product of the setting's learning environment. In turn, the learning environment is created by how the setting use labels. The sub sections of each overarching theme, will be discussed in some depth later.

This now appeared as a blanket one size fits all from a very painstaking and through initial analysis. A broader area (umbrella), therefore, had to be identified to understand the findings and stop becoming lost; unable to see the wood for the trees, which I had experienced for some time. I had identified too many areas to investigate to make it productive. Yet this is something I grossly overlooked during the vital stage of analysis. The research itself had aimed to not change any systems in school during the research period and I had moved from the most natural system and form of analysis I knew and had made things far too complicated. I had wrongly believed that I had to analyse and understand every comment and theme individually and that somehow, in doing this it would knit itself together and create the bigger picture. I had approached the analysis not as a teacher and instead looked at it from the bottom up, instead of the bigger picture and working down. In hindsight, I would attribute this to the pressure of wanting to have answers and 'fix things' in order make improvements to benefit students and the participants.

However, despite this, the research design's strengths had allowed for such vast collection of data initially and the reflective diaries recorded over a year permitted time to analyse this. The monthly participant reviews had reduced researcher bias and the interviews being repeated at the start and end of the research year had allowed any changes in views to be noted. The design of the research had been sufficient, it was a novice researcher's naivety that delayed the analysis and in turn made conclusions difficult initially.

Once I had I the overarching themes, I could then work backwards and create subsets within each theme, based on the initial colour coding.



Attitudes and behaviour would cover: confidence, resilience, concerns and work ethic. How the pupils expressed their thoughts and feelings and why they were expressing them.

The learning environment would cover: how the school had created a climate for learning and how each class had addressed this. I had already given my word to the participants that in participating, their teaching was not under investigation, so I had to be very careful that this did not come across in the research as criticism of their teaching or individual classroom systems. The monthly check in or thoughts should not focus on teaching styles or how individual participants had organised classrooms for example - the focus clearly had to be the pupils and their reactions. It had to focus on what the school generally was/was not doing if it was to make long term changes and impact. Linking my first theme of how pupils were expressing their thoughts and feelings, then lead naturally into why something is being expressed a particular way. I also, by now, had the luxury of having the actual SAT results. Given there were not any significant differences in results between either class, it seemed counterproductive to investigate the teaching styles against results and attitudes of pupils, for the two teachers regarding classes that had now left, with a cohort that they will never completely replicate again. Staff may remain in that particular year group going forward, but they will not have the same pupils again and so looking at school or class procedures that we were going to be continuing with, regardless of who the pupils were would give us the way forward. The research aimed to look at how I can improve things for all pupils as we move through the new assessment changes, therefore the school environment and the ethos and tone it sets would cover the areas such as pressure and testing. It would also help to understand how the school was linked to the development of attitudes and labelling.

The final theme of labelling was the last one I arrived at and came across it more by accident than analysis. I had colour codes of ability groups, pressures, concerns and so on and I began to wonder who had identified something as a concern or pressure in the first place

and what was the basis for this? Who had labelled it as such? This would then bring together the two other umbrella themes. Were the attitudes generated by labels or the school environment? Was the learning environment a product of the attitudes and behaviours and the labels we had attached? I had initially begun this research to look at if a pass or fail label would impact on our pupils and yet had not thought to have this very notion of labelling as an area within my analysis.

### 3.6 Data Analysis

This section explains the process from research design, carrying out the research to finally analysing the data. The literature review did not yield any links to the research area of focus, specifically in evaluating the impact of the new testing regime, as opposed to testing and impact in general. The results of the literature review are to inform the research conclusions and allow areas identified in the initial literature review to be revisited once data had been collected and analysed, in order to understand the impact. There are many pieces of research that have investigated testing, impact, labelling etc but little, if any, set within the context of the new national curriculum and the new pass or fail regime. I do acknowledge their existence but feel that the testing research available, in general terms, was too broad a spectrum to correlate

As the research is a qualitative piece of research, the eight areas identified by Miles and Huberman, (1994, p.6/7) are included within the design.

- 1- It is conducted through a yearlong contact in a life situation, reflective of the everyday life of the participants.
- 2- I, as researcher, aim to gain a holistic view on the context being studied.
- 3- Data will be captured from the inside through a process of empathy, understanding and suspending any preconceptions.
- 4- When reading through data and the reflective diaries, themes were reviewed with participants.
- 5- A main task/aim is to explicate the way in which the participants manage and view their day-to-day experiences.
- 6- Understanding that many interpretations of the data may be found.
- 7- There was little if any standardised instrumentation at the onset.
- 8- The analysis will be completed in words.

Each step will be discussed in depth. Figure 1 shows the flow chart of the research design.



Figure 1. Flow chart of research design.

### 3:7 Conceptual Framework for Analysis

Oliver (2014, p.29) highlights that a problem with terms such as theoretical framework, perspective and paradigm is that they are used by 'different writers in different ways'. Within this research, the theoretical framework will be defined as the specific assumptions made during conducting the research.

The theoretical framework underpinning the research is that of Bronfenbrenner's (1979) ecological model. Education, whether we agree with them or not, is set within a framework of policies and legislation. Bronfenbrenner (1979) was a 'visionary' able to grasp the

essence of major theoretical, as well as policy issues 'long before most had come to understand them' (Wertsch, 2015, p.144).

Bronfenbrenner's (1979) model focusses very clearly upon context and the personal development achieved within each context. Bronfenbrenner (1979) refers to these contexts as an ecological system.

This ecological system consists of five subsystems. The innermost level (microsystem) is the immediate setting of the developing subject; places such as home, school or a laboratory in the terms of experiments.

'A pattern of activities, roles, and inter-personal relations experienced by the developing person in a given setting with particular physical and material features and containing other persons with distinctive characteristics of temperament, personality and systems of belief' (Bronfenbrenner, 1979. p.148).

The next subsystem (mesosystem) looks beyond one setting and looks at the interconnective relationship between the two subsystems, home and school for example.

'the linkages and processes taking place between two or more settings containing the developing person. . . or a system of microsystems' (Bronfenbrenner, 1979, p.148).

As one moves to the third subsystem (exosystem), it presents a hypothesis that development is affected by events occurring in settings in which the subject is not present and where they are not an active participant, for example the neighbourhood, mass media, industry, and policies.

'One or more settings that do not involve the developing person as a participant, but in which events occur that affect, or are affected by, what happens in the setting containing the developing person' (Bronfenbrenner, 1979, p.25).

The fourth subsystem (macrosystem) is the largest of all subsystems and relates to the consistencies of the three previous subsystems that could exist in the culture, alongside their belief systems. It also encompasses political and economic systems.

‘Macrosystem are consistencies that exist, or could exist, at the level of the subculture or the culture as a whole, along with any belief systems or ideology underlying that system’(Bronfenbrenner 1979, p. 26).

Later, Bronfenbrenner went on to reference this ecological system further as

‘opportunity structures, life course options as part of a ‘societal blueprint for a particular culture’ (Bronfenbrenner, 2005, p. 150).

The final subsystem (chronosystem) is the dimension of time and includes the influence of change, good or bad, and consistency. Figure 2 highlights Bronfenbrenner’s Ecological systems and how the systems fit within each other.

‘The influence on the person's developmental changes (and continuities) overtime in the environments in which the person is living’ (Bronfenbrenner, 1986, p. 724).



Figure 2. Bronfenbrenner’s Ecological System (MacBlaine 2018. p42)

To thoroughly understand one dimension, Bronfenbrenner (1979) states that the whole ecological system should be considered. Human development reflects, not only potential and individual personalities but also the interplay between family members, neighbourhoods, communities and the influence of factors such as social policy, economics, and the legal system (Tregaskis, 1995). This was referred to in previous chapters, in relation to only copying one ecological system (education) from other countries.

Like a set of Russian dolls, the context of human development work in a nested fashion, each one expanding beyond but containing the smaller ones. Each one also simultaneously influences and is influenced by the others. (Bronfenbrenner, 1979, p.3)

While each sub system influences on the next, some have more influence and stronger boundaries than others, I would suggest not all sub systems have genuine 'two-way traffic' or as Bronfenbrenner (1979) states 'bi-directional influences.' In references to the bridge metaphor already used, traffic can be stopped, slowed or increased by the bridge traffic controller. Otherwise known as the headteacher.

The ecology of human development is the scientific study of the progressive, mutual accommodation between an active, growing human being and the changing properties of the immediate settings in which, the developing person lives, as this process is affected by relations between these settings, and the larger contexts in which they are embedded. (Bronfenbrenner, 1979, p.21)

It is for this reason this theoretical framework has been chosen. What is impacting more on the child- school or the testing legislations; the exosystem, or macrosystem? These systems will be referred to within this research and used within the data analysis and conclusion sections.

Figure 3 highlights the links between the ecological system to the general education sector, Crooke (2011).

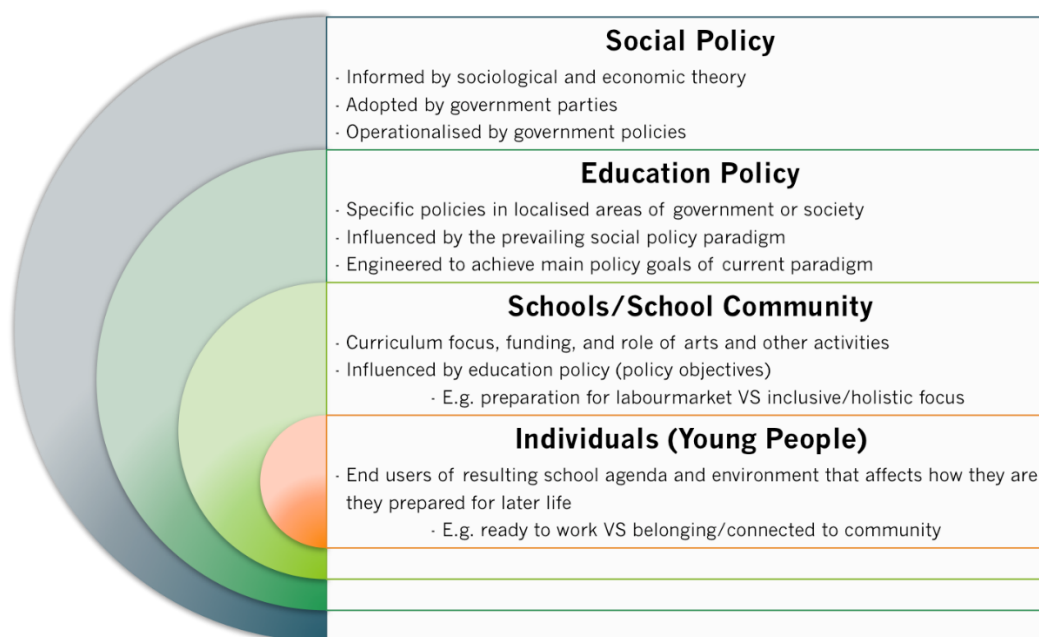


Figure 3. Education System Application of Bronfenbrenner's Ecological Systems. (Crooke 2011. P 13)

### 3.7.1 Application of the Theoretical Framework

Within the context of this research, a brief overview of each subsystem would be:

Microsystem	The close relationships between pupil, peers, and teacher
Mesosystem	The school stakeholders such as teacher, parents and headteacher; the school.
Exosystem	Outside agencies such as LA advisors, Ofsted
Macrosystem	The education system such as DfE including educational policy.
Chronosystem	Policy impact over time

Within this research the microsystem is that of the year 6 classroom. It contains the relationship between pupils and their relationship with each other and their classroom staff.

It is where most of the time is spent and is different from general school interactions. The tight focus on testing is specific to one aspect of school life and one main set of interactions. This set of interactions and in turn allocation of position, is not replicated outside of these classrooms. The environment of these classrooms, the personalities within them and the resulting behaviours due to this will be the main basis of the research.

The mesosystem is, simply speaking, a system of several microsystems. For the purpose of this research, I would define this as the relationship between class teacher and pupils within the year 6 setting and those of the class teacher and their place of employment. It would therefore be apt to acknowledge that the positionality of teachers may change as they move from the microsystem to the mesosystem. How the teachers interact with pupils in comparison to the headteacher or governors. – what they want to do and what they must do. This is reflected in the values and aims of the school and therefore the wider school as a whole would form part of this subsystem. While parents are a significant factor within Bronfenbrenner's (1979) subsystem definitions, parents directly have not been studied as part of this research. It is acknowledged that they will have a significant impact on the child and may influence their child's attitude toward testing, but because of time and organisation pressures, their views were not included. Instead, the teacher and other support/cover staff were the focus as they are with the child at the point of test, preparation for and end of testing. Teachers also know families well enough to understand the family dynamics and will, by the very nature of their job, interact with them. The home-school learning is not an essential aspect of this research as, in brutal honesty, very little control over this can be given. What this research is concerned with is how the children are affected by the new testing regime and what I, as headteacher, can do about this. As the researcher you can speak to parents, but they do not need or have to agree with whatever findings the researcher may conclude. Research cannot control the whole aspect of this subsystem and the main aim of this research is to inform practice in this particular setting and within this



subsystem, with a view to it improving the two further subsystems nested within it – pupils and staff.

The exosystem within this research links to boundaries in which the school is set. By this I mean the Ofsted report it had, the data it had to improve and the advice it had been given by the local authority. School systems and policies were implemented and created to address these and some, such as Growth Mindset, have been unpicked somewhat already previously. The pupils do have some say on school systems, but these are usually only after something has been implemented and they are asked for an opinion. The balance between done too the child and done with the child is very fine. Pupils certainly do have far more voice today than they ever had historically but the aim of this research is to look at what the school is doing to the children (within this research this is testing) and in that, unfortunately they have very little choice. The school cannot change the whole ecological system, but the aim is to influence the correct subsystem to in turn improve the others.

The actual data collection element of this research focussed very clearly upon the first two subsystems of the micro and mesosystem, but the conclusions and ways forward will very clearly be within this subsystem – the exosystem. The findings will impact future school policy and systems, which will either be in support or not support current external advice. This in turn will impact and affect pupils without their knowledge or input. Perhaps this particular subsystem should be the one that yields the most bi-directional influence Bronfenbrenner (1979) refers too. This is the subsystem that permeates a very top-down approach, as depicted in the policy reviews earlier.

The Macrosystem of any educational research will always be the educational system as a whole and therefore national educational policy. It is this subsystem that has been addressed in some depth within the literature review and the impact on this subsystem regarding the previous three, that will be discussed within the findings. Has the policy of

testing impacted on pupils? If so, has it permeated all other subsystems or just some? If this policy has not impacted on pupils, then which subsystem had the weaker barriers and why? Whilst not a criticism of Bronfenbrenner's (1979) ecological system, it is at times unclear if subsystems can bypass each other. It is clear from his writings that everything is interrelated and interacts to various degrees, but can any subsystems in particular circumstances have had no interaction or impact from others? This research has the potential to look at each subsystem, in relation to testing, and identify the role (if any) each played and answer this question.

The Chronosystem has to some extent been addressed within the history of assessment section and has identified the journey to the current assessment system. The impact of previous testing systems is known, but the impact of this new system is still not clear, and research is in its infancy. However, this research, while aiming to only change my own school and its subsystem, has the potential to impact on this wider subsystem and therefore the chronosystem.

Within this research, the context in question is that of the school and the new testing regime and the impact of this on the pupils. Within the context of this research one clear system (school) will be analysed but the whole ecological system in which schools' function will be acknowledged.

In ecological research, the properties of the person and the environment, the structure of the environment settings, and the processes taking place within and between them must be viewed as interdependent and analysed in system terms. (Bronfenbrenner, 1979, p.41).

If I link Bronfenbrenner's (1979) ecological systems to case studies and understanding a particular case, I can see the parallels between Stake's (1995) explanation of a case and the eco systems described. Stake (1995, p17) identifies those issues are not clean and simple, but intricately wired to social, historical, political and personal contexts. These meanings are

important when studying cases and draw the researcher to tease out the problem, the conflicting outpouring, and complex backgrounds. This mirrors the ecological systems and their impact on each other, the influence and interpretation each system will create for its members. My research explores the case of a new testing regime and investigate the complexity of its impact within the various contexts and ecological systems.

While Bronfenbrenner's (1979) theory has been criticised based on not sufficiently paying enough attention to the individual development of children, it nevertheless offers a 'useful means of thinking of the wider and often unseen impact of society on development and learning' (MacBlain, 2018, p.4).

Given the age of Bronfenbrenner's (1979) ecological system, there is no doubt that the world in which we live today is not fully covered within the 5 original systems. Drakenberg (2004, cited in Christianson, 2016) identifies a new subsystem, ex-macro. This is an international system reflecting the globalisation we take for granted today. This too is mirrored in the education system and world league tables and should be reflected within the ecological systems.

A further area to raise regarding the ecological system and the things important within the school context of this research, is resilience. Growth Mindset has already been discussed within this research and I would agree with Engler (2007) and Christensen (2016) that this is missing from the ecological system model. The notion of resilience and the ability to overcome obstacles is prevalent in many educational policies today and certainly a priority within my own school. While this theoretical framework is the best suited for my own epistemological stance and that of this research, I acknowledge Engler's (2007) stance that including resilience could help explain some of the unexplainable ways in which people have overcome trauma and tragedy.

However, given this research aims to explore teachers' perceptions of the impact of a testing regime, with participants being both product and producer, this model allows for the various ecological systems to be analysed and potentially address some of the weaknesses identified. It will also allow the role of headteacher to then be analysed.

What sets Bronfenbrenner apart is his general assumption that one cannot improve the developmental trajectory of individuals by focussing primarily on individuals. Instead, one must go 'beyond the human organism, not only to understand, but to change the development trajectory of that organism. And the key to that is changing the social environment in which individuals are nurtured' (Wertsch, 2015, p.147)

This research will look at the interplay between the ecological systems and as Bronfenbrenner himself states, looks beyond just the pupils as individuals, but the role in which I have played, as headteacher. It allows for the findings to be further analysed through the learning environment, behaviour and school's use of labels. For the ecological systems to be seen as just that, a whole complex system and not a singular being or ecological system.

Coe (2012, p.9) identifies the aims of a piece of research as a 'separate dimension from its values, assumptions and methodology'. Using the typology of aims identified by Coe (2012), this research is predominately overlapping with scientific and political aims. The reason being that the main aim is to express a lived experience, while hopefully creating a new piece of knowledge and therefore impacting on educational policy, even if only at grass roots level.

Given the research aims to describe and understand the impact of formal assessment in primary school and the role of the headteacher within this, it will have recommendations and inform my own practice and therefore will be an applied piece of empirical educational research. Bassey (2008, p.39) states educational research is a critical enquiry aimed at

informing educational judgements and decisions in order to improve educational action. In this case, the actions the school takes following the conclusion of the research.

Elliot (2007, p.149) argues a strong case for distinguishing between educational research and research on education. He identifies that educational research is a form of 'common sense enquiry rather than a science', aimed at the formation of practical insights and judgements.

### 3:8 Positionality and Ethics

My own epistemological stance is that of an interpretative perspective, in that I acknowledge a variety of factors that individuals will apply when relating to the social world. An interpretivist approach not only 'seeks people as a primary data source but seeks the perceptions of an insider view' (Mason, 2009, p.56). Within interpretative research there is an appreciation that participants, the researcher, and research field are involved in an 'interaction and dynamic process, with each impacting on the other' (Sage, 2014, p.187). Interpretivist studies tend to be small scale and aim for detail and understanding, rather than statistical representativeness. It seeks to understand and portray the 'participants perceptions and understandings' of a particular issue or event (Burton and Bartlett, 2005, p.22).

A constructivism stance alongside this, would be fair to state and understood as:

Knowledge lies in the minds of the individuals, who construct what they know on the basis of their own experiences. It suggests that the process of knowledge construction is active rather than passive. Researchers who adopt this approach believe that research involves an attempt to understand individual construction of knowledge and believe their role to understand the ways in which individuals construct meaning, since knowledge, truth and reality are created rather than constructed. (Savin-Baden, 2013, p.29).

My epistemological stance mirrors that of the theoretical framework used.

“To a greater extent than any other species, human beings create the environments that shape the course of human development “(Bronfenbrenner cited in Wertsch, 2015, p.147). Bronfenbrenner (1979) views the effects of social environment to be the main issue and in this regard has ‘constructed a theoretical framework that stands in opposition to methodological individualism and atomism’ (Wertsch, 2015, p.147).

The data collected, will be the perspective and ‘truth’ of the research participants. It makes an original contribution to knowledge by giving a voice to teachers involved in the testing systems within a primary school and reflects experiences of year six pupils.

The willingness of staff and support of governors highlights just how vital this ‘voice’ is. With this, it is important to note that research relationships are two sided and that those being researched will make their own interpretations of what is going on, regardless of the researcher’s intentions (Opie, 2008).

Kvale (1996, p.296) highlights that we exist in a circle of conversation, with our understanding of the world dependent on conversation and our understanding of conversation dependent on our knowledge of the world. He believes it is not a vicious circle but instead in a hermeneutical sense, a circulous fructuosi, going on to raise that the issue is not getting out of the conversational circle, but to get into it in the right way. This research aims to use the conversations as a valid leaning tool for improvement within my own setting by using conversation ‘in the right way’.

The style in which the research has been conducted is that of a pragmatic approach, in that my starting point was a problem in need of a solution. Punch (2011, p.19) defines a pragmatic approach as beginning with a ‘research question that needs answering and then choosing the methods to answer them’. Pragmatic research should also be carried out in a ‘natural context’ (Savin-Baden, 2013, p.24). The research will dip into the realms of substantive theory, due to being a content-based theory with the aim of explaining and

describing the area of research. If the research is to be used to inform future plans and adaptations, then it is important that I not only describe a teacher's perspective but have some understanding of why something is happening. Explanation goes further than description. Description focuses on what is the case, while explanation focuses on why something is the case and thus 'explanatory knowledge is more powerful than descriptive knowledge' (Punch, 2011, p.22).

The research is based within the school in which I am employed, with the school's staff and therefore the element of staff saying what they feel I may want to hear has to be raised. The power I held or was perceived to hold, had to be confronted at the very beginning of the design process.

At the very start of the planning stage, staff were reassured that the research was not a reflection of their teaching, and it would not involve any observations by myself. Instead, the research aimed to look at the impacting of testing from their observations and thoughts. The research sought their professional opinions of the events in their class over a year, with an aim of improving things for the pupils and therefore in turn staff. At the end of the research year, we would then, as a collective look at the research findings to look at the ways forward based on the findings. The research was a tool for school development and would inform future changes within our own setting. One advantage of own classroom research is the insider knowledge and understanding. This type of research can bring an understanding not only of the research question but the social, cultural, and micro political aspects. This can 'enrich and deepen' the research (Punch, 2011, p.44). As already highlighted, as the research evolved it became clear that whilst the perspectives of staff regarding testing impact had been the main reason for undertaking the research, my role as part of this impact was also no part of the research. This therefore meant a genuine whole school analysis of our ecological system.

The ethical considerations of obtaining data from children unable to give informed consent was prohibited in this context. However, despite this there are benefits to the research using the viewpoints of staff. The class teacher firstly has the vocabulary to describe their observations and feelings, a rich descriptive narrative may not be given by pupils aged 10 and 11 who may not know how they feel. The teachers also know their pupils exceptionally well and would notice when a behaviour changed, or a pupil behaved out of character. The shared culture staff had would be reflected within the research. If the research involves children, the balance of power can be pronounced. In asking pupils to participate in a questionnaire or interview there is the danger they will not perceive me as a researcher but as the teacher or headteacher. The answers given may therefore not be honest but instead be those deemed by each pupil to be right (Burton, Brundrett, Jones, 2008). This aspect has already been raised in relation to staff undertaking research within their own setting and any research undertaken within the researcher's institution will raise the same issues. Therefore, I would suggest that while staff may feel like pupils in regard to being honest, they are better placed than pupils to see the long-term gains of participation and, therefore, I feel the decision to base the research on the opinions of teachers rather than pupils is the correct one.

Much of what we cannot observe for ourselves has been or is being observed by others. Two principle uses of case studies are to obtain the descriptions and interpretations of others. The case will not be seen the same by everyone. Qualitative researchers take pride in discovering and portraying the multiple views of the case. (Stake, 1995, p.64).

In using a variety of school staff, though all do teach or cover lessons, several viewpoints would be obtained and each of these viewpoints could be articulated.

Qualitative advocates such Guba & Lincoln (1982), Eisner & Peshkin (1990) place a higher priority on direct interpretation of events than those interpretations from measurable data. Erickson (1986 cited in Stake, 1995, p.42) claimed the primary characteristic of qualitative



research is the 'centrality of interpretation and as such, findings are assertions.' This research would be based on my interpretation of teachers' assertions. It can be argued that a great deal of empirical knowledge is brought to case studies from outside the evidence of the case itself and informs both the description and analysis of the case (Evers and Wu, 2007). This would need to be considered during the analysis of the data given by the participants and factored into the discussions. The research had to be an accurate interpretation by myself of the interpretations of the participants of their pupils.

Ethical considerations were approved by the university and adhered to throughout the research, both at university and school level. This involved informed consent, the right to withdraw and the protection of identify and data. This consent was also sought from the chair of governors to undertake the research within the school itself. All participants were required to give written consent, which is included within Appendix 4, and informed of their right to withdraw, that all information would be securely stored on an encrypted USB and shared only with university supervisors. It was also agreed how the reflective diaries completed by staff would be stored securely within their own classrooms and my own office. This was agreed to be my locked filing cabinet and the classroom locked store cupboard. Records would be destroyed following completion of the research. Throughout the research, participants were given my summary of monthly findings based on their notes, to ensure correct interpretation. An example of this is given in Appendix 5. As discussed, to reduce the work life balance I did not expect staff to discuss my summary of thoughts unless they felt I had misunderstood their diary. That said, all participants had the right to challenge and clarify comments. This system was also followed with the notes from the participant group discussion at the start and the end of the year. Throughout the research all participants were treated with respect and care.

As not only a researcher, but as the headteacher, it was made clear at the outset, that I had a professional duty to report any findings in which I believed the participants were in danger to themselves and others.

In order to protect the identify of participants, every effort has been made to not give away details that would point to their identity. Gibbs (2012) suggests that group interviews can present additional dilemmas and risks, as it could be argued that confidentiality may be difficult to ensure because all participants hear the discussion. As the research is also based on formal assessments in primary school and these only occur in two school years, there is always the risk that the participants (both staff and the pupils observed) would be easily identifiable. It is, therefore, just as vital that the school's anonymity is placed in high regard also. However, many pieces of qualitative research occur in natural settings, to understand the context of the phenomena they are aiming to interpret (Savin-Baden and Howell-Major, 2013). While conducting the research within this setting does create some element of risk of identification, it is the only way to truly understand the impact of testing within this particular setting.

The research, as already discussed, was to inform school changes and identify next steps. While the research would benefit the school, it was my own personal choice to make the research part of an EdD. This meant that initially, I personally funded the university fees. However, due to both work and personal circumstances, then a global pandemic; this took longer than planned and school therefore financed the research also. I also needed dedicated research time, which meant the school were permitting me time away from my paid day to day job. This created a perceived expectation on my account, that I would have answers to the problems I had raised (pupils responding to a pass or fail, staff morals etc). That not changing anything for a year would reap a significant reward and give me the knowledge to improve. If it did not, I risked being a year behind with making any future

changes. The pressures on me as researcher, to understand the problem, solve it and not jeopardise the gains we had made had also to be factored into any analysis. The data had to be true not what I or staff felt people would want to hear.

Miles and Huberman (1994, p290-297) raise eleven ethical issues for qualitative research. While many of these have already been addressed indirectly, it is worth noting clearly how all considerations have been addressed.

1. Worthiness of the project: The research was not undertaken to further a career, paid research or publication. It was undertaken simply to enable a better understanding of a new testing system and how I as headteacher, and the staff, could prepare for the impact.
2. Competence boundaries: This was perhaps the hardest issue as a novice researcher. While I had enthusiasm and a passion to undertake the research, the skills were not naturally there, and I found myself with more lows than high. I did not feel 'academic' or professional and therefore doubted my own methods and spent many hours reading literature on how to conduct research, alongside speaking to my supervisors. I also had very honest participants, who expressed their views on my methods and findings along the way.
3. Informed consent: All participants had a full understanding of what the research would involve but more importantly, they were aware of why the research was being undertaken.
4. Benefits, cost and reciprocity: As already discussed the participants were aware of why the research was important and therefore the benefits to the school in the long term. If we could understand the impact of testing, we could address it. Participants did not have to give vast amounts of time and when they did, these were arranged to suit them. Where possible any time given for the research was arranged to have

least impact on the class and thus in turn create less workload for the participants.

Diaries were one way of reducing face-to-face time and any face-to-face time, was either time limited to a maximum time slot, or during sessions that were taught by external agencies i.e., PE with a PE coach – freeing up the teacher in ‘school time’.

5. Harm and risk: In not changing the school systems in place during the research period, there was a risk that results may dip, and we would be reacting too late to a new system. This was not a risk taken lightly and had the research genuinely not aimed to understand in order to improve, then the risk would have been far greater as not only would the school be ‘on the back foot’ but I would not know why.
6. Honesty and trust: As already discussed, the honesty and trust of both me as researcher and that of the participants, at times made the research almost too hard for a first-time researcher and it would be fair to say at times were taken personally. However, while these issues were difficult to hear, it greatly reflected the trust between the researcher and participants.
7. Privacy, confidentiality, and anonymity: While all steps were taken to ensure these were in place. The bottom line is that these cannot be guaranteed. The research took place over a set time and so staffing could be identifiable to anyone willing to investigate. I am also a headteacher and while the school has not and will not be mentioned, again my career timeline could be easily identified and thus in turn the school. Once the research is concluded, the findings and data will be out there to view. What I have done, is attempted to ensure the participants are not readily identifiable and in using a range of staff and clear system for changing pupils, made this as secure as possible. However, any sensitive information that made it particularly clear who a pupil or participant was, has not been included in direct quotes.

8. Intervention and advocacy: I had made it clear at the start of the research that action would be taken if needed, should any harmful or wrongful behaviour be observed or noted. While this may be one on Punch's (2011) ethical considerations, it is a daily part of school life.
8. Research integrity and quality: The research had reasonable standards set by both me as researcher and that of the participants. Given this is a novice piece of research, those standards may not be as high as other pieces of research but there were standards in place. The research was also conducted carefully and thoughtfully, taking time to allow the participants to express their views, understand these views yet not overburden them with my interpretations.
9. Ownership of data and conclusions: I am very clear that the finding of this research are my own and do not reflect the school. The data belongs to myself and while partly funded by the school, the research belongs to me. Staff and governors will be aware of the findings at the point of analysis, but do not control them.
10. Use and misuse of results: The overarching aim of this research is to use the finding to improve things within my own school. There is the risk that any individual can use the findings for their own gains, as any research can. Most things are open to manipulation and interpretation. While the results will be made as clear as possible, qualitative research is rarely black and white and clear cut.

### 3:9 Reliability, Validity and Generalisations

The research itself is idiographic, in that it is focussing on the sole case of my professional context. It is not designed to be representative of schools in general. Thomas (2016) argues that whilst case studies do not allow for generalisations, generalisations are not always wanted. The researcher may not want or need to generalise, with some of the most insightful research being from case studies. I am interested in it, not because by studying I

will learn about other cases or a general problem but because I want to learn about this particular case, in my own setting, (Stake, 1995). Stake (1995) further develops this, to conclude that experience is essential and whilst this research is a single case, the vast experience as a leader meets the principal qualification of qualitative research. I would argue this outweighs concerns over using a single case. Yin (2009) acknowledged the concerns that case studies are seen to provide very little basis for scientific generalizations. Yet also raised the same point for a single experiment. Whilst single cases are not as strong for basing generalisations on. Things can be learnt from a single case because they can be familiar with other cases, thus making a new group from which to generalise (Stake, 1995). Many qualitative researchers do not aim to generalise their work to the whole research population. Instead, they seek to explain and describe what is happening within a smaller group of people with the belief that this may provide an insight to a wider research population (Dawson, 2009). Furthermore, Hamilton and Corbett-Whittier (2013, p.6) refer to Merriam's revised definition of a case study to be the 'focus on the case, rather than the outcome to understand that the case is a bounded unit'.

Using a single case study would also increase the vulnerability, in that the case may not turn out as I had hoped (Yin, 2009). This was indeed a risk but having conducted a pilot study and ensuring my initial analysis would be checked with staff for accuracy, I felt the risk outweighed the gains.

Stake (1994) identifies 3 forms of generalisation (naturalistic, statistical, and analytical) and this research addresses both naturalistic and analytical. As it is based on personal experience, derived from tacit knowledge leading to expectations and therefore moving from tacit knowledge to explicit personal knowledge, it can be classified as naturalistic. It also covers Stake's analytical term, as reasoned judgements about the findings of this study could be used as a guide to what may occur in other studies. While the school context of

the research may never be fully replicated, other schools will have similar systems in place, use similar strategies etc.

The biggest challenge Yin (2009) identifies within case study research, is that the investigator can make subjective judgements when collecting the data. This has to be acknowledged and as such the measures already discussed, such as checking diary interpretations, were an attempt to be less subjective.

Reliability in using a single case study and within my own setting are potentially the weaker aspects of this methodology. The format is clear in that a group interview scribed would be the initial and final task, with a reflective diary kept by all participants and analysed monthly running between the two group interviews. Whilst this would appear relatively straight forward, another researcher using my design may not yield the same conclusions.

Yin (2009) identifies two objectives to an interview: To follow the line of enquiry and ask the questions in an unbiased manner. Having a list of questions had already proven unsuccessful in the pilot study, so semi structured interviews had to be used with the final research project. I acknowledge that had other lines of enquiry during the interview been identified, then different answers may have been given.

That said, the analysis was always checked and therefore this has little room for any other interpretation. The main findings are based within the reflective diaries and the interviews purely used to gauge the teacher stance at the start and end of the research.

The word validity has been associated and used within qualitative research for many years, yet there is generally no agreed definition. Within qualitative research there are currently 17 different terms associated with it (Gray, 2014, p.182). Thomas (2016, p.73) also highlights that case study validation is no longer accredited from reference to vast bodies of work or generalised knowledge but instead from the 'connections and insights it offers between the experiences of others and your own'. This research would represent the experience of this

school, the staff and pupils in comparison to those portrayed in other schools and as such is therefore valid. Thomas (2016, p.66) also identifies that validity is not solely concerned with 'plausibility and credibility'.

While validity in terms of the research has been addressed, Kvale (1996, p.218) also discusses validity in terms of the 'contexts of the subjects' statements.' Again, I would suggest that the participants/staff had a vested interest in understanding the new testing regime. They would want to reflect it accurately. What was of concern was their understanding, and my interpretation of events, which has been discussed earlier. Kvale (1996, p.241) also raises validity beyond methods and looks towards the 'moral integrity of the researcher'. This is of relevance within this research as put simply; it is the school in which I work, and systems I have created or permitted. Would I really admit I had got things wrong? Absolutely, is the only answer. It is the very ethos I had been determined to create – one where nobody is perfect, where pupils and staff are allowed to get things wrong and indeed making mistakes is simply a part of life. As, will be discussed within the findings, hindsight also allows me to think that in even undertaking the research I felt I had got things wrong.

Qualitative researchers must seriously question the validity of their work, if fellow researchers reading their field notes feel the situation depicted is not supported. However, they would not expect other researchers in a similar or even the same situation to replicate their findings in the same sense of conceptualization (Ward- Schofield, 2000). To counteract some aspects of this argument, Kvale's (1996) three aspects of validity were adhered to. At all stages the findings were explicitly shared with the participants to reduce my bias during the analysis (aspects one). I also answered the questions of what I wanted to know and why I wanted to know it before I decided on how, in terms of the research design (aspect 2).



Finally, a review of methods to ensure that what is intended to be investigated is, involves a theoretical conception (aspect 3).

Gray (2014, p.151/153) distinguishes further forms of validity; face, internal, external, criterion, construct, content, predictive and statistical. Again, I focus on each aspect in turn and relate each to the research design.

Face validity discusses the importance that the instrument at least appears to measure what it is designed to measure. While, the interpretation may be subjective, the research does measure and report on the teacher's perspective of testing. Internal validity refers to the cause and effect and the extent to which conclusions can be drawn. Within this research conclusions are shared with participants along the way and interpretations as to impact will and have been made. External validity has already been addressed. Criterion validity compares how participants have answered in relation to a new measure, comparable with existing and accepted measures of the context. As this research is investigating a new concept, it is difficult to compare to another piece of research. Comparisons to themes within the data and findings/conclusions can be drawn and where possible this has occurred.

Construct validity looks at the measurement of abstract concepts such as ability, attitude, knowledge etc. It is concerned with if the indicators show the expected relationships among the concept being researched. This is addressed during the findings and conclusions, and the links made to other research areas. Predictive validity is concerned with how well a test can predict a future trait. The design of this research would allow for this, it acknowledges the strengths and weaknesses and would allow for the design to be replicated.

A further area within this research to address is that of statistical validity. Whilst this is not required in all research. I acknowledge that very little statistical data has been included and this has been a deliberate choice. This research is interested in a lived experience and the

impact of that experience. It requires rich descriptions and conversations and as such has little scope for statistics.

### 3:10 Conclusion

Ball (2006) notes that vast pieces of research about education are not about policies. He argues that research is seen as about: the classroom, teachers or progress with policies that underpin this as free standing, identifying the requirements of testing, National Curriculum, league tables, teacher standards are irrelevant. Ball (2006) reports the problem is identified as the teacher, the classroom or the school but rarely the policy. Whilst I acknowledge that this research will not single handily change government policy, it will investigate the impact of it within a particular context and be used to inform school policy and decisions.

Bassey (2008, p.23) raises the point that although case studies have made considerable contributions to the corpus of knowledge and practical wisdom of education. They are still regarded with 'suspicion and at times, hostility'

However, I agree with Travers (2001) that a good piece of research, even if non-academic, is difficult to do without being aware how the same topic could be approached by someone with a different conception of epistemology, theory, and method. It is also acknowledged that while qualitative research interviews can potentially yield interesting results, they have also been dismissed as not being scientific and as such critical objections will remain endemic to the practice, Kvale (1996, p.59)

## 4 Chapter Four: Results and Analysis

### 4.1 Introduction

This chapter outlines the research findings and underpins this with the research and policies within the literature review. It also explains in more depth, some of the research touched upon, albeit limited, within the literature review. It provides a brief overview to set the context and then raises the main findings.

The data collection took place in the 2017/18 academic year, and during this time media coverage and comments on testing were prevalent. Headlines from O'Grady (2018), Turner (2017) and Weall (2017) raised in previous chapters, before the testing week highlighted the impact on pupils before they had even undertaken any test.

The impact these may have had on various agencies would vary from parents worried over their child's mental health, children themselves becoming anxious, headteachers feeling the pressure of the tests to come and teachers facing pressure from both headteachers and parents. These headlines are created and interpreted without understanding. They create impact (either positive or negative) and as this research will show, may not be reflective of all school situations.

The media discourse surrounding assessment, since the data collection period, is still clearly developing and not only highlighting issues, but also influencing public opinion. While not so much evident recently in regards to SAT testing, but certainly with GCSEs and the recent testing fiascos during and following the Covid epidemic.

The Let Kids be Kids social media campaign published a template letter for parents to withdraw their child from SATs, which was downloaded more than 2,000 times at the time of

the announcement on Heart Radio (10.4.18). Not only was this a successful campaign, but the beginnings of assessments almost being 'threats' to society. The moral panic that both parents and school staff were now feeling were fuelling government responses to the 'misinformation' published. It set the tone for future behaviours from educators, governments and parents and created further influences within Bronfenbrenner's mesosystem, which were not at play when the ecological systems were first created.

As already mentioned, Turner's (1987) test anxiety research and the Children's Society Report (2020) some 33 years later were supporting this notion of the significant impact of testing.

Given all this coverage and quite hard-hitting comments I began to doubt my own analysis and set about double checking my records and checking the conclusions with participants. The themes already identified within previous chapters, were reviewed again and the overarching conclusions remained the same; the new testing regime had not impacted negatively on pupils or staff. Extracts from the reflective diaries during the testing week and results day, highlight the thoughts of participants, showing they mainly noted no changes in the pupils. This was also support by the final group interviews where participants' opinion changed very little from their thoughts in September.

#### 4.2 Main findings

This section is written in dialogue with the literature review and focuses initially on the reaction of pupils during their test week. Unlike the policies reviewed within the literature review, this section focuses very much on pupils. I then compare this to participant accounts from the rest of the academic year. I then focus on the themes identified during analysis and highlight the evidence for my findings for each area and interweave these to form a narrative, linking back to the literature and policies reviewed, alongside Bronfenbrenner's

Ecological systems. This is then linked to the context of the research setting and my own role as headteacher. It refers back to concepts and research from previous chapters; mainly mental health, the impact of a new assessment system, how I have approached testing, teacher workload and pracademia. It also aims to answer the research questions.

The main findings of the research are;

- Pupils are not opposed to testing and the impact of testing is limited, albeit in this setting only
- Teachers did not perceive a significant change in the habits of pupils
- The current education system does allow teacher freedom.
- The positionality of the headteacher has significant impact.
- The pass/fail testing system is whatever schools make of it.
- Achieving the correct ecological system in a school is key.
- The impact of testing on staff is also minimal. The opinion of staff during their initial start of year discussion did not change when asked the same questions at the end of the year.

#### 4.2.1 Microsystem Analysis of Testing and the Impact on Pupils

At the heart of this research was always the pupils. How would they fare with a pass or fail testing system and what could I do about it? It is therefore logical to begin the finding sections with the test week itself.

“I wouldn’t say there have been any pupils whose behaviour has changed in a negative way on the build-up to the test.” (Participant C during the end of year interview, July 2018).

“Pupils seemed relaxed and happy at breakfast. All had something to eat and drink.” (Participant P diary, 14.5.18 ).

This reflective diary, written on the first day of the testing week and at breakfast club, highlighted that the pupils showed no signs of nerves. Breakfast is not something they usually have in school but during SATs week all pupils are invited in. This ensures they are not late but also ensures they are not hungry and have time, if needed, to calm down and be reassured. Despite breakfast being a new experience and it being the first day of a week of tests, no behaviour linked to testing was evident in any of the participant diaries.

“As yesterday, pupils were relaxed at breakfast and were chatty. They certainly ate more!” (Participant P diary, 15.5.18).

The following day, having already completed one test, the pupils’ attitude had not waned. Breakfast club was still attended and the pupils appeared unfazed by the tests they had already sat.

“Children were happy and no behaviour issues. Some chat around the days questions but only generally expressing if it (the test) was easy or hard or what answers they had put.” (Reflective diary participant C written at the end of the testing week).

These extracts highlight that as tests continued, the attitude of the pupils remained calm and unfazed. Discussions between peers were not worry and anxiety but general curiosity. The reflective diaries did not highlight any discomfort or upset, if during discussions pupils had given different answers.

“All year 6 approached their SATs with very mature attitudes. They worked remarkably hard, even when faced with some very challenging questions. Really proud of them.” (Reflective diary participant M, completed at the end of the testing week).

It could be that the pupils were still in ‘test mode’ and without the capacity to think clearly while in the moment of a week of tests. The reflective diary of staff participant P however,

from the week after SATs (21.5.18) also highlights the little thought given to tests taken and instead highlights that the pupils had indeed already moved on.

“Not much said about SATs or school work generally. The class were very excited about going to London and the World Cup.”

This is the same the week before SATS (7.5.18) with participant M writing “Most children seem excited about SATs tests, with some expressing confidence.”

At the time of analysis, I clarified if ‘most’ in this comment had meant that some were nervous and anxious. The reply given was ‘not really’, that these pupils had not commented either way.

“Mrs R had sat next to T during lunch on Tuesday (reading test day) and mentioned how nice it was to sit with him and how positive he was being.” (Participant C, reflective diary wb 14.5.18).

After days of testing and more interestingly after the reading paper, which is the one the pupils usually find the hardest, this particular pupil had no negative comments regarding the test. This pupil (according to the participant diary) is one that does not always find school easy and needs lots of support with their own confidence,

“Overheard S saying to dinner staff that SATs was good this morning.” (Participant diary C, 14.5.18).

The fact pupils were discussing the week positively and more importantly it was their choice and in their own time (over lunch) to discuss them shows that testing was not having a detrimental impact. If we refer back to Reay and Wiliam’s (1999) research, this research is in stark contrast and has pupils who ‘will be something’, as opposed to ‘being nothing’. The test is not having the impact, it perhaps once did.

In the weeks after SATS, pupils were not commenting on if they thought they had passed or not, and lessons continued as normal. Reflective diaries had no comments about tests or

pass marks from pupils, but as illustrated already they had moved on and were now focussed on the World Cup.

“Energisers and arithmetic continued as normal. Puzzle type activities in maths with M, K,T and E showing good resilience when working together. Rang R’s mum regarding being rude to dinner staff. (Participant C, diary extract 21.5.18).

In a similar way to how the pupils had moved on, participant diaries also highlighted learning continued and the school expectations remained. The school day and expectations did not stop because the tests were over, school was preparing them for more than the tests. This extract highlights not only how learning was still expected, but how pupils still demonstrated resilience and did not give up. It would suggest that the pupils saw learning as more than just for the benefit of passing tests.

The tests were over but learning still continued in the same way – problem solving/resilience developing challenges. Neither would poor behaviour be excused just because pupils had sat their tests. SATs were not the end goal and the expectations of pupils are there permanently. The participant diary did not indicate that this behaviour was as a consequence of testing, more so that school life continued regardless of testing and certain behaviour is not tolerated.

“Continued with maths energisers/GPS energisers and Friday arithmetic with the class completing as normal. None have questioned why they are still doing it.” (Participant E, diary extract WB 4.6.18).

“New text started: Last Kids on Earth. Amazing work on audience and purpose. Maths worked on mixed decimal calculations. Quite a few misconceptions but children were better Tuesday. Rounding fine after modelling.” (Participant J, diary extract wb 21.5.18).

Despite the SATs being over, lessons continued and misconceptions still addressed. The participant diaries highlight that new texts were introduced, misconceptions addressed and



school continued. It was not raised in any diary that pupils would be leaving in a few months and nothing now would impact on any grades, so not to bother addressing things. The diaries suggested that teaching was not and is not to pass the test, pupils need to know decimals for life and therefore it would be addressed. Books will still be read and discussed even though the reading test is over. Reading is a key skill and will continue to be addressed in school.

Joseph dress rehearsal went well and children tried their best. Production on the evening was a hit and only 1 child a no-show. No noticeable changes in behaviour since transition – majority are pleased to be back and don't want to leave, (staff participant E, diary extract 9.7.18).

This extract, one week before the pupils left, identifies the behaviour of pupils toward the end of the year had not really changed either, other than not wanting to leave. They had at this point, received their SAT score but still attended a school production in their own time. They had not felt despair with education if they had not passed, as the many media outlets at the time had implied. The child who did not attend was due to personal circumstances.

"Telling parents results went well but there were no children who were a shock or surprise." (Participant C, diary extract wb 9.7.18).

All results were given to parents and most children turned up for the meeting. All parents were positive even when their child hadn't passed. They could see their child had tried. A few parents spoke about rewarding pupils and O's dad was very pleased as O had no English 4 years ago. M, K and C should have passed but there had just been too many questions from them. Explained it didn't mean they couldn't read. K and B passed higher than expected. No tears with results and nobody questioned the result given (Participant W, diary extract WB 9.7.18)

This particular example does not reflect the research already mentioned by Ferretti, Ganley, Kofler (2009) or the (2007) research of Harold, Aitken &, Shelton, which identify the links between parents fear of failure and academic attainment. This research identifies those parents did not see the 'failure' and were supportive of their child's attainments and

achievements, despite technically not 'passing.' Perhaps a reflection of the social media impact, previously discussed.

"Thought C may be disappointed with not passing as they had really tried hard. He didn't seem that fussed. Said he had tried and his mum agreed he had. Think I was more bothered than he was!" (Participant M, diary extract WB 9.7.18).

There were many examples of how results day, with a pass or fail grade also did little to impact on pupil attitudes. The excerpts highlight those pupils knew they had worked hard or knew why they had not passed and how to improve, which would link to the Growth Mindset in place within the setting and previously reviewed. The pupils had not felt defined by a test score. Unfortunately, while the pupils at the micro level have not necessarily been impacted, the institutional apparatus of ability setting at secondary school education, highlights the exosystem's alternative view and this may be simply felt at a later date.

I definitely feel that starting the support in September and then gently being able to ramp up the level for then is far better for the pupils than a 'SATs frenzy' after Easter or whatever would be. I think the strategies that are in place - targeted support, SATs breakfasts etc. should remain and continue to develop and evolve alongside the actual teaching. Testing to death is not it! They are continually prepared and so other than breakfast club during SAT week, they notice no change. Though breakfast club is more about a rite of passage and ensuring they have been fed before the tests and not any last-minute boosting. (Participant A Group interview July 2018).

"All the cogs together make the machine and I do not think there is one single element that is the key to success." (Participant A group interview July 2018).

"I love that the majority of pupils are prepared to undertake new challenges and with confidence in the support network around them if they fail." (Participant K, group interview September 2017).

On transition days, the incoming year 6s (current year 5) are always aware of the fact they will do the SATS tests at the end of year 6. They often say this is a worry or something they are not looking forward to. I've always found, in the spring build-up, that pupils really build a determined attitude towards the tests. They know they're

coming and they want to do their absolute best. (Staff participant W, start of year interview, Sept 2017).

Having seen lots of year 6s through the assessments, I don't feel that our pupils have negative reactions to the tests. I often hear horror stories from other teachers about children completely breaking down, storming out and having big reactions. Our pupils tend to just get on with them. We've had tears over the years, but this is usually one or two pupils during one of the tests and it isn't repeated throughout the entire test week. Some pupils even seem to enjoy doing the test, enjoying the quiet of the classroom and the chance to better themselves (Staff participant A during the group interview at the end of the year, July 2018).

Pupil BB during the leavers assembly and reflecting on his whole primary school education stated: "My favourite memory is completing my year 6 SATs and feeling I had done well."

I acknowledge that this particular pupil did achieve in the tests and got the 'pass' mark but he didn't say "passed" he said 'had done well'.

Pupil AA at the same leaver's assembly stated, "Something I have learnt is to not be afraid of tests. Before SATs I was really worried because I thought they would be hard but they were actually pretty easy and not scary."

Pupils understandably are unsure about the testing regime in school and know regardless of their opinions and feelings, that they will happen. The ethos of the school and attitude of staff are able to put the pupils at ease. The examples within the diaries, consistently refer to attitudes and effort, not attainment and achievement. This is the message passed to pupils.

Participants were also clear throughout their reflective diaries that pupils are ready for tests without continuous testing. The day-to-day teaching, expectations and school ethos mean that pupils are well prepared and supported, with SATs being one aspect of school life, instead of THE one aspect of school life.

This reflects Bronfenbrenner's (1968) eco systems and the ability of school as the mesosystem to limit impact on the microsystem. In this research, the pupils were not impacted on by the pressures of testing and policies. The previous assessment system of pupils working through vertical sub levels to achieve final levels, was solely based on assessment and assessment leading teaching, to work through each rung of the ladder.

I am an optimist about our society, because there is something deeply rooted in our culture: we are pragmatists. If something isn't working, we don't stick with it. It may take a long while, but we eventually face up to our problems and do something about them. (Bronfenbrenner, 1979, p.462).

This previous system, while giving pupils more information than pass or fail, did create an endless list of criteria to achieve. Therefore, if the general research findings are also linked back to policies previously reviewed, it could also be suggested that in allowing schools to devise their own assessment criteria following the removal of levels; this research setting, which did not create a list of criteria and ongoing formal assessments, allowed pupils to see more to school than testing. Despite policies being created and enforced by other ecological systems, the very notion of these policies allowed some interpretation and eased the 'assessment conflict' in place prior to the changes. Thus, in turn, they either reinforced or in the case of this research setting, lessened the impact between the different ecological systems. Despite the criticism I made in previous chapters of the educational policies which underpin this research. The Final Report on the Commission on Assessment without Levels (2015, p.10) ideal that 'Schools should be free to develop an assessment approach which aligns with their curriculum and works for their pupils and staff', has allowed the research setting to reflect their ethos of a lesser value to high stakes testing than previous policies permitted. It allowed for the removal of testing other than statutory testing. Despite the worries of a new testing system initially, it had given freedom to develop a different system away from an assessment-based curriculum. I would suggest, that policies can live in a grey

area and not necessarily the black and white of their text. The choices of a headteacher, in implementing policies, sit firmly within their control of the ecological system within which they are positioned.

Referring back to Hannon's (1990) perceptions in the reduction of 'teacher autonomy as a result of government policies leading to a distrust of the profession'. I would argue, this research shows that government policy has trusted headteachers and the policies reviewed have permitted them to make choices. However, in stating this, perhaps headteachers need to be looking for the grey area in the first instance and then, in turn, trust their teachers.

Unlike the research of Conner (1991) raising teachers had become burdened with highly prescriptive curriculums, increased paperwork and work load and heavy inspections systems to ensure compliance and get results. This research had found:

The pupils don't see a difference really between year 6 and any other year group. They don't suddenly get to us and everything is different. They don't complete reams of tests and only have maths and English lessons. They have the wow days, topics and lessons the same as everyone else. Actually, they do get a residential which the other year groups don't get. (Participant A, July 2018 end of year interview).

Within this setting the participants made it clear during discussions that they had previously, that prior to 2015 that they felt they had not been free to use their professional judgement, but were teaching to the test and following a strict teaching regime. Prior to my headship, topics were not chosen by the participants, and neither were they really interested in what they taught, as they had little say in what happened in their own classroom. The reflective diaries mirror the view that this mindset and shift away from the importance of testing as central to school improvement, was embedded. The removal of levels and new National Curriculum had allowed change within the setting.

The Educational Excellence Everywhere 2016 White Paper has echoed this also. Whilst acknowledging higher expectations for pupil knowledge and skills, they highlighted more

teacher freedom on how to teach and assess. This is reflected within this setting and teaching is not to the test or a narrowing of the curriculum. The year of reflective diaries did not show testing, in test conditions, at any point over the academic year other than the statutory test week.

The diaries also highlighted all national curriculum subjects and did not just focus on maths or English. As participants were not given guidance on what to include, it would suggest that reflecting on all subjects shows the value given to all curriculum subjects, not just those tested. Despite Sutherland's stance that the three R's are the fundamental thread of education and were seen as the basic function of schools during the initial introduction of compulsory education. It could be argued that the priorities, while still taught and the main source of testing, are no longer suitable for the world today. The value of arithmetic when many adults all reach for mobile phones to work out a calculation, is one example. Despite the SATs historically having a calculator paper, the new assessment system within this research, has a 30-minute written arithmetic paper. Pupils also need to demonstrate an ability to write in a range of genres, demonstrating set criteria for each genre. This shows a focus on writing, when very few adults hand write things. While the research setting still legally had to teach the 3 Rs, they had tried extremely hard to make them relevant in context to pupils and moved them away from being the core time of education, to being equal in relation to time spent compared with other subjects such as art, music, history etc. As already highlighted, subjects other than English and maths were a substantial part of the reflective diaries over the year; this shows the value placed on a broad and balanced curriculum. While Sutherland's (1988) argument does show the differences between what is tested and the skills needed for a global market, despite the White Paper of 2016 using the global market and worldwide skill set as the foundation for reform. The research setting had attempted to make real world links to show a genuine purpose to learning and develop the skills for a global market. Kilpatrick (1925, p. 266) had promoted an education 'free of fixed

subject matter' and one that would rely on thinking and methods of action. While the research setting was not free of subject matter and had statutory lessons to teach, it was a setting that had become more than testing. It was real world learning. It was also reflective of the principles of Expansive Educators.

The setting's changes reflected this and their website highlights careers and skills are interwoven to curriculum design, their maths strategy links to real world skills such as percentage teaching in relation to loans and tax. Where possible the context for any learning and real-life application, is always given. This directly links back to William's (2013) desirable features of authentic tasks and James' (2008) focus on skills not facts.

Testing has been taken in their stride as pupils were prepared for the world daily and the education continued after tests had been sat and results given.

#### 4.2.2 Microsystem Analysis of The Headteacher's Role In Relation To The Impact on Testing

Within Bronfenbrenner's theoretical framework the main findings highlight the limited impact of the various outer ecological systems on each other, when there is consensus between the participants in each system. Within this research setting, the pupils and their own needs and achievements are at the heart of decision making and therefore any policy or testing system is implemented with this in mind. The policies from the outer ecological systems are implemented but with limited impact due to passing through the same leadership, teacher and support staff filters, all of whom create an additional layer of protection. The end product has been filtered in the same way by all those involved in the setting, resulting in a finished product and thus impact which means testing is not allowed to be detrimental. The impact of testing has been filtered away by the various levels within school through their

ethos, curriculum, timetables and so on. The microsystem created within this setting has multiple protective layers that any out ecological system must penetrate to impact.

The headteacher's role in this essential. The microsystem had been led and created with a shared value of education, which did not have testing as a significant stakeholder. The curriculum design and ethos of the school was reflected in all that happened. As headteacher I had actively employed and developed staff with a shared goal, to allow the equilibrium of the ecological system desired to flourish. In doing this the demands from external ecological systems and demands within this ecological system were limited. As teachers themselves did not have data targets, then they themselves as subject leaders did not require data from other teachers, or pupils. The demands within this ecological system of school and testing were viewed as a statutory requirement, no more and no less. Yes, the data would be published and shared with parents and pupils, but there would be no professional repercussions from this from the headteacher. Any 'concerns' from the macrosystem via the local authority or Ofsted, for example, would go to the headteacher and remain there. They would not be passed down to teachers and would remain within the protective layer they, as headteacher, provided. The responsibility for my decisions should and must remain with me. I am that pracademic bridge between school and other ecological systems and between the many elements within the ecological system of this one setting. I now see my role more clearly as the traffic controller, deciding who enters our ecological system, controlling the flow of traffic and more importantly not letting people see just how busy or fragile the bridge is.

In hindsight, all of this has to come from the headteacher. Whilst at the start of this research I was not aware of how much impact I genuinely had, it is now clear that this stance is not in every school and other staff and pupils have a different lived experience. Indeed, it was not there when I was the deputy and we tested everything that moved. That is by no means a



criticism of my predecessors, simply a reflection on the impact headteachers do have. In honesty, it is difficult to understand how my own views have been so influential. I would always deflect by stating I am the headteacher and so of course people will do as I ask, whether they really agree or not but this simply is not quite the whole story. Mainly because I do not ask anything specific really. I do not decide on topics taught, time spent per subject, how lessons are taught etc. Staff collectively decide this, they are in the classrooms not me. They create the knowledge organisers of what they want the pupils to know and actively listen to what the pupils ask for. I am of course aware and they inform me, but I do not ask or tell. That said, if I thought something was not right, then I would say and it is this that leads me to realise that staff would also respond in the same way towards me. If my view was in some way distorted, staff would be confident to say so and that is how this ecological system survives, on a basis on trust, of honesty and a shared passion for our pupils. The many examples highlighted during test and results week, show that staff have that freedom and trust to make decisions. This is then reflected in them listening to and valuing the views of pupils. The attitudes and actions are reflected initially from myself as headteacher, to then staff and then pupils. The setting's ecological system works in unison and due to the work on resilience and mindset, as will be discussed, created a climate of honesty, possibility and trust.

#### 4.2.3 Research Context

So how did that trust and honesty within ecological system come into being? Given the findings show the new testing regime has not impacted on pupils any more than other testing regimes, it is important that the setting in which the research sits, is firstly understood in order to make links and understand the conclusions. In doing so, I acknowledge as not only a researcher but also a headteacher, that the context of this particular setting may never be

fully replicated. The findings are based on the setting's staff, pupils and community at the time of the research.

The setting of this research had implemented the principal ideas of Growth Mindset in 2014/2015 as a logical progression from the multiple intelligences and VAK (visual, audial, kinaesthetic) that had previously been in place. Within this setting, it had been implemented to raise aspirations and develop resilience, more than improve the quality of teaching or raise standards. There were posters in every class informing the children they just could not do it 'yet' but eventually would. The school held parent information sessions, had staff CPD and went through the conveyor belt of 'we are a growth mindset school'. In reality, lip service was paid but the idea of not 'writing children off' was the main gain from this newly adopted philosophy. While the setting left the notion of mindset in place after a year of 'selling it', it was not something actively enforced or monitored, it was always something in the background. It was felt that without showing the children how to achieve something and that simply acknowledging with pupils, that the teachers they genuinely believe know and can do everything, actually also struggle and get things wrong. No number of posters and courses would change that. This reinforces my earlier stance in relation to top graduates and the notion that you also have to be able to 'walk the walk'. It also relates back to Williams and Thompson (2007) and the notion of teachers also being learners. The staff within this ecological system, as already stated, share many of the characteristics in their own backgrounds as the school and community they now serve. This in itself creates a shared ecological system of belief. Staff are also not afraid to show the things they find difficult. They too, as the pupils do, know that effort counts and that finding things difficult is not anything to be ashamed of. When appointing staff, how they 'fit' and interacted with staff and pupils was always the most weighted objective, more than paper applications. Again, the role of the headteacher in this was not obvious until analysing this research. In appointing staff, I was actively reflecting the pupils and creating that shared purpose.

Sometimes, 'wild card' teachers were appointed who had not had the best lesson observation but there was something there. It was the staff who would be applying the same filters and not fighting against the system in which they were to be employed. Those that were not afraid to say I do not agree, this doesn't feel morally right or I'm not sure I understand.

The research setting acknowledge that they haven't changed mindsets completely but do now have pupils that are prepared to try and trust in staff to get them where they need to be. The staff have been in their shoes and the three mindsets of expansive educators, previously discussed, were embedded.

This setting always acknowledged that students need to get the right answers in tests, but the critical issue was how do they achieve them. The pupils needed to be able to do more than just pass tests and as already raised by Lucas, Claxton and Spencer (2012) key life skills are also good exam passing techniques. This was also the ethos of the research setting at the start of this research. They are being prepared for life and to change the world, not pass a test.

However, achieving this ethos had not been without its issues. Prior to the research being conducted, the setting identified that in removing all testing they had removed too much and when pupils came to sit their SATs, they had never seen a test paper and felt far from ready.

Bushby (2018) highlights that one of the most common complaints is that the whole of Y6 is pulled out of shape by the looming tests, so simply avoiding the test date does not necessarily deal with the root of the problem. As the headteacher, I accept that my own ideologies and views of testing had impacted on pupils in this instance in a negative way. I had removed the test but not necessarily addressed the real issue. While the whole of year 6 is not designed towards SATs, as Bushby (2018) raises, they are part of the fabric of year 6 life and I had simply now created a blot on that tapestry. The solution agreed was that

work in books would be 'test style' and therefore the layout, format and wording would be similar, while continuing with the removal of optional SATs in all other year groups. In Y6, work completed in class each day, would be very similar to work they would expect in a test. To show the Y6 pupils this, the setting introduced them to a past test paper but not always completed as a test. To show them, in the hope of reassurance, that the skills they had and the lesson they participated in would allow them to sit the test they were all aware they had to sit. High stakes testing can focus teaching towards the test. Instead, as already discussed, the participants pointed out the links between work in class and the format of tests and completed them together as a class. Participants sold to pupils that work in class had a clear purpose and that was not just to pass a test- it was a skill to keep for life. This reflects the expansive educator view that real life learning would boost exam performance, not hinder it. Pupils could see a test but not have to complete them as a test.

The biggest shift implemented within the setting, prior to the research, was that of the teacher giving answers and moving to the children finding out for themselves. The setting moved away from staff imparting their knowledge to pupils, to show they knew it, to facilitating pupils thinking and problem solving. Skills that would be needed beyond the four walls of the classroom and could be transferred. As the headteacher, this involved a significant deal of trust between myself and staff. It involved allowing pupils to be wrong and to some extent struggle. Early years foundation staff (EYFS) are very adept at allowing pupils to learn through play and trial and error, yet as this phase of education ends pupils are suddenly not expected to ask questions and discover for themselves. How pupils are taught (mainly at a desk with little play or limited hands-on learning) and what they are taught (now key facts and not enquiry based) shift. Perhaps this is due to the formal testing within the next key stage (KS1) following EYFS, but if as a setting, the formal assessment data was not our end goal and thus the research setting's end goal, then the shift away from EYFS learning did not have to abruptly end. This was the start of our journey. Pupils should enjoy

school in the same way they enjoy those first years. They need to remain active learners and not simply passengers now placed on a conveyer belt of testing.

For staff this meant changing how they taught, trusting that the leadership team would not be observing and questioning pupils making errors and being stuck. Quite the opposite, it would now be expected. The relationship with the ecological system of school permitted this shift, as there was a shared collective of it is the pupils that count not our egos. As the headteacher, I had not suddenly decided one day that this is what we were going to do and the next day we would all do it. Instead, it was a discussion over a long period of time and an agreed action, based on our collective thoughts and feelings.

Initially to support with 'being stuck' the research setting used Dweck's (2012) Growth Mindset and then Claxton's (2002) 'building learner powered schools' followed finally by Lucas, Claxton and Spencer's (2013) 'expansive education.' However, as already addressed, the setting did not know they were building learning power or being expansive educators! The premise of these is that education is about preparation for the future. At its core is that pupils will have the confidence and skills to achieve in the world. This can also be linked back to Chandler's (2014) work on resilience needing to be part of the everyday school world. Failure, in some guise or another is part of school life and as raised by Hart, Blincoe and Thomas (2007) is also about things not becoming any worse. As Prilleltensky (2005) discussed, resilience must go beyond a phrase limited to how individuals cope with adversity. It was how we were now expecting them to cope with being stuck.

To support with implementing a more pupil led, enquiry focus curriculum, as Lucas, Claxton and Spencer (2013) had argued. The setting supported staff with this by using Wenger's notion of a community of practice. A community of teachers, with a shared goal to improve the education of their pupils, who would learn to do this together. Alongside this ran Daniel

and Wassell's (2002) 6 domains of resilience to staff: Secure base, education, friendships, talents and interests, positive values and social competencies

As raised in previous chapters, teachers worry that if they say they do not know, that they do not know their job and will lose respect from parents and students. They fail to also seem themselves as a learner (Claxton, Chambers, Powell and Lucas, 2011, p.81). This was a significant shift in ethos but one that was vital to create real change.

Claxton, Chambers et al (2011, p.208) highlight it can be difficult for teachers to discover or be told that there are hidden costs to 'good teaching' in terms of doing too much for pupils and therefore creating passive and receptive pupils rather than proactive and reactive ones. Headteachers need to ensure that staff do not feel judged or blamed and have time to understand for themselves, why they might benefit from understanding the learning process and creating a culture of enquiry. Headteachers also need to demonstrate that they too are learning and can get stuck. A vital element missed when first undertaking this research, was allowing myself and the school to be stuck to see what happened. I expected to see it in classrooms, to see struggle and how pupils respond to this, but worried that not stepping in at the start of the new testing regime would be detrimental. Initially, I failed to extend my own expectations to my own research.

Whilst not all the ethos of the school was reflected in the policies reviewed and despite the weaknesses of the policies, which have already been identified, resilience had been an area of focus before the policy release.

By the time of this research – some 3 years later, the setting did have a complete shift in the culture of the school and had achieved an environment led by pupils finding their way, with the guidance of staff. The school had a safe environment and that love of learning, which was reflected in their Ofsted report. Unity within the ecological system was clear and this in

turn had created a more solid wall between the other ecological systems making them harder to penetrate.

It is perhaps due to these changes that testing had not impacted on pupils in the way the media had suggested it had. The data findings also highlighted similar themes to those areas adapted and implemented prior to the research.

### 4.3 Micro System Analysis of Attitudes and Behaviour

In this chapter I will focus on my three umbrella themes and analyse these in relation to the research setting, while linking these to the final conclusions, theoretical framework and the literature review. In starting with this particular theme, attitudes and behaviours, the smallest ecological system (micro) and the pupils themselves are the starting point. When looking at a year of data from participants the overarching finding that testing had not impacted, suddenly needed explaining; this was now apparent, was due to 'something' and the context of the school then became open to analysis, as did my role as headteacher. The findings already identified and listed at the beginning of this chapter had to be due to the ecological system in which it was based, but what were the main factors within this eco system which were impacting? After much colour coding of all the data, three themes emerged: the attitude and behaviour of staff towards not just testing but education being the most dominant. This, in turn, then leads into the school's learning environment and finally to how labels both by staff, pupils and within other eco systems has impacted. In these next sections, I will discuss in depth these three themes and then demonstrate how the role of the headteacher has influenced these.

Policy makers and practitioners who advocate the need for resilience do not always seem clear on how it works and how it can be developed. As was made clear earlier, the very definition is open to debate. Within academic literature, resilience appears to be linked to

testing, socio economic groups or adversity, without an overview of resilience outside of these areas. Whilst adversity could be at play within this research setting, it is not a something significantly focussed on. Issues beyond the school's control, are just that – beyond our control. They are acknowledged but not an excuse. Achieving academic results are also not the main focus of the setting or research, neither are the backgrounds of the pupils. Therefore, it is difficult to grapple with academic literature in this field due to a lack of consensus on what resilience is within an educational context. However, the stance of Mohaupt (2009) referred to earlier as a process and not a personality trait, is the stance the setting had worked on.

Pupils had not been allowed to give up, we just needed to work more on the process with some pupils than others.

“Z and R very defeatist but once had intervention in assembly, saw it wasn't as bad as they thought.” (Participant C, diary extract 12.3.18).

In the changes the research setting had made, Bronfenbrenner's microsystem of the classroom had a consistent message in all interactions that aided resilience.

Resilience is something that the setting had worked hard on prior to undertaking this research and as already identified gave it the description of 'bounce back ability', The international Resilience Project cited in Daniel and Wassell (2002 p12) and the 15 check points they identify are beyond the remit and control of a school. However, there are some that the setting had clearly focussed on, these being points 3,9,11 and 15. Staff noted in their reflective diary in February that one pupil approached reading tasks using a different strategy:

“T decided to work backwards and do the most difficult text first. Seemed to boost confidence and didn't effect speed.” (Participant C diary extract 26.2.18).



Staff could have addressed this and indeed had ample time to do so before SATs in May, but instead allowed the pupil to do things their way. During the reading test paper, pupil T did indeed start at the harder text and worked backwards. This was a strategy she felt happy with and it was her plan. Rather than try to enforce the method the rest of the class were using – starting at the easier of the three texts and work through them, staff were confident in letting this pupil develop her own plan and influencing her own outcomes. Staff followed the school ethos in that they are not there to tell pupils how to learn, they are there to facilitate them taking control of their own learning. This particular pupil may well have had completed the wrong strategy at perhaps one of the most vital points of her education, but it was her choice. Should it be the wrong choice, staff would not know until results day. What they did know, was that her finding her way was more important to her than the results were to the school. She was fully aware of the fact she was sitting a reading SAT test, yet approached it as any other reading lesson.

Another link to resilience seen in staff's reflective diaries were during art lessons in November:

Lots of 'I can't draw' comments today. I spoke to them about when they were little and how they would paint, sing, dance totally without any hint of being self-conscious and somehow that can get lost as we grow up. RM was still reluctant to draw. I asked her to listen to my instructions, follow my guidance and just enjoy the experience regardless of the end product. She did and a good piece of work was produced that she both enjoyed doing and was pleased with. (Staff participant A reflective diary 13.11.17).

These examples also continue to reflect the culture within the mesosystem and that consistent messages throughout school creating a barrier against anything other than resilience and pupils being permitted to find their own way. Had this consensus not been in place, staff would have been advising of alternative reading methods or agreeing with pupils that they actually could not do something. The 'rules' of this settings ecological system were enforced and pupils were allowed to discover their own paths.

However, it would also be fair to acknowledge that without fully understanding the research setting, these actions could be interpreted as the participants simply not caring or having an ethos that cared so little about testing, that it was irrelevant which strategies pupils used as the data and test scores did not matter. This is not the case and is highlighted as such later in the year, when a reluctance to just try at art was still there but now pupils were prepared to listen and try, some progress within the development of building resilience had been seen:

A few general 'I can't' comments as the task was introduced. I asked them if they had any idea what they were being asked to do? Had they received any instructions or tip at this point? Had they watched any demos? Answers were all obviously 'no'. Pupils were told to wait until all those things were done before thinking out loud they can't. All pupils did this and there was no more 'I can't', all prepared to have a go. (Staff participant A, reflective diary 30.1.18).

As the researcher, when I first read this comment, I was under the impression the staff member had been upset or perhaps annoyed, that pupils had given up before even knowing what the task was. During my monthly check ins, I raised this issue, and the staff member did comment that she was frustrated they are not afraid to get things wrong in maths, but in art when actually there is no right and wrong, pupils will not try. In hindsight, we had clearly spent the time developing resilience within our reading, writing and maths lessons and this had developed the required results in these subjects, yet pupils were not making those connections outside these subjects. Had pupils only thought staff did not mind if they got things wrong in reading, writing and maths. Did we need to be more explicit and physically say, it is ok to get things wrong anytime, in any lesson? As staff had not been asked to make any changes this year while the research was being undertaken, this staff member was not sure she could "point out the obvious" or if they had to "discover it themselves." Once it was clarified that we can make small tweaks to the teaching and classroom, the staff member tackled future art lessons making links to resilience in all lessons. Art is not

something that is testing, and while only two extracts from many have been identified, both examples show staff's passion for the pupils to improve and achieve for themselves, not achieve a good test score or produce a masterpiece, but simply to try and be better than they were. This is not the participants not caring, it is them allowing the pupils to control their learning and for participants to support this, not control it. It is resilience at work.

If we refer back to Hunnybun (2012, p.19) and the 4 characteristics of resilience:

1. It is a process not a trait.
2. Resilience is not stable and will vary at different times of a person's life.
3. Protective factors can be located within families and the community.
4. Resilience is multi-dimensional and should not be implied in all domains.

While the four characteristics have been addressed, albeit briefly, the second characteristic can be shown in a diary extract:

DB, more academic than creative was in tears as he couldn't draw faces and 'was no good at art.' I spoke to him and the class about perseverance, determination to do the best job possible, aiming for a sense of personal achievement in completing the task, rather than the end result." (Staff participant A, reflective diary 8.2.18).

When the art lesson continued later that week, the diary extract read "Tackled to work more confidently and no tears!"

These extracts also highlight that pupils have tears and upsets not just linked to tests and core subjects but in other, more creative subjects. If I refer back to Dweck (2012) and the notion of not showing weakness, this particular pupil was 'able' in core subjects and as Dweck suggests may not want to show weaknesses in other areas. What I have learnt is that I needed staff to show and develop mindsets outside of reading, writing, maths. It had been introduced within this context and therefore had remained stuck within the confinement of those lessons. It was an opportunity to show pupils and staff that everyone finds things hard at some point, that another individual may find easy. This pupil may be one of the

better readers in the class, but they really struggled to draw. Pupils needed to see this and as one pupil put it

“We’re smart in different things but we’re all smart.” Pupil TC (reflective diary participant C, 8.2.18).

If we link this statement to the actual tests and refer back to Robinson’s (2014) notion that any child’s ability to concentrate and persist in an activity is fundamentally related to how they feel about the activity.

“A great deal of resilience was evident throughout all the tests.” (Participant diary M, WB 14.5.18), and “Arithmetic paper was challenging and not all finished. In previous years pupils seemed to have time to spare. Lots of questions for 1 marks that required lots of working out.” (16.5.18, diary extract participant J).

Referring back to expansive education and that when pupils are confident and independent learners, under test conditions they are less likely to go to pieces; they are more resilient when faced with difficulties and don’t give up so easily (Claxton, Chambers, Powell and Lucas, 2011). That good life skills are good test techniques:

“A different format for three-mark questions but pupils didn’t seem too phased. The test seemed accessible, but 40 questions was a lot and far more than any other year.” (Participant W diary extract 15.5.18).

“Reading was not as I expected but the children were great. Not all finished due to the increase in questions, but they didn’t appear flustered.” (Participant M diary extract 15.5.18).

Despite the reading test being a new format and not one known to them, pupils were not phased. Neither were they phased by the significant number of questions, which was an increase on what they had been expecting (based on previous years). They felt safe in their approach to the test and their ability to unpick the questions. The environment of the classroom and ethos of the school had created a climate for learning and application of

skills. Strategies, such as starting at the back, were accepted as valid strategies for that particular child. The strategies the pupils had, whatever those may be, were nurtured and valued and in turn the pupils felt comfortable being 'them' and able to behave as such.

#### 4.3.1 Microsystem Analysis of the Role of the Headteacher in Shaping Behaviour and Attitudes

If behaviour and attitude has not changed, then why? For this, we must look at how the behaviour and attitudes of pupils and staff have impacted on the research findings.

I have seen pupils who cannot take a chance of 'failing' and so will kick against the tests and all they stand for. Fortunately, at our school these pupils have been few and far between but I believe it is a fine line that most of our pupils walk and that often some only make it to the end of the tightrope because of the level of tailored support the school puts in place for them. (Participant A, Initial interview).

The vast amounts of work on resilience, self-esteem and the value placed on effort appears to have created a 'buffer zone' or deeper barrier between the various ecological systems within the school. Pressures to get results appear to not be passed to the pupils from staff and this is also reflected in pressure for good data from the SLT to the teachers.

The treatment of pupils as people and not statistics, targets or simply part of the job, aids Dweck's (2012) stance that removing labels impact on the pupils' behaviour and attitude. The class are treated as individuals yet also a unified class, with diverse needs, strengths and weaknesses all of which should and are to be embraced. This has to be permitted by the headteacher, to ensure the values of the school and its strap line is embodied in everything it does. Staff were confident that it was not the data that was the most important aspect of the test examples given, but that it was the pupils taking control of their own learning and destiny that was the ultimate goal and this is not something one individual can do, it has to be the norm within the ecological system.

Social influence is easily identifiable by the way individuals conform to the attitudes and behaviours of the majority. Building upon the work of Turner (1982, p.165) previously raised, three reasons for individuals to deny the evidence of their own sense to go along with a majority view are identified: the need to depend on others for information, to test the validity of our own opinions and the achievement of group goals – achieved through a unified purpose and need for approval.

One example of this is:

Maths investigation today and all enjoyed it. Children could pick their own groups. J headed to B (they are good friends). B backed off stating 'I'm not working with you. Last time you were in my group, you didn't do anything.' (Participant C diary extract 11.12.17).

In this example, the pupils are not wanting 'clever' or 'able' pupils, or even their friends, as they know that everyone has different skills they can bring. What they do expect however is that everyone will do something. They expect the class to conform and try, regardless of who they are. This is also a further example of the ecological system living its 'Everyone is important' strapline. Everyone is important. And regardless of who you are you will work!

As already highlighted, the ecological system had created a shift towards the pupils working harder than staff and not having pupils reliant on staff for their thinking. This extract also supports the ecological system's insistence upon this. It also highlights the collective responsibility for sharing this aim, it reflects the community of being in it together, which I, as headteacher, had facilitated many years ago. Pupils were also aware that answers would not be spoon fed to them, and that they needed to try, try and try again before staff would intervene. This again supports Savin-Baden (1996) belief that children should be taught to be stuck, as it is an important part of learning. As pupils were seated in mixed ability groups this initially created frustration as some pupils needed to try again more often than others. Martin Seligman (1967) identifies the concept of learned helplessness, which arises when

students are not given the opportunity to think for themselves. Expansive educators identify helplessness of this kind becomes 'habituated to giving up or waiting for someone else to give the answers.' These pupils' lack resilience and resourcefulness (Lucas, Claxton, Spencer, 2013, p.33)

My decision as headteacher to ask staff to not step in, to give pupils time and allow that EYFS method of learning, had moved the setting away from the many cases of learned helplessness it had prior to my headship.

If we refer back to the 16 habits of mind identified by Costa (cited in Lucas, Claxton and Spencer, 2013), These are very much reflected in the diary extracts below.

All tried and had a go. There was so much to do in some sections and it almost seemed unfair as it was only worth one mark. Nobody (thank God) looked at the marks on the corner and thought this isn't worth the effort. (Participant K, diary extract 16.5.18)

Have been looking at working out missing sides of shapes to find perimeter. Challenging at first, but most children got it. D very upset because she just couldn't see it. Looked at using bar models to work it out. (Participant W, diary extract 5.2.18.)

The above diary extracts show pupils being left to try and even allowed to get upset in order to support their own educational journey, before support was given. The extract also shows that despite so many test questions, pupils simply continued and put in the effort to get through.

While Costa's (2008) habits of mind were not purposefully focussed on, they are reflected in the data:

K spent far too long drawing circles to divide as she clearly had the wrong answer when she had a remainder. Pleased she continued, as she got there eventually, but it wasted time for her and we have done inverse operations forever! A little frustrated

she had forgotten year 6 methods, but at least she had a backup. (Diary M extract 6.5.18).

In this example, rather than get upset, leave it or leave it and then come back to it. The pupils took the stance of: this method has not worked, what else can I do?

K was sitting next to the new girl in maths. It was paired work on area and perimeter. I heard K give a clear and correct explanation and how to calculate it. D (new girl) did not understand so K tried again, this time drawing and making annotations. D did still not understand. K then got very flustered and burst into tears. I asked what the matter was and K replied she didn't understand area and perimeter and can't do maths. I told K she had given a very clear explanation and had done everything asked of her but I would take over and work with D now. (Participant P, diary extract 19.2.18).

Despite pupil K being correct and having an understanding, she had doubted herself as she had not been able to help others. This example also links to the attitude of testing. When the first explanation by K, had not worked, she tried again slightly differently. In the tests, pupils tried things differently when their first method had not worked. The decision made by the headteacher in suggesting to staff to take that step back, to allow pupils to be 'stuck' is evidenced in these examples and reflects the principles of expansive educators already discussed. Pupils will have many plans to use and when they don't, that is when staff will step in and not before. Whilst this may seem a very simplistic statement, the concept is actually very difficult. It goes against your natural responses to help, takes up valuable lesson time and to those not aware, looking in – may appear as though the teaching is poor. Whilst not challenged at the time by staff, it is only now that I realise the significant ask this was of staff and the trust that we both had in each other.

Referring back to Bourdieu (1990, p.77) and his 'habitus as a system of dispositions linked to a set practice'. It is the basis for regular modes of behaviour and therefore the effect of the habitus is that those who are equipped with it, will behave in a certain way, in certain



circumstances. Through sheer perseverance of staff and pupils, habits were instilled and soon became the norm. It is also testament to the ecological system that the trust in the greater good allowed staff to move outside their comfort zone and take the very chances they were asking of their pupils. Again, the staff reflecting very much the pupils within this system.

The focus within classrooms had to be learning and not simply showing their ability by completing lists of sums or writing pages of a narrative. Performance goals are associated with measuring ability. It focuses on pupils measuring themselves on their performance and if they do poorly, they may condemn their intelligence and fall into a helpless response. Learning goals focus on mastering new things. The attention falls on strategies for learning and so when things do not go well, it is not a reflection of intelligence but simply that the correct strategy had not been found/applied (Dweck, 2000, p.16). This was the lesson pupil K needed to understand.

In Dweck's (2000) research, students with learning goals were much more mastery-orientated in their approach to the challenging new problems. Those with performance goals, despite being just as able, were thrown off by the novelty of the problems. They spent too much time worrying about their ability:

Primary school pupils are small human beings and like adults their attitudes to test differ. Their attitudes change like the wind. They are children, they respond to success and so a hint of perceived failure can change their attitude in an instant. (Participant A, initial interview).

Spent some time discussing with R that he needed to show his working out. He was adamant he didn't as the answer was correct, which it was. It wasn't until we moved to decimals, and we then had to spend 5 minutes instead of 2, unpicking what the issue was that he finally agreed with me and I could not see inside his head to help him. (Diary extract A, 17.1.18).

This links to Rhodewalt's (1994) notion self-handicapping and deliberately withholding efforts when confronted with new challenges. The reasoning for such a lack of effort allows the pupil to deflect from their ability and retain the self-belief that had they studied and worked hard, then they could have achieved better. The examples of pupils becoming frustrated highlight this quite clearly. They tried and then staff stepped in to unpick when things were going wrong. Staff had worked hard to not step in too quickly and to allow pupils ample time to grapple with the issue. If staff were going to help, they needed a starting point of where things were becoming an issue and therefore needed to physically see the pupils' processes. This does not follow the natural testing process, where no support or unpicking could be given. In tests, pupils could make mistakes and nothing be mentioned; the pupil not realising they had made a mistake and when told after the test, there was nothing they could do to change it. It was therefore vital, that when faced with the unknown a learned helplessness did not take control. Not just for the passing of tests, but for the benefit of future employers and the collective desire to be ready for anything.

This same principle was threaded throughout the test week, it was not about getting everything correct. As previous quotes highlight, participants did not comment afterwards on strategies used in the tests or indeed pupils not finishing.

Dweck (2000) identifies intelligence as a growing issue that gains importance as pupils move through school years and experience successes and failures themselves and observe those of their peers.

This is not to say that participants did not have concerns over pupils, of course they did.

What they did not do was pass on these concerns to the pupils:

Was concerned about H and S and the pressure of the test. H coped really well and worked her socks off. S worked better when read to but accepted the confidence booster of a smile or a gentle 'keep going,' (Participant C, diary extract 14.5.18-test week).

L is notorious for speeding through things and making silly errors. He really slowed it down and used his extra time well. GT can be really slow and unproductive in class, I didn't see any of that today! (Participant C, diary extract 15.5.18 – test week).

Participants did not express their concerns to the pupils. They had not told pupil L to slow down. When I enquired with this participant, during the monthly check in if they had expressed their concerns to the pupils, I was told not. The participant felt the issues were theirs and not the pupils:

I feel they work too fast, they don't! I have been saying since September to slow down and to read carefully but it goes in one ear and out the other. I can't change it, so instead we agree that when they finish, and they do finish early, they have to go back and check it. He probably didn't want to read all that again! I can't make them do anything, I can just drip feed the advice and hope they listen.

Introduced long division in maths. A, announced her love for this method very early on. J pushed against it and continued to use short division whenever he could. At one point he was asked to use long division and he said no. A, fell out with the method during the week when errors crept in but stuck with it and loved it again by Friday. (Participant C, diary extract 6.11.17).

What is clear from the diary extracts is that participants did not attempt to change the learning habits of pupils by telling them they had things wrong. They present the options and allow for choice. Perhaps eventually the pupils will run out of choices and agree the one suggested in the first place, but the important aspect in this, is that the pupils had the choice and discovered for themselves. They were not told.

“AT worked from the trickier text to the easiest in today's session. Seemed to help her, as she was more resilient as it went on.” (Participant W, diary extract 26.2.18).

Participants did not stop the pupils starting from the back of the reading text, tell the pupils to speed up or slow down or to give up with long division if it was too hard. They allowed them

to work with the skills they had. In the extracts above, suddenly telling a pupil to work slower or to start at the front of the book may have created doubt and gone against the value of pupils in control of their own learning and strategies that worked for them. The pupil had chosen their path and were permitted to do so, despite the views of the participants. That said, if pupil strategies create a 'wrong' answer than obviously they were addressed:

Misconceptions when reading and using timetables addressed. Children were happy to arrive to meetings late – “It’s only a few minutes!”. Worked on not rounding to the nearest when using time. Always get the earlier bus or train. (Participant W, diary extract 16.4.18).

This extract also highlights that participants are preparing pupils for the real world. It isn’t just about the strategies to work out answers to questions, but simply that it is not acceptable to arrive late.

Maths- multiplying decimals by 2 digits. Advised the children not to use the column multiplication method. Children were vocal in their disagreement in this. Allowed them to give it a go their way. They quickly fed back it was too tricky and wanted to try it my way. (Participant C, diary extract 29.1.18).

In this scenario the staff member could simply have said they were the teacher, they know better etc and not lost teaching time letting pupils find this out for themselves. Instead, the staff member allowed them to try it, developed that trust and at the same time pupils were going to fail and would be able to bounce back once the new method had been shown.

There are many examples of staff addressing misconceptions and even obvious refusals from pupils to listen:

Pupils remembered key angle skills from the last time we looked. Introduced protractors and they enjoyed the practical task. Developed confidence but need to develop accuracy. Informed the class they need to look carefully. (Participant W, diary extract 23.4.18).

O reluctant to show his working out and appears to want to be the first to finish. Asked by teacher to show working out but he did not. Stayed with me. Explained that:

- The teacher had asked him to, so he was to do it.
  - It shows the understanding of the method.
  - It allows errors to be fixed quicker as we can see the process.
  - In a test, showing working out may gain marks, even if the answer is wrong.
- O accepted this and we went through it again, this time showing his working out.  
Participant P, diary extract 17.1.18

“T given some advice to check the question in reading. Refused and said he didn’t need help and thought what he was putting was the correct answer even when told differently.” (Diary extract 26.2.18, participant E).

When I clarified with the participant if the pupil continued to not listen to advice, they replied he did not listen for the rest of the week. He even cried when he got his reading scores that week but did not see it as his fault, it was the ‘stupid’ text. The staff member and pupil agreed that another ‘stupid text’ may come up next week in the reading sessions, so perhaps the pupil should look at another way of doing things. The next week, the pupil did as the teacher had asked but not because it was the teacher method but because he had thought about another way to do the task. The staff member while annoyed they had not listened in the first place, was wise enough to know that sometimes, you ‘choose your battles’. They knew from the tears that the child was bothered enough to listen, even if they would not admit it. They also had the vital skill of understanding and knowing him, as earlier diary extracts show:

“T brought three homework pieces back in a row, which is unheard of. He said Dad is making him do it!” (Diary extract from participant C, 5.2.18).

“T returned his homework three weeks in a row. Think he is secretly proud of himself but reluctant to show it.” (Diary extract 6.2.18, participant K).

These excerpts also show that staff have a very good knowledge of the pupils they work with and therefore show that they are best placed to show the impact on pupils, rather than the pupils themselves. My role, as headteacher, has allowed staff to genuinely know their pupils, to see them as people, not data, and to be invested in them. Whether we agree or not, to do this takes time. Time that could be spent on core subjects or testing, not conversations. Allowing staff to manage their own time and curriculum coverage gives the permission to do this.

“Repeated lesson from last week as not confident investigating on their own.” (Participant J, diary extract 4.12.17).

“Independent writing planned but no idea what happened with boxing up. Will look again at it tomorrow.” (Participant M, diary extract 11.12.17).

This extracts also highlight the participants stance that learners are central and that knowledge is not imparted from teacher to student, as previously highlighted. It also supports the statement earlier, that learning is lifelong and not about passing a test. Neither of the examples above are things tested, yet they were still repeated as the skill was needed. It would not impact on a test score.

I was pleased with Y6 maths and S passing was a nice surprise. M wasn't far off and perhaps she could have tipped into a pass on another day. Reading was disappointing. I was frustrated for the pupils that were so close and would have passed if the pass rate had not moved or through a subjective mark scheme. I knew how much work had gone into reading this year and felt the results didn't reflect that. I also knew it would impact on RWM (reading, writing, maths) and that bothered me the most. Although it wasn't a drop, I felt the staff and pupils deserved better. I felt better once I had worked out progress and that looked good. Perhaps I felt more responsible than normal due to the number of years I'd taught this class. (Participant C, diary extract wb 9.7.18).

Despite the findings showing testing and results were a positive way for pupils, the above extracts highlights that the participants themselves were frustrated.

When I first read this, I was unsure why the staff member was upset by the combined score and if it was the pressure that this would be published. I was advised it was not the

publication of the grades that was the issue but that the significant jump of the pass mark (4 more than previous years) would now mean pupils did not achieve across the board and they deserved to. The staff member, after discussion, was more annoyed at the system than the impact it would have on pupils.

“I believe that over the years the tests have become less about the child and more about scrutinising the school but, unfairly, the child is still the one expected to give the performance.” (Staff participant 1 during interview July 2017).

Whilst evidence of the impact of testing is identifiable in these extracts, this example was not an impact on pupils or one that was fed back to pupils, it was a professional observation highlighting frustration with the exosystem impacting on the mesosystem.

In conclusion, the interactions within the mesosystem interconnected with each pupil's microsystem, to assert the positive influence needed to allow pupils to be themselves and in charge of their own learning. While the impact of the exosystem and macrosystem is clear at a staffing level, the mechanisms such as ethos and staff values, meant that the impact of these ecological systems does not filter down to individual pupil micro systems but can be contained within other eco systems. The ability to filter information between the various ecological systems within this research setting is established and influences individual behaviours.

In relation to the headteacher's role, in creating and reflecting the microsystem in terms of staff and pupils, developing the trust and ethos of the school values in everything required. This is largely significant and has contributed to the limited impact testing has had on pupils. However, despite the natural desire to shield pupils and staff from the pressures of other ecological systems, this is not something that can ever be truly diminished. The perfect culture within a school can exist but staff will still feel that pain, feel things have been unfair, upset for their students etc, as they are human. However, in the same stance that I, as

headteacher, wish to limit the impact on staff from other ecological systems, staff also mirror this towards their pupils and as I model, do not pass on their concerns of frustrations. Perhaps basic leading by example is at play, but that does not fully explain why staff follow my lead or indeed trust in my leadership. For that we also need to analyse the next theme that emerged; the learning environment.

#### 4: 4 Mesosystem Analysis of The Learning Environment

Here we move from the micro system to the mesosystem and the impact of the research setting on these two systems. This is the next theme identified from the data; the learning environment. The learning environment allows the behaviour and attitudes, whatever they may be, to flourish.

As already highlighted staff are not afraid to change and adapt the school day to address the needs of their pupils, whether this is repeating lessons or teaching it a different way. The setting's learning environment allows for flexibility.

Planned ratio using the bar model in the hope this would support K. She didn't use it and continued with her own method. D and R did use it though and so it hadn't been a complete waste. Will try with this method again tomorrow (Reflective diary participant J, 1.2.18).

'Tried again with bar model. K attempted it and appeared more confident. Quite a few children prefer this model so we will continue to use both systems.' (Reflective diary participant J 2.2.18).

Assembly ran over (bonfire issues I think) and so rather than rush our text, we will move this to next week and continued with a short text extract for today. Informed the pupils and they agreed to do another extract from last week's text instead. (Participant A, reflective diary 6.11.17).

Great teachers teach students how to reach high standards. Dweck (2012) suggests that teachers spend hours on planning alongside careful thought on individual pupils. Nothing is



left to chance. The staff plans account for most eventualities and daily changes/issues are seen as part of school life. This is not to say they spend hours on planning, they do not. What they have is the freedom to react and the time to really know what is needed. Time is not wasted on things that do not impact, Staff do not leave things to chance because, as already raised, they know their pupils and have the freedom they need to run their class as they need to.

The ratio examples also highlight that just because some people can do something with little or no training, does not mean that others cannot do it with training. This is important because many people with a fixed mindset think that a person's early performance tells you everything you need to know about their talent and their future (Dweck, 2012). In this example K needed additional training and an alternative method in order to understand the lesson. The staff offered this, even though the need was not evident for the majority of the class.

“Started looking at equivalent fractions using diagrams. Pupils struggled to draw their own but seemed to get the concept. Will provide diagrams tomorrow.”  
(Participant J, diary extract 8.1.18).

In this example, we see the staff participant admitting they had got something wrong and addressing it. When this issue was raised at the staff check in, I asked if the lesson was repeated. I was informed it was and when the pupils had asked why they were doing it again, the staff member had informed them it was because the lesson had not been planned well enough and they (the teacher) had got it wrong. The learning environment is one where anyone can get things wrong, even teachers. Again, reflecting students flourish when their teachers are visible learners, as already highlighted by Williams and Thompson (2007).

Following the new reading test, the three-mark questions had been formulated in a new grid method. The staff participants had not seen this method and actually believed it to be a better model than the ones they were using. After the test, the classes had a discussion on

which format they preferred and allowed the pupils the option. This also reflects that teachers should and are willing to adapt if there are 'better' ways to do something. The learning environment is an ever changing one and one where opinions are valued, as highlighted by discussing an alternative grid method for reading and whether or not they would prefer it. In the pupils also voting to use an alternative model, they were also moving away from a system created by the teacher but were confident enough to agree with their teacher that they wanted to leave the current system and start a different one. The pupils were confident they would not upset the teacher and the teachers were not offended. The learning environment is one of honest conversations, which are not taken personally. Everyone in the school is learning, staff included.

Another factor the learning environment caters for: is addressing the needs of their pupils gently, yet constantly preparing them for life:

The preparation begins in September, although the pupils would not really notice it at first. A lot of effort has gone into providing a system of almost drip-feeding the things that are required. For example, on a daily basis pupils will be given questions in maths, reading and GPS that look like questions on a test paper – the type face, the answer boxes, the types of problems, the vocabulary and language used etc. The pupils just become familiar with the style so when they eventually see a test paper, it seems more comforting and less daunting. (Staff participant 1 during the initial group interview, September 2017).

I think that in recent years, we seem to be preparing pupils a lot better than we used to. Spending days looking at old papers has been less beneficial than giving pupils 'snapshots' of test style questions on a more frequent basis. Pupils need to know how to prepare to deal with the unknown when it comes to a test. We need to ensure that pupils have positive learning habits. (Participant 4 during the group interview, September 2017).

While this may appear to be teaching to the test, I would suggest that it is different. The pupils are not completing practise test after practice test or only being taught things they will need for a test (as the art lessons demonstrate). As the research participant makes clear, it is to expose the language and layout pupils will encounter in everyday life but in a less threatening way and to see it enough times so it is 'normal'.

The weekend after SATs, staff were saying:

Spoke to my year six about the weekend and to have a normal weekend, good night's sleep and about the unique skills SATs didn't test. Children seemed to appreciate this chat. M was beaming from ear to ear as she left. (Participant C diary extract 11.5.18).

"Didn't want to overdo reading so no session today. Instead, we used the time to finish independent writing and redraft." (Participant W, diary extract 11.5.18).

The above diary reflections also highlight that after the test week itself, it had not been identified that the tests were the final destination; instead, that school life continued. They were advised they had unique skills and also redrafted work; work did not end.

Classic text read: The Speckled Band (Sherlock Holmes). Used reciprocal reading approach. Lots of work needed on vocab due to the age of the text. Some confusion 'considerable state of excitement' as pupils expected her to be happy rather than agitated. (Participant J, diary extract 12.3.18).

The above extract also highlights the needs of pupils being addressed, in this case vocabulary, during every day lessons. Language is a focus and a part of reading sessions, classic texts are covered to support this.

While resilience has already been discussed, it does also link to the school environment and the extracts below link clearly to the three fundamental building blocks of resilience, identified by Daniel and Wassell (2002)

S has the role of checking the behaviour chart each morning. It is becoming clear it is impacting on her morning time and was beginning to feel rushed and as a consequence is not finishing her morning energisers. She has been told not to worry but has asked to resign to focus on her work instead. We agreed. (Participant A, diary extract 17.1.18).

R spoke to me today. He has been working with Mrs M during PE on areas he has struggled with. He has asked to not miss PE as he enjoys it and it isn't fair he misses it to do things he doesn't like. Made sure he knew it wasn't a punishment but was to help him. Agreed to move his session. (Participant E, diary extract 22.1.18).

“Class enjoying their creative topic and asked if we are covering the same D.T project as last year as they wanted to do it.” (5.3.18 diary extract, participant C).

The reflective diaries highlight that these essential building blocks are embedded within the school environment. These examples show that pupils are confident to express their concerns to staff. They are secure enough within their class to do this and also know that their opinion is valued and learning is not a done to process, where they are simply passengers. They are active in their learning. It also highlights that staff are not focussed on data. Pupil R needed interventions and support but the teacher was not prepared to do this as the cost of areas they enjoyed, in this case PE. The sessions would have already been timetabled and planned, but the teacher was prepared to change all that once the pupil had raised concerns. In the other example, despite the teacher not being worried over the pupils missing her maths energisers, the pupil clearly was and therefore the staff member took this seriously and offered them the security of remaining in class. Again, this role would also need to be addressed, creating more work for the teacher. Growth mindset leaders believe very little in hierarchy. They put the right people, regardless of position, who can help solve problems in the meetings (Dweck, 2012). Staff are not afraid to listen to pupils if they have a ‘better’ idea and in turn the leadership team are not afraid to listen to staff.

In terms of self-worth, the above examples also highlight that pupils feel their views are valued. Pupils also demonstrated confidence on many occasions:

“Investigating angles. Unsure at first but all listened carefully and by the end of the lesson all confident and happy.” (18.4.18, participant A diary extract).

“High confidence to tackle radius and diameter.” (20.4.18, participant P diary extract).

“Reading: Children showed a good understanding and appreciated the imagery created and understood how it was created. Children had contrasting opinions which we explored.” (18.12.17, participant J, diary extract).

While most quotes highlight general confidence in terms of skill, the final quote raises the idea that children are also confident to share their opinions even when they may not be popular or in disagreement with their peers. This is testament to the learning environment created within the school. Again, we could link this to the consensus within the meso and exosystem within the research setting. With each agreement and 'consensus' creating an additional layer between itself and the eco system within it, meaning the opportunity to filter through the outer eco systems to the pupil becomes harder. The mesosystem of the research setting was careful in what reached down to the pupils themselves. As headteacher, not only have I realised I was traffic controller of the ecological bridges, but also ensuring barriers and additional layers are in place within our own ecological system. I had created a 2-pronged defence system.

In contrast, the same consensus in that pupils and their thoughts and feelings are valued allow flow between their microsystem and the mesosystem of the school. The research setting's 'starting with child' ethos allowed and encouraged the micro system of the pupil to impact and influence the mesosystem of the classroom and school. In this research it could be concluded that the microsystem impacted far more on the mesosystems activities and adaptations than any other ecological system interaction. The ethos of the research setting reflects starting with the child and thus the microsystem.

Finally, linked to growth mindset, pupils are aware of their strengths and weaknesses and see neither as a burden or failing.

O used the word pioneer from a text studied in October. He has worked really hard with his vocabulary. K has been upset this week. This has been an ongoing issue. Told her she would get there but will look at three-mark questions next week with her. (15.1.18, diary extract, participant W).

This example shows that pupils are aware of their weaknesses and are working on them, not sticking with the knowledge they have. The following examples shows that the pupils want honest accounts of what they can or cannot do. As Dweck (2000) raises, in order to improve one has to accurately know where one is.

“English: when given their work back to review they were very honest if they had been given a mark but had got it wrong or not punctuated accurately.” (16.4.18, participant C diary extract).

As the headteacher, reading back through the diary extract did give some sense of pride regarding the learning environment created but when reading the statement without context, it could again appear that the participants do not care and the pupils can do as they like – choosing which lessons they will miss, who they will work with etc. However, where possible context has been provided to show the positive learning environment established, though I accept the black and white context given within this research as a researcher, is not the same as the colourful context I see as the headteacher.

#### 4.4.1 Analysis of the Headteacher’s Role in the Mesosystem

Here I will look at how the headteacher’s role in creating the school’s learning environment has impacted on the research and to link back to Bronfenbrenner’s (1979) ecological systems and the impact of the systems on each other.

I think our pupils have opportunities to ask about the tests and know they receive honest answers. They are informed about the tests and know they have to be done and why. I think we should continue to be open and honest about the year 6 tests with the pupils and we should take the time to discuss and explore their feelings or worries, as they arise. We should build on the culture that tests are an important feature of school life, whether we agree with them or not, and we need to encourage the pupils at every turn to treat them seriously and respectfully. However, it just as important to let the pupils know that the tests are not the be all and end all of everything. They are not and never should be. (Participant A during final interview, July 2018).

“I feel the majority of pupils have a positive attitude towards the y6 tests and accept and respond to them almost as a rite of passage.” (Participant A, initial interview, July 2017).

From these comments, it is apparent that testing is simply seen as another part of school. No more important than any other and this is reflected in every day practice. The learning environment created by the actions and decisions of the headteacher reflect very clearly the status placed on tested.

When asked during the group interviews – If I walked in your classroom for the first time what would I see? The participants stated:

“A positive learning environment where adults demonstrate that they still enjoy learning even though they are grown-ups!” (Participant J).

“Pupils ready to learn, not only with a safe classroom and full pencil case but also a positive, can-do attitude.” (Participant P).

“Lessons that are relevant and engaging.” (Participant A).

“Hopefully lots of learning, lots of discussion about work and pupils supporting one another.” (Participant M).

“I feel very happy in my classroom and I hope the children do too.” (Participant W).

In analysing these comments, it could be concluded that the learning environment is one for all those within it – staff or pupil. I, as headteacher, listen to the ideas and views of staff and am not afraid to change course if someone has a better idea than I do. Similarly, staff are also open to changing and going with pupil ideas, rather than their own. This is another example of not only the ecological system working together, but also reflecting each other. The research participants want pupils to enjoy school and expect the pupils to be learning themselves, not being told what they need to know. As previously discussed, there has to

be a value to tasks and this value needs to be clear. Teaching, when possible, is taught in a context of real life. If those links are not clear, then pupils will challenge their purpose. This is evident in the maths test, when pupils had to identify the various combinations available of football tops, shirts and socks.

“R clearly annoyed in the maths test that the question didn’t seem to have anything to do with maths.” (16.5.18, diary extract, Participant J- test week).

This pupil was not opposed to the test, but the fact that in their opinion it was not maths.

The learning environment focusses on non-negotiables that pupils must have by the time they leave. These are not just the test areas but real-life skills.

‘All excited for the water fight. They are already discussing what ‘guns’ they are bringing.’ (Participant diary P, 2.7.18).

More lessons on time needed. They have no idea you can’t arrive late and that in real life if you don’t cook the cake for 30 minutes you can’t eat it. We will look at running through these issues practically!’ (Participant diary A, 9.1.18).

Despite Lucas, Claxton and Spencer (2013) raising that as the pressure to get the grades increases schools focus on test skills. The above extracts highlight in this setting, skills are taught for life and that a pupil’s whole school experience is not just academic learning. An earlier example relating to time has already been given but this further reflection by different participants highlights the value given to this particular area of maths. Time is often only one or two questions in a whole SAT paper, yet it is a life skill and so covered many times throughout the reflective diaries, especially when the pupils have not grasped the concept. At the opposite end of the scale: rotation, translation and symmetry often appear in tests, making lots of quick wins in terms of test scores, but within this school environment they are not seen as such a key life skill and so are taught far fewer times. They are not mentioned within the year of reflective diaries. As the water fight extracts suggests, time is also



invested in things not associated with school in a traditional sense. The water fight is one of the activities pupils wanted to participate in before leaving and the school environment allocates significant time to this. The whole school water fight is a full afternoon of school. This highlights that the curriculum and focus of the school is far from only teaching and covering things in the test. The tests are not ignored or not spoken about, but are simply just one jigsaw piece in the whole puzzle.

“SR upset at end of paper 2 as they didn’t finish. Spoke to him about how he had done his best and this result wouldn’t affect his future.” (17.5.18, diary extract participant J).

This is further highlighted in the staff response to a pupil that should have easily completed the maths paper and passed with no real difficulty.

The staff member did not promise everything would be ok and that they might have passed, as that cannot be promised. What they did do was praise the effort and put the test in context of the bigger picture. This also reflects the thoughts of Growth mindset and that it is effort that is the focus and not the finished product.

“Some children wanting and seemed to be spending lots of time with me and staying in over lunch. Many expressed mixed emotions about leaving.” (Last week of school, diary extract participant C).

“Transition days so not here for two days. More than half a dozen kids came back after school to tell me about their day.” (2.7.18, participant C diary extract).

A further testament to the school environment and that pupils’ feel happy and safe is the sheer volume of pupils that return to see staff. However, this can be a double-edged sword with pupils worried about leaving as they trust staff and enjoy school and worry what the future will hold elsewhere.

While many children are apprehensive about new schools, the secure climate we create can make that jump even more daunting. This is supported by the reflective diary accounts and the sheer volume of children that returned to school following transition days.

The expansive classroom, while not known at the time, was something the setting had created and as a consequence had very different relationships with our pupils and families than in previous years.

When analysing this research and trying to unpick why testing had not appeared to impact, the staff raised the notion that we look after every single child and noted that we appear to have a higher percentage of pupils excluded, or in isolation/behaviour units when at high school, than other feeder schools, yet the research setting's behaviour is not a concern.

School can act as a capacity builder for children. It can provide children with opportunities for positive experiences which provide children with experiences of mastery and success and in turn may bolster self-esteem.

However, in creating the environment, the research setting may not have considered the entire ecological system within which they are placed. That within the mesosystem are other schools and whilst this research had highlighted some notable links between pressures not penetrating each other in relation to testing and the micro and mesosystems, the same could not be said for different establishments within the eco system. This research setting's learning environment was impacting on the feeder high school. The setting was creating pupils who could express opinions, decide how to learn and discover from themselves through discussion. Yet was this environment reflected in other school within the eco system. At high school, could they give opinions? Could they discover from themselves or were they expected to work through text books in silence? My actions, while creating the culture for learning within this eco system and in shielding pupils and staff from this mesosystem, may simply have delayed the inevitable. That once outside the safety of this

microsystem and school, they are not prepared for the pressures that will be exerted. This is a difficult action to reflect upon but I genuinely feel this ecological system is the right one. That is not to say that others should change, just that I am confident the decisions made are for the right reasons. If the behaviour and attitudes, as already discussed are there, then this is due to the environment being there to allow it. However, a further action following this research is more work with our feeder high school on the types of learners we are developing.

Resilience has formed a significant part in the learning environment of the research setting and is identified as one of the reasons the impact on testing has been limited. However, while there appear to be many advantages to resilience and the theories associated with it. I acknowledge the criticisms regarding resilience raised within the literature review chapter. Although I would argue that the benefits outweigh the negatives and the very findings of this research support resilience as an essential part of school life.

The research setting had not lowered standards and had indeed increased expectation. This reflected the 2016 White Paper; Educational Excellence Everywhere higher expectations for every child. Yet this setting had walked the fine line of failure versus success without real incident. Within this research setting, the exosystem of policy within this theme had not impacted negatively on outcomes. This is supported by the White Paper's (p.9) acceptance that 'outcomes not methods' are the main goal. Despite this statement having the ability to be interpreted, those outcomes (therefore testing) are the most important aspect and impact negatively. Within this research the methods within the setting have impacted positively on the outcome, in relation to the impact of testing. The statement from Dr Boustead (Bushby 2018) highlighting pupils are experiencing too much failure and are unable to deal with this, is not reflected within this setting. However, the antecedent to this is how that failure is presented. That is not set, and is within the control of the teacher and has to originate from

the headteacher. The role in the headteacher in creating a successful learning environment is needed and is achievable.

The moral struggles identified by Ball (2013) do not need to exist within this setting. I would suggest the research setting and staff are not at odds or conflict with their personal views and school demands. It could be that the research setting has only been protecting their pupils in the short term and time will tell, as these pupils leave high school and enter the world of work, if that shielding has impacted beyond our limitations in a positive or negative way. When you read comments about leaders and teachers leaving education due to the pressures from the different ecological systems, Ball (2013), you realise that a headteacher's decisions can make or break careers, not just impact on pupil outcomes.

The by-product of this research: the hope of reassurance to parents regarding their concerns at the number of practice papers pupils complete (Bushby, 2018) and the testing of solely core subjects about which parents raised concerns (Bradbury, 2019) was not needed. My decisions and actions would offer this, but as is clear, has not been needed. It also highlights, the many headlines available that simply instil panic and fear but that are simply not at all reflective of this setting.

Within this research setting, pupils are not faced with any other tests, other than the statutory week of tests. As already identified testing is not a significant aspect of school and nor is the curriculum sacrificed to focus on testing, or additional work given, as the diary extracts have shown and, as highlighted in Bushby's (2018) stance that the year 6 curriculum is hijacked for the benefit of SATs tests.

As a researcher and headteacher it is easy to assume that all leaders have the same wrenching in their stomachs and feel morally obliged to react, yet are we at risk of swinging too far one way and pupils having a see-saw education off ups and downs, i.e. don't worry about test – worry about tests? While to some extent it is pleasing that testing and a pass

or fail grade has not impacted on my pupils there is a concern that this pressure has simply deferred this for them and it will be experienced later. As a researcher, this offers further study and possibilities, but as a headteacher it makes you doubt the ecological system you have created and if you have simply 'set pupils up to fail later'. This is one aspect of the research I am still personally grappling with. However, as a headteacher I am preparing pupils for a world of unknown knowledge and skills and due to it being unknown, I resolve my inner conflict with the notion that thinking and resilience will hopefully give pupils a solid foundation on which to build their future educational journey.

The principle of Building Learning Power (2011) is not to raise conventional results. It is to expand the range of valued outcomes, such as the development of confidence or capacity to learn. This occurs outside of school as well as in it. Expanding pupils' capacity to learn and their appetite for learning is as valuable an end of education in its own right, not just as a way of improving scores.

While in the short term, within primary school years, this research setting's learning environment has not allowed testing to impact as drastically as in some school, if the headlines are to be believed. But it is now a worry, when hearing one of the biggest attendance reasons for absence is anxiety (Prof Ken Reid 2013), as evidenced in ITV's Time to Talk adverts to reduce the increase in mental health challenges, that perhaps the setting has simply deferred the anxiety and anguish and not halted it. Despite the power of the mesosystem, this has not impacted on the school's microsystems as the headlines and policies already reviewed would have us believe.

#### 4:5 The Use of Labels and the Bridge Between the Micro and Mesosystem

This final theme came to fruition linked to the use of Dweck's (2000) mindset research, i.e.

the notion of the pass or fail label originally being researched and the reflective diaries using

labels in some form. The participants whether conscious or not had used labels – ‘not usually academic’, ‘emotional’. It was therefore an area that developed as within the monthly analysis, labels were also created to quantify thoughts and classify the data.

This final theme also straddles both the microsystem and mesosystem, with both labels used by pupils and the research participants and myself as researcher. It highlights the fragile interactions between the micro and mesosystem.

This theme links directly to the use of labels in Dweck’s (2000) initial thoughts of performance and learning goals and the review of her research in previous chapters. Labelling either knowingly or not by staff will impact on the choices the pupils may make. Dweck’s relationship between a pupil’s theory of intelligence and their goal choices are clear. Those with an entity theory of intelligence are more likely to choose performance goals. Those holding an incremental theory, were more likely to choose a learning goal (Dweck, 2000).

Labels for ability are not something actively used with pupils within the research setting. They are used in planning documents to highlight differentiation and the broad bands that pupils are working in. There has to be some starting point. Due to Dweck’s (2000) Growth Mindset implementation some years earlier, the setting had stopped labelling what pupils could and could not do and focussing on ability, to focussing on effort

Labels regarding the abilities of pupils are not mentioned in any of the reflective diaries at any point including following test results. The caution Gipps and Stobart (1995) and Sumner (1987) raised regarding labels and the limitations they then give, were not evident within the research setting.

As the earlier example of the pupil that would not show working out or the pupil that refused to listen to the advice of the teacher shows judgements and labels may naturally occur.

When teachers are judging students, the student will sabotage the teacher by not trying.

When students understand that school is for them and see a way to grow, they do not insist on safeguarding themselves. I have seen tough guys shed a tear when they realise they can become smarter. Despite their bravado, pupils do care. (Dweck, 2012)

While labels are not actively used, it would be untrue to say judgements and therefore some form of labels do not exist.

“Hull University visit – O negative attitude.” (9.7.18, diary extract participant C).

“Enjoyed bike festival.” (18.6.18, diary extract, participant C).

“London visit. Children excellent all day. Some behaviour issues on the train home (9.00pm) but may be due to tiredness.” (11.6.18, Participant E diary extract).

“Pupils excited about the upcoming performances.” (Diary extract, Participant A, 2.7.18).

“General air of trepidation about algebra.” (Diary extract, participant P, 10.3.18).

“R absent for a day. On return was negative towards his writing. Said he isn’t capable, when he is.” (4.6.18, Participant E, diary extract).

These extracts show that the participants are making judgements and labelling and are deciding if pupils are enjoying something, why they are acting in a certain way etc. As previous examples highlighted – labels such as ‘defeatist’ or ‘more academic than creative’, show labels have been given. However, while labelling, the focus is very clearly on the attitudes and actions, not ability. Participants appeared to be naturally focussing on effort and attitude and labelling that, not achievements and ability. Feedback given to pupils is on effort and attitude and the learning environment appears to ensure that it was this that formed the focus of observations. While levels and ability were obviously monitored, as all

schools legally do, they were addressed through effort and attitude. This is reflected in how participants wrote in their diaries:

“Focussed on 3-mark questions in reading. S reluctant to have a go. A and B slow and not getting finished. M and H giving them a good go.” (29.1.18, participant C, diary extract).

The staff participant has not listed scores or which part of the three marks they are struggling with. They have focussed on the approach. Participants were given little advice on what to include and so it is interesting to note that while scores will have been recorded somewhere, what the participants focussed on to write was effort.

“Introduced graffiti art. Children enthusiastic at the prospect but when drafting their own, soon became negative. R was very vocal in her frustrations. M was really proud of her efforts.” (Participant diary A 5.3.18).

The reflective diary analysis suggests that any labels used within the setting are a result of the learning environment in which they are situated. Had the environment been data driven and ability led, diaries may well have reflected this and more comments on ability and data would have been made.

This highlights how the views of one ecological system can directly influence and thus impact on the actions and views of another eco system, as will be discussed.

During the results week when ‘labels’ had to be given due to the new statutory creation of pass or fail. These labels held little value and pupils did not see them as something permanent. Their reactions were not disappointment or upset, no pupil cried or became upset, but instead the label was taken in their stride. It was another piece of educational knowledge they were given, as they are given daily. The meaning behind it they would attach for themselves, as they do with all information they are given.



“Told K her results and she thanked me and then asked if we still had PE.”  
(Participant E – results day)

“Nobody came in this morning and asked for their results. Not sure if they know it is results day or they don’t care!” (Participant A -results day)

#### 4.5.1 Analysis of the Headteacher’s Role in the Use of Labels Between the Micro and Mesosystem

Dweck’s (2012) view that assessments are not valid indicators of future success would support the notion that ability labels should come with a warning sign. However, labels for effort and attitude have, within this setting been successful. They created a non-threatening way to feedback to pupils. Rather than saying ‘let’s improve your test scores’ the reflective diaries would say ‘let’s look at some strategies to improve your speed/what are you happier doing/great effort so far now let’s look at...’

Labelling theorists argue that once a label is attached to a person, a tendency to see themselves as the label is formed and actions follow accordingly. Others then see this action and act towards them on this basis. This can lead to the self-fulfilling prophecy discussed previously, which argues predictions made by teachers about the future success or failure of student tend to come true, because the predication has been made. This research would suggest that saying labels should not be used in school, would be virtually impossible. Whether it is acknowledged out loud or not, judgements and labels will always be naturally formed, as identified already. What was evident in this research setting was that the labels that do occur have the least amount of negative impact as possible. This research would suggest that Dweck’s (2000) stance on effort being the sole focus of any label, has aided the overall findings that testing has not impacted on pupils negatively. In not labelling performance or ability, the sheer impact of testing and performance was removed. As with behaviour and attitudes, the pressures between the micro and mesosystem appear to have not impacted on each other.

Rosenthal and Jackson's (1968) research on the self-fulfilling prophecy was not permitted to exist. Within this setting, all participants expect all pupils to work. There are no excuses made for the disadvantaged pupils, or those with additional needs with regard to the effort they put in. The reflective diaries show no differences in expectation between the pupils, despite the participants class having pupils with SEN (special educational needs) and disadvantaged pupils. Many of the examples already quoted are pupils with SEN or who would be classed as disadvantaged. The diaries raised that every child gets support from the teacher and the expectation is the same. This is reflected in the national policies reviewed where the national standards are for all pupils, regardless of personal circumstances.

Referring back to Ball's (1981) three-year study of Beachside Comprehensive school, he identified that most students were eager and conformist when first entering school but this attitude changed due to teachers' stereotypes. Ball claimed the teachers' expectations of the different attainment bands led to strong correlation between banding and performance.

'Partner work and R didn't want to work with K and made it clear to her.' (Participant diary K, 1.11.17).

'CT made a great attempt at Spanish and joined in confidently. He even smiled' (Participant diary A 4.12.17).

During the regular check in sessions with the participants, it was explained attainment bands are used as an initial starting point they are not set and certainly not set against whole school life. Pupils sit in a variety of seats and in the first extract above, both pupils participating in the partner work would therefore, academically, be at different stages. During the discussion, it was queried if the student had not wanted to work with his partner due to this but the participant explained that he wanted to work with another male pupil due

to the teasing he felt he would receive working with a female. The monthly check in had also allowed clarification that the data was not being misinterpreted.

The second extract highlights a pupil being confident in languages, yet in previous examples he had given up at art. Both extracts support that the participants have not labelled pupils with preconceived ideas, or based on previous experiences. This particular pupil had not been labelled following his difficulties with art earlier in the year and was now being praised.

It would not be true to say attainment bands are not used full stop, there has to be a basis for the setting of work and pitch of a lesson in any setting. What this setting has highlighted is that attainment bands which when used as the starting point, are seen more as an elastic band, moveable, stretchable and each band can stretch more than others. The reflective diaries have many examples of pupils who have had things recapped or given support in order to achieve. The ability of the setting's pupils is not set and there are no examples of the participants stating a pupil has not grasped a concept and has been left behind with no further support given. In the examples used within this research, the participants have instead adapted, revisited and persevered with the teaching of the concept. The ability of the pupil has not been questioned in any of the accounts within the reflective diaries. As raised by Lucas, Claxton and Spencer (2013), in expansive environments the word ability should be used extremely carefully to avoid the mislabelling of talent and ability.

Given this research has taken place in a significant area of deprivation, indeed in an area deemed within the top 10% most disadvantaged nationally. It would be easy to assume that staff therefore had little expectation of pupils, as raised by Keddie (1973) or Ball (1981). However, the vast majority of teachers within this setting are first generation university attendees from similar social backgrounds to their own pupils. Perhaps one reason that ability labels are not used by the participants is the lived experience they have had regarding perceived ability and class. A further link to the literature review and teachers having first

class degrees in comparison to those who do not but have the lived experiences, can also be made here. As already raised my own personal opinion is that there appears to be some form of social justice-based motivation to work in particular communities; those that reflect very much their own experiences. As stated, the 'gut' feeling on relationships and understanding, far outweighs the paper exercise of the application form and grades achieved.

The many comments within the diaries were all linked to attitude, effort and behaviours. Staff were writing comments such as; "said he can't do it, but can." These were not aimed at the ability of the pupil but their attitude towards their work. Pupils were believed in, even when they did not believe in themselves.

Within the research setting the behaviour and attitudes are derived from the learning environment and the learning environment is created by the participants and the labels they give and use. Had the behaviour and attitudes or learning environment already discussed in previous chapters be different, then the impact and use of labels would in turn be very different. The culture of each ecological system within the research setting clearly impacted on, to varying degrees, the other ecological systems and any change within these sub systems would, in turn, impact on the change between the two sub systems. It highlights the fragile boundaries and interactions between each ecological system and how one small sub system of labels, environment and attitudes (as this research has identified) can directly impact or not on each other and into further sub systems. The role of the headteacher in developing and sustaining all these small but intricate relationships is essential. Developing effort as the key guiding principle, not attainment, has to come from the headteacher. Setting the expectation that pupils are not the finished product and their educational journey is far from complete has to be embedded. Prior to this research, I genuinely believed all headteachers gave this freedom, trust and respect to all their school community.

Unfortunately, this is not the case. The pressures to achieve, in relation to tests and data, are drivers in some schools and for this, I hope this research allows a little more flexibility in their decision making.

#### 4.6 Future thoughts

While I am mainly happy with the research methods, there are things that with hindsight, could have been improved. The pilot study, while useful, took significant time that had not really been factored in. Future research would need sufficient time allocated to a pilot study in order to address the issues raised within this research. Time with participants is perhaps the biggest factor that any future researchers would need to give significant thought to. It is a very fine line between checking the conclusions drawn as a researcher and encroaching on the participants' time. While this had been planned monthly, schools are ever changing places and the best laid plans usually need an alternative.

Larger sample sizes and samples across a variety of schools may also yield more benefits than a single case study within one setting and would allow for future comparisons of the outcomes of this research.

Referring back to the notion that the 'principal qualifications of qualitative research is experience' (Stake, 1995, p.49), while I have 20 plus years' experience of schools and the education system, I had little experience of research and therefore initially time was spent on understanding doctoral research and indeed how to undertake research. This is time that had not been planned within the research timeline and was a painful lesson to learn. At times, the findings were not really what I wanted to hear. It did not prove what I had hoped (that testing was wrong) and so why put myself through a Viva and weeks of writing up research? The findings would not be changed and I actually had nothing to change within my setting. In the end it came down to a moral decision that firstly governors had paid a

significant sum of money and secondly what example was I setting to staff to pupils in not finishing. I cannot expect pupils to persevere when things become hard, if I am not prepared to do the same. This was a further lesson learned, when analysing my decisions as headteacher. That, I lead by example. I do not ask anything, if I am not prepared to do it myself. I have been cleaner, dinner staff, support assistant and covered the office. I am no more important than any other staff member or pupil – Everyone is important and has a value.

A further thought is that of teacher perspective research and the tendency for these to be small scale and for the purpose of the researcher themselves. They are rarely undertaken on a large scale to influence outside the research setting. Therefore, whilst this research is aimed at making improvements within the research setting and to improve my own understanding of the consequences of educational policy and implications on practice. It also has to be acknowledged that it is a small piece of research in a large field of education research. As such, the impact will be small scale, in that it will be the research setting. However, while that was the simple aim of the research, to inform this setting and microsystem only, the findings have been shared with the local authority via the improvement partner and with other local headteachers. While, this has been done informally, it has been shared and the doorway opened for those future professional discussions. It is there to potentially influence future decisions, strategies and policies.

It would also be worth noting previous discussions here, in that teachers' despite wanting to and acknowledging its usefulness, rarely read research. This is not a criticism of the teaching profession. Had there been research initially on the new testing regime, in honesty I am not sure if I would have read it or contemplated implementing any of its recommendations. I too am guilty of saying 'but it isn't us, that isn't my school.'

With this in mind this research has limitations in that replicating it and the exact same ecological systems, even if the research setting were to repeat it, would never be exactly the same.

## 5 Conclusion.

In this chapter I will draw together the main findings and link these to the literature review, theoretical framework and discussion points.

At the start of this research, there were no preconceived ideas on what the research would uncover and the basic premise was simply to understand the impact of a new testing regime in order to make future changes within my own school. What could I do as headteacher? However, once results had been analysed and conclusions drawn, I was somewhat disappointed. The findings suggested testing had little, if any, impact on pupils and as stated, I even doubted my own findings, they went against the 'norm' and the many extracts already included which highlighted testing as having a negative impact. Whilst I did not know at the start what would be discovered, with hindsight, perhaps I had not been honest with myself. I did want to make changes and for the right reasons- to help pupils, but perhaps I had already decided, subconsciously, that the new testing regime would have a negative impact on pupils and as an educational setting, assumed we would need to change something.

The initial research plan allocated a significant amount of time to data analysis, in order to find practical suggestions for ways forward. Following the conclusion of the research, however, I found myself having to spend time investigating why the research setting did not need to make changes and this took much longer. It also altered the landscape of the research, turning the focus very much to the role of the headteacher and the power and influence they have within educational settings. Reflecting on my own decisions and their

impact, was not part of the original plan. As a first time headteacher, it was difficult to accept the positive steps the school had already taken which had served to reduce the impact of testing on pupils. Praise is not something that sits well with me, I always strive for improvement and whilst areas of development remained, I had to accept that what we had built was generally a success. It has become clear that not all headteachers feel able to make the difference and become lost in the 'traffic' directed from external ecological systems, losing sight of the reason they chose education as a career initially. The sheer volume of headteachers leaving the profession, highlights the fragility of the current situation. The five qualities of power Cossa (1996) raises are fundamental in understanding this. With reference to these qualities, parallels can be drawn between my own views and the impact these had on the power I held (as headteacher) and the extent to which I have, or have not, exercised that power. The regulatory power, in allowing policy implementation, has to some extent been addressed during the literature review of policy, but this can also be linked back to the very definition of power. It does not need to a top-down, single person entity but, as Bourdieu (2010) and Christensen (2023) conclude, power stems from 'analytical constructs developed for and through analyses of practice.'

In simple terms, the new testing system had little impact on the pupils or staff, with participants identifying that the children had not really had any behavioural changes.

What this research has identified is that testing does not always impact negatively on pupils. Indeed, within this research it proved to be a positive experience, taken in the stride of education.

While not fully agreeing with Gibb (2018), this research has shown that testing can be just another part of school life. It has also identified, that if the school culture is child-focused, testing does not need to impact on mental health.



The manipulation of power Cossa (1996) raises can be linked to positive manipulation here. The culture and definition of 'doing well' and 'achievement' is what a school can control, the label of pass or fail, as this research aimed to investigate, is actually whatever a setting makes of it. It how it is manipulated. Headteachers do have the power to persuade stakeholders that testing is not the sole purpose of school and, in so doing are able to create a change in behaviour that reflects this. While I am not sure I fully agree with Gibb's (2018) notion that more testing will help, this research has highlighted that testing within primary school does not need be anything to fear.

The stress felt by pupils and affecting their attitude, as identified by O'Grady (2018) is not evident within this setting. As already identified SATs were a positive leavers memory. Fear of tests and testing does not need to be common practice in primary schools and following this research, I would suggest that the statistics raised in earlier chapters do not reflect the national picture. The other side of these headlines is why are pupils aged 10 and 11, aware of the implications of results? Why have they not been shielded from that pressure? Sadly, this lies firmly within the school and the decisions made by the headteacher. It lies within the informational and interpretative qualities of power, as Cossa (1996) identifies. The power of the headteacher to gatekeep the information shared and interpret the actions of pupils should not be underestimated. If this research has given any glimmer of hope with regard to an everchanging assessment and testing regime, it is that things can change. Schools can and have, controlled the impact of testing and this needs to the headline which is also publicised. The media headlines raised within this research give little hope to schools, pupils or parents that things are within their control and that the ecological systems within which they are set can offer protection. Bronfenbrenner's ecological walls are permeable and power is far more complicated than a powerful and powerless system. The ecological systems of policy and testing are not powerful, in their own right. As already explained, it is given power and kudos by other ecological systems and agencies.

As Wertsch (2015) identified, Bronfenbrenner's basic assumption that one individual alone cannot be analysed, nor can headline data. The context, the role of the headteacher and ecological systems have to play their part. If schools, as the social environment and one ecological system, can master the balance of the pressures of testing against pupil needs and well-being, then they can also change the trajectory of its pupils. This supports the research of Zuccollo, Dias, Jimenez and Braakmann (2023), as raised in the literature review and the vital role in education, played by headteachers. This research suggests that all schools, whether they acknowledge it or not, will in some form teach to the tests. How they do this, is the key. Pupils can be prepared for tests, as this research shows, through discrete and subtle teaching, not going through practice paper after practice paper. Only being able to only pass tests through test training and being able to pass tests through the application of skills and thinking need to be raised as two different entities. The examples within this research show that the setting and participants have not hidden the fact from pupils that they will face exams and tests throughout their life and they will require particular skills to do this. As Claxon and Chambers (2011) identified it is the job of a twenty-first century school not just to build these habits in school, but to ensure they are transferred beyond the school gates, that pupils have lifelong skills.

Referring back to James' (2008, p.25) belief that learning is not simply absorbing information but an active process of meaning-making' is something that had been successfully created within this setting.

Within this research setting, the ethos and the lack of high stakes attached to testing allowed the micro ecological system and therefore pupils, to be protected from the pressures of the mesosystem. Thus, in turn, the mesosystem of staff are protected from the exosystem of policies and data. The role of the headteacher, and their application of Cossa's five qualities of power, is vital in creating the delicate balance between each of these ecological systems.

It could, therefore, be suggested that each ecological system can limit, in some circumstances, the influence of other ecological systems. As Christensen (2023) stated, power is something 'floating' between agents and not something someone possesses and others do not. It moves between the various ecological systems. The cultural capital and ethos of a school can determine its success, not test results. Referring back to expansive education; Lucas, Claxton and Spencer (2013) are clear that evidence shows when learners are helped to become more confident, independent and articulate about the process of learning, the results go up and not down. Within this research setting, students achieve good results through resilience and resourcefulness. Given that the research setting has achieved their highest test scores (across all key stages) since moving away from testing and teaching to the test, supports this viewpoint.

As identified by Torrance (1988), moving away from the traditional views of assessments and what education is, should be as one. The curriculum within the research setting and their view of assessment were not in conflict. The role of the headteacher in creating this equilibrium is at the heart of this research. Testing and assessment can be viewed differently and as this research suggests, seen as part of life and something that pupils are prepared for without teaching to tests. This is the consensus referred to throughout this research- the view that testing is not the end goal and that the day-to-day informal assessments, tweaks and conversations are how schools can impact. Once this is the accepted norm for the mesosystem, which I acknowledge is not a simple or easy feat, then the individual pupil microsystems can flourish without fear of tests and see assessment as part of daily life and a helpful development tool. However, as should be made clear, this has to come from the headteacher. The power to interpret, control, manipulate and regulate are intrinsically linked in the same way that Bronfenbrenner's ecological systems are linked. Neither can be completely separated and analysed as they are more than a singular entity.

They are complex interactions and decisions interwoven within a vast number of ecological subsystems.

Linking back to Wiliam (2008), in conclusion, tests are simply whatever we allow them to become. They are a product of how they are interpreted, manipulated, controlled and regulated.

The two original questions of this research aimed to understand:

- How can headteachers limit the impact of testing on pupils?
- How do we understand the headteacher's role through the lens of Bronfenbrenner's ecological systems theory?

In response to the first question; this research suggests that formal assessments do have a place and value and that testing does not implicitly need to impact negatively on teachers or pupils. While it would not be solely true to say pupils wanted to sit tests, they were not opposed to them either. Therefore, firstly, headteachers do not need to fear the notion of testing. They are able to control the narrative. Within this research tests were seen as part of life and something to be taken in their stride and not feared. As already highlighted, they also give pupils the official reassurance they required. My first response, to 'ban' testing as a first-time head teacher is not the answer. What headteachers can do, is ensure the status given to testing is something with which they are comfortable. They have the power to manipulate beliefs and behaviours in a more positive way.

The new testing system allows teachers to teach as they choose, pupils to enjoy learning and still provide governments with the data they require for policies, reform and league tables. It could be argued that any testing system implemented, can work alongside learning and not just teaching to pass a test. This is because the power really lies with the headteacher and not with the test themselves. Headteachers can work with their staff to ensure they know that the curriculum is not just a conveyor belt of tests/assessments to tick off. Schools can be values driven not data driven, if the qualities of the headteacher permit.

The power of headteachers is there to be used and controlled, it is not something which is predetermined. If the values and ethos in schools correct, the data will follow. Whilst not officially part of this research data period, the setting did achieve 100% in maths and were placed in the top 2% nationally in 2019. Data for reading and writing were also significantly higher than floor targets. Given only 15% of pupils were on track on entry to the setting (at the start of my headship), while reluctant to take any credit (as I do not teach the pupils) my actions and choices have allowed staff the freedom to teach, and to teach with a purpose. Results were never the aim of my actions but a by-product of choices. Attainment data forms no part of the research setting's appraisal process, the focus is on progress only. The judgements of staff are completely trusted. If they feel a pupil is at a set level, they do not need to justify that professional decision. The evidence will be there in the many books completed by pupils. As headteacher, I give power to staff to make their own decisions and also trust in their judgements; the need to test is removed. The headteacher positionality on this is vital, not just in terms of the power they give but also the value they place on their staff and their professionalism.

Teacher workload does not need to increase in order to do this; headteachers can treat staff as the professionals they are, with the professional knowledge they hold. In giving teachers the time to plan lessons and to have those live discussions as opposed to hours of marking, to be handed back the next day when the child has forgotten, are what head teachers should be doing. Micromanaging staff and their behaviours are more aligned to the concept of powerful and powerless, not the shared power that I, as headteacher, promote. To genuinely limit the impact of testing, headteachers must ensure there is an alternative which provides the required information, yet is also beneficial for all stakeholders

I am, like other headteachers, under pressure to look at academisation. While this research is not directly linked to the academisation process, it has clearly highlighted a style of leadership which does not fit readily with the local wider collectives. I have heard my

particular leadership style referred to as ethical leadership. It is how I choose to exercise the five qualities of power as identified by Cossa (1996).

However, a negative aspect of distributing power within the setting, is that staff then also have a level of responsibility. To be explicit, this is not a responsibility to deliver results, but a responsibility due to the relationships they have developed and the ecological system created. Whilst pupils were not impacted on by testing, I would conclude that the impact on staff was greater. On results day, the teacher of one class felt unable to register her class. She had received the results and whilst not upset with her pupils' effort, she believed that the increase in expectation to pass was unfair. The teacher did not want her class to see and misinterpret her disappointment with the system as disappointed with them. Regardless of how much the headteacher may feel they are shielding staff from other ecological systems and how well they feel they are doing in gatekeeping information, staff are human.

Creating a school with one collective, where the pupils' matter and permitting staff to be human, feel and even make mistakes should be promoted but monitored. The headteacher's role in ensuring that balance is difficult but essential. A shift either way can very easily puncture the equilibrium.

In reflecting on my role as headteacher and the ecosystem I have created, it was not until analysing the research setting, that I realised staff do not leave. Not because they do not possess the skill set, they absolutely do. You have to be able to genuinely teach in areas of high disadvantage - there is no escape or let up. They stay because they are valued, allowed to make a difference and more importantly to shape lives. The school's entire leadership team, including myself, were all newly qualified teachers starting within this setting. All, but one teacher has started their career there and has been moulded and shaped to become the amazing teachers they are. This supports the findings of Zuccollo, Dias, Jimenez and Braakmann (2023) who argue that good leadership reduces staff turnover.

The staff within this setting are not there to get data, to achieve good Ofsted outcomes or even to implement the many interventions and guidance documents required of schools. They trust that I, as headteacher, will allow them to do what is genuinely needed and close that bridge between the two ecological systems as required. The main long-standing priority within the school's improvement plan is staff well-being, followed closely by pupils' wellbeing. Data is not mentioned, it is a by-product of all the other actions. The workload of staff should be manageable, the support they need never questioned and they should be free to simply teach. They did not choose to be senior leaders and should not be exposed to the burdens, constraints and constant firefighting that headteachers face daily. The role of the headteacher and demonstrating their five qualities of power, is once again at play. If the teachers are not in the best place, they cannot implement good teaching and develop good teacher/pupil relationships. If pupils are not valued and allowed to also be themselves, then regardless of how good the teaching is, they are not ready to learn. Well-being of staff and pupils is what will make a school great. Not just because, in the end it will give you data, but because you are dealing with people and they need the tools to survive in wider contexts than just school.

The headteacher should be the traffic controller of the bridges withing ecological systems and make the decisions to close to traffic, as required; to slow the speed of travel and more importantly to make sure the right cars are on the bridge and travelling in the right direction. Headteachers also need to make sure that those at the end of the bridge (the staff), are also creating that additional safety barrier developed through a shared vision. Several bodies behind one another or holding hands side by side, are a harder border to penetrate. The headteacher needs to ensure that while they are 'traffic controlling' on the bridge, those left in the ecological system are still safe and fully understand their role in protecting those within their remit. The power does not live solely with the headteacher, it has to remain within the ecological system of the school setting.

Given, that assessment is also now taken to be testing, it would be apt to return to an earlier question in that 'are assessments working in the way intended?' This is not black and white and the 'intention' is very much open to the interpretation. Within this research, participants were not opposed to the tests, but how the tests were used and this view did not change over the course of the year. Therefore, if they are intended to be part of school life, this stance could be supported.

The political picture is complicated by the use of assessment results to create league tables. This is at a time when the UK government has created a market system for publicly provided and funded schools. As Torrance (1995) stated, this may be at the expense of 'real education.' Education will continue to be subject to considerable and at times heated debate because people want many different things from schools and have differing opinions on what education is. Even when there is agreement on goals, there will be disagreement on how to achieve them. As Sugrue (2008) stated education is so important and our children are so important, that educational issues will continue to engage people's feelings and commitments strongly. This was also reflected in the teacher becoming upset with test results. What needs to be ensured is that any assessments do not become central to school life. Schools do need to consider whether assessment measures what it claims to measure.

Gipps and Stobart (1995) exemplify this through their uncertainty of what is being tested and querying if a single test score on a reading test genuinely represents a pupil's attainment in all skills in reading. There are a number of ways to view validity, but for teachers the most important aspect is that the test closely matches the objective it is teaching. Another concern is that if the tests were given on a number of different occasions to the same child, would they achieve similarly? Furthermore, if it was marked by different people each time, would we still get the same score?



Perhaps the use of testing is too subjective and an agreement on its use is not one this research can answer. The impact of testing on pupils, this research has identified is minimal but it would be untrue to claim that the research has shown whether or not testing works as intended. Therefore, the wider language of the aim and the word 'assessment' warrants further discussion. As previously raised, assessment can be used too often to mean testing and so, therefore, to answer the question as to whether or not assessment is working as intended, we need to refer back to the 7 original reasons.

If we look back at the 7 reasons for assessment, this research would suggest that all 7 are being utilised within the research setting. I would conclude, therefore, that while my heart, as a head teacher, tells me assessment is not being used correctly nationally, the evidence within this research suggests that it is.

Alongside this are the HMI original 4 purposes of assessment and again, the research has found little evidence that these four purposes are not being utilised within the research setting and, therefore within this research. I can argue more weight is given to some aspects than others but I am unable to argue that they are not being implemented.

The research setting does not do things for the sake of doing them, and while initially I may have expected not to find 'assessment' being used correctly, this research concludes that it is. We must remember, however, that assessment is not a synonym for testing. If this notion is linked back to Williams' (2013) desired features of assessment, these are not reflected, nor are HMI or the seven original reasons in the statutory tests' pupils sit. They are reflected in the planning and day to day lived experiences of the classroom – in the informal assessment scenario. The original starting point, as a first-time headteacher, was to revert to the professional knowledge staff possess simply through doing their job, through the observations they do, the discussion with pupils on why, the tell me more about and the planning they write and tweak. Assessment in this guise, this research reveals, is working

as intended for both pupil and teacher. As Glaser (1980) suggests, assessment must be to support and not simply show past achievement.

Referring back to Davis (1998), I would agree that the rich knowledge, which is the goal of education, should not be assessed according to narrow and restricted procedures only capable of measuring procedural knowledge. Again, the definition of testing and assessment are pivotal to this.

What this research has highlighted is the vast grey area in relation to the understanding of the term assessment in comparison to testing and that each assessment may measure very different things and therefore 'does assessment measure what it claims to measure?' This will depend on several factors and criterion and the concepts involved are too complex to be answered simply. What does impact on this, however, is the power the headteacher has to interpret and manipulate this grey area.

A further area this research has highlighted is that headteachers can decide how they use testing and the impact they allow it to have within their school. The use of results by parents, governments and local authorities is beyond the control of the school. Focussing clearly on the impact of the school's own ecological system and the pressures it faces from other ecological systems can be controlled and not feared. Schools are able to filter and control impact through the actions of the headteacher and the staff. They are not powerless.

If schools take control of what they can and have a clear understanding of what is important to them and remain true to this, then testing does not need to hold any negative connotations. While this is very simple to write, the trust and determination to achieve it are far from simple. Creating a culture where staff do not fear results, a school where the needs of pupils are central and where pupils are free to make mistakes and be themselves, yet be ready for the next steps in their educational journey, is not achieved easily. The cultural capital of the school has to promote this eventuality.

While I feel able to conclude that this research has given a clear impact statement in regards to the effects of testing, I do not claim to have unequivocal proof of this. I simply argue that, within the context of this research, the systems in place within the research setting and the ethos of staff and pupils have given this specific, perhaps unexpected, new knowledge on primary school testing. Interview inquiry leads neither to a subjective reality or to an absolute objective knowledge. This research, while focussing upon testing, does not reflect the findings of McAllister's (2018) research around American teachers' perspectives and concerns with regard to their education system. His findings suggest that the policies and practises of progressive education had made students unable to comprehend subject matter and simply had a strategy of memorisation. That pupils had a lack of independent thought and instead mirrored their peers (and also cut corners to make short term gains), are issues that I may easily have encountered within the research setting had I not worked on the values and ethos of the school initially. This research highlights that the framework and ecological system in which a school finds itself are not unique. McAllister's (2018) research, although based in America, shows similar problems to the British education system. I would argue however, that this research has shown that those in schools, regardless of the external policies, can change the impact this has on pupils.

The knowledge I now have as a school leader will allow systems to continue and rather than an overhaul, subtle alterations can be made. The school environment is one for genuine learning, not just for passing tests.

In relation to the second research question. The role within the educational ecological system appears to be independent of external forces yet sits perfectly within the Russian doll scenario Bronfenbrenner (1968) offers. The actions I have taken, alongside staff, have created a microsystem which meets the demands of external ecological systems yet remains self-sufficient. Schools can and have, limited the impact from other external systems. This

setting has not given in to the pressures of teaching to the tests, focussing on core subjects or even on teachers having to get results. They have remained comfortably within their microsystems and allowed other systems to trickle in, rather than create an overwhelming waterfall and risk drowning. They have measures in place to control the water levels. These measures, in hindsight, are the headteacher and the micro system they run.

Human development is the process through which the growing person acquires a more extended differentiated, and valid conception of the ecological environment, and becomes motivated and able to engage in activities that reveal the properties of, sustain, or restructure that environment at levels of similar or greater complexity in form and content (Bronfenbrenner, 1979 p.27).

The notion raised by Bronfenbrenner (1979) can be identified within the research setting, as the setting has adapted to the changes and requirements set by the exosystem. The competition aspect, threaded throughout this research, is simply not part of this research setting. Referring back to Sumner (1987, p.29)

In my own experience a substantial proportion of children have said that tests were interesting and enjoyable to do. When the results are promised beforehand, the motivation of pupils can be very high, especially when the purpose is explained fully and the conditions under which the tests are done are understood.

I would not say I agree that testing is enjoyable or that results should ever be promised, but I can conclude that tests are accepted within this research setting.

Perhaps, however, this is the statement that shows testing does not need to be negative and whilst in the minority of literature, positive reactions to testing do exist. Testing has not had a negative impact within this research setting because the research setting would not allow it to. The school, the pupils, staff, ethos etc can never be completely replicated in another school but what can be changed, is the attitude as to what really does matter and how you address this. While the setting did not promise students any results, they did promise them that their best is good enough. Just knowing that testing can be a friend and not an enemy

is a good starting point for this. It would be wise to refer back to one of the early comments within this research and Eggleston's (1991) thoughts around social phenomena and testing. Testing and the impacts of it could be argued to be just that. The impact is not from the label given as such, but by the social reaction to it. The research setting controlled the mesosystem's reaction to testing. It manipulated, interpreted and regulated its control to fit within the values of its own ecological system. This ensured the school's focus was on effort and not attainment.

Despite a review of the policies surrounding testing and assessment appearing to be more negative than positive, the policies implemented by successive governments and reviewed within the literature, have allowed the research setting to create a system where the testing has not impacted and for this reason, they have achieved their desired outcomes. The power of headteachers to regulate and interpret policies has been successfully exerted. Initially, and perhaps subconsciously, this research had hoped to impact upon government policy and to highlight the errors and hope for change. However, it is now clear it is not about impacting on external policies, but about working on the interpretation of these. It is about focussing on the status that is given to policies and the level of impact permitted, as this ecological study has demonstrated. Even when something appears to be at odds with our own personal beliefs (as I originally felt), headteachers can impact upon outcomes through their implementation of policies, their power and their influence. As stated within the introduction, government policies have to be implemented. Trying to change them and impact on that ecological system is perhaps, currently, outside of a headteacher's control. Not many headteachers have the stamina for a full revolt, but this research has demonstrated that it is actually unnecessary. Changing policies, testing, pressures etc are not the issue. The issue is the significance and kudos attached to testing and as this research demonstrates, headteachers control that. That is not to say that this research cannot be used to inform educational policy, the 6 key findings would be beneficial to both

policy makers and school leaders. Within this setting the role of the headteacher has remained firmly within their own ecological system and focussed only upon the things that they are able to control. They have focussed on one small subsection of the ecological system, not the entirety. Perhaps this the key, to know what you can change, through the control you have and not the control you want. In terms of understanding a headteachers role within Bronfenbrenner's ecological systems, they are the agent of their own ecological system. Change starts with them, in this small yet significant microsystem.

Bronfenbrenner (1968) remained an optimist throughout his long career. In reflection of this, as educational leaders, there is no need to despair. Leaders can change things and their actions do matter. To hold school leaders to account for the decisions made in relation to the maximisation of results is the main rationale within the many policy documents published in education and reviewed within this research. It is here that school leaders perhaps need to firstly acknowledge that they do possess a significant amount of power. They then need to decide how they exercise this power. Local authorities and governments hold school leaders to account but, as headteachers, we also create accountability. This is done through the systems we create, the power we share, the views we hold and the values we instill.

Testing did not impact negatively within this research setting, due to the ecological system that it sits within. This ecological system has been created by the headteacher and the choices they have made and the actions they have taken. While the research findings may not reflect those commonly stated, particularly in the media. It is a finding that offers some hope to school leaders and teachers. Testing is not necessarily a negative experience and headteachers have a vital role in limiting the impact within their own ecological ecosystem.

## Appendix One:

### Pilot Interview Questions

- *What do you think of the y2 tests?*

Honestly? I'm not sure what use they are. They inform the teacher assessment but don't tell you anything new. Sometimes one or two might surprise you and do really well but it's usually luck rather than understanding. They take ages to do also. I don't mean we give them extra time – don't panic. We have to rearrange the tables; we need extra staff to help read things to the pupils and because there were two of 2 last year, we had to keep them separate until they had all done it.

*How many tests are there?*

*Too many! There's reading and maths you have to do. We also do the GPS one but that's optional, after they leaked it the other year. It gives us a baseline though so we still do it, as we teach it. Actually, I nearly forgot. Some of them will also need to re-sit the phonic test if they failed it in year 1. That's later on though.*

- *How do you think pupils respond to the tests? Can you illustrate/give examples?*

It doesn't matter that you say 'just try your best and don't worry' it's still so different for them. They do amazingly well though and all have a go. I think the reading one is the worst for them as we can't help them at all, we can't read anything to them. We tell them any other time – just ask if your stuck and then we make them sit there with something they can't do and don't help them. They just look so helpless.

- *What do you feel we need to do as a school to ensure pupils are equipped for the test? (or similar dependant on the answer to previous question)*

I don't know. Not do them? Well do them but not tell them it's a test! There so little and to sit for so long without speaking or us being able to help them must be awful for them. I'm not sure we can do much can we?

- *Do you feel that the tests impact on the attitudes of pupils? Can you illustrate/give examples?*

They get worried about any test. B cries on a Friday when it's spelling test day. I don't think it's the tests so much, they can do them. It's the fact they're different. They're sat quietly and my class is never quiet! It's awful for them. It must be strange to have to sit there and not be how they usually are in class.

- *If I walked in your classroom for the first time today, what would I See? (this may show a change in attitude when asked after SATs)*

Pupils loving school. A noisy but purposeful class. Pupils engaged with lessons and having a go. Watch you come in now and they just freeze!

## Appendix 2:

### Final Agreed Areas for Discussion.

Can you tell me how your tables are organised in your classroom?

- Carpet area
- Ability groups/subject
- Vulnerable groups
- Gender

Why did you group them that way?

- Ability
- Subject
- Classroom constraints i.e. floor space

How would you describe your pupils' current attitude towards their learning?

- Mindset (FM/GM)
- Behaviour
- Resilience
- Particular subjects/times
- Groups of pupils (EAL/SEN/FSM/MA)

How do you engage pupils with your lessons?

- Hooks
- Practical lessons
- topics

How do your pupils respond to any tests (spellings, maths)?

- Groups
- Type of test
- Time or day



## Final Interview Question Example

### **Can you tell me how your tables are organised in your classroom?**

*We had behaviour training from Jason Bangbala and he suggested mixed ability groups in this shape. Which is fine but it doesn't work for guided reading so we have ability tables for that.*

*Why does he suggest mixed ability groups and why does it not work for guided reading?*

He says that the lower attaining pupils will be pulled on by the rest of the group and the higher achieving will show their understanding by helping the other group members. Otherwise, you have the lower ability all together being stuck together and nobody to raise the bar for them. This shape means the teacher can work with everyone. It does work so far but in guided reading when it's not a free choice week or whole class reading week, then the group need the same book so we have sets for guided reading.

*And you have a carpet area still?*

*Yes, we use that for whole class teaching. We tend to start the lessons there before moving to their tables.*

### **Why did you group them that way?**

*As I said it is based on the behaviour training, we had in July. It's mixed ability so I used the previous year attainment grade to put each child on a table. I also looked at the other vulnerable groups such as EAL and SEN. Each groups have a mixture of pupils. One boy has to sit on this table though as his wheelchair is easier for him there and his 1:1 can fit in.*

*Are the groups fixed now or will you change them again?*

*No, I have moved them a few times as originally, they were based on results and not knowing the pupils. Now I know which pupils can't sit together and which work well. I think it will stay like this but obviously as the pupil's progress and develop they may need to move.*

### **How would you describe your pupils' current attitude towards their learning?**

*Oh gosh that's a hard question! Generally, they enjoy learning or I think they do. We keep getting attendance rewards so they are coming to school. They do the energisers when they first arrive and they're happy to show their parents what they're doing. Most get onto the mastery challenges in maths and we have loads of work in our books.*

*What would you say behaviour is like?*

*Good. We have nobody that has moved to red or black, only the time outs.*

*Would you say the class are resilient?*

*Most of them, yeah, they give it another go or they keep going but I would say some don't. A handful give up and say they can't do it.*

**Are these pupils in any particular group?**

*Do you mean like SEN or something? No, I wouldn't say so. It isn't always my bottoms.*

**Are there any groups of pupils that you would say demonstrate strong resilience?**

*Hmm, I've not really watched for groups but the EAL pupils generally seem unfazed and keep going but that could be because the work more specific and they get support.*

**So, do you think pupils are more resilient when they have an adult sat with them?**

*I wouldn't say more resilient but they don't sit there doing nothing or saying they can't do it. They ask for help more though.*

**When would you say the pupils are most engaged in their learning?**

*Probably when it's more hands on. They love the hooks and science now we don't have books. Practical makes a difference as they are finding out themselves and I can get to grips with their thinking more. The current topic is quite boy friendly too.*

**Would you say the day or time impacts on their engagement?**

*No not really. Mornings tend to be English and maths and the afternoon topic and things so that's what they're used to. We still doing maths and writing in topic so it's not as though they don't do work on an afternoon. They like all days I think but I'm out Tuesday afternoon, so you would have to ask Janet.*

**What would you say the mindset of the class is generally?**

*Well, we have been doing growth mindset for a few years so I'd like to think we aren't wasting our time. Most pupils have a growth mindset but they need reminding sometimes!*

**How do you engage pupils with your lessons?**

*We have hooks, practical lessons, the outdoor area, we go with the interests of the class. We do loads, you don't need me to name them all do you?*

**How do your pupils respond to any tests (spellings, maths)?**

*Well, we have a spelling test each week and some just cry when they get their scores. Some aren't bothered but we'll keep trying with it. The CLIC maths test isn't as bad and we don't get as many tears, even though just as many do as badly.*

**Why do you think that is?**

*I have no idea! I suppose they did CLIC last year so are used to it and spellings are new. Spellings go home also, so I get parents asking the scores more than the pupils. I just send the scores home weekly now as it's just easier.*

**Do you call them tests?**

Hmm we definitely have a spelling test but I can't remember if I say CLIC test or CLIC. Do you think it makes a difference? Do you want me to check what I call it with Rachel?

**No. It was just a question. Please don't worry!**

**Do you think the fact parents get the spelling scores impacts on those pupils that you mention cry?**

*I don't know. They cry when they get the score and that's before the parents get the slip. I suppose the pupils can always just not take the slip home if they don't want their parents to know though, can't they?*

**Is there anything else you would like to add or comment on?**

*No. There probably will be later though!*

### Appendix Three:

#### Example Monthly Thoughts from Reflective Diaries

Analysis Theme: Behaviour and Attitude The Learning Environment Labels

October	1	2	3	4
WB 9.10.17	First SAT test, using previous papers. Time warnings given (half way) SSC realised not half way and put head in hands. Given gentle reassurance and carried on. Lower score than expected. Third test and all reassured by teacher happy with effort. RS asked what would happen if scored zero in real test. Reassured			
WB 16.10.17				
WB 23.10.17				
Thoughts?	Attitudes towards SATs, worried? Not prepared (timewise)-will they get trained?			

November	1	2	3	4
WB 6.11.17	<p>Difference in mindset. Girl TA and boy OJ. Long division – girl got in a muddle first time and didn't like it. Revisited and loved it. Confidence grew. Boy didn't want to move from his method, which didn't work.</p> <p>Reading style questions – KE. Always gives 100% but overcome with frustration when 3 mark question she answered (vey well) why ,when it was how. Upset and tearful.</p>	<p>Writing (HA girls) – fantastic imagination, only using text map as guide on structure. Clear growth in confidence. Perseverance when up levelling punctuation. Reading test – girls showed positive attitude, eager to score well. Aware of marks available for each question. Boys gave in easily. Sulked when faced with challenge.</p>	<p>Long division – TA loved it at first, then errors but stuck with it and loved it again. OJ Pushed against it and continued to say 'no' when support offered. Majority of class enjoyed a method they found accessible. Arithmetic test – SSC expressed desire to do well to tell her dad. SH achieved 2/7 and was disappointed.</p>	Email!
WB 13.11.17	<p>Art (obs of cola bottle). Lots of I can't draw. MR reluctant to start and lots of reassurance needed. Pleased with end product.</p> <p>Maths Y6 - JW refused to ask for help. Producing</p>	<p>Whole class positive approach to Andy Warhol despite claims they can't draw. High quality sketches produced and sense of pride evident.</p> <p>Highest spell score week. Writing - Slow to put ideas forward, lack of</p>	<p>Spelling scores gd this week. Applying skills taught to test.</p> <p>Y6 girls eager to know house captain results. Reading – chn not consistently applying skills taught.</p>	Email!

	less work than others and making more errors.	understanding. Demonstrated by teacher and brilliant engagement.		
WB 20.11.17		Maths y6- able boys taking short cuts in methods, resulting in stroppy behaviour when pointed out. Perseverance by most with prime numbers. All expressed they found it difficult but displayed determination. Shift in boys behaviour when covered by supply teacher. Negative approach to lessons and disruptive. Vulnerable pupil in meltdown all day due to supply.	EL- sudden change in handwriting, not prompted but a big improvement. Maths y6 – Prime numbers. Chn generally didn't have patience to work logically through the numbers. Arithmetic tests scores improving. GD pupils getting amber in energisers so spoken to. Adult support not in today – grp productivity and behaviour not as good.	
WB 27.11.17		CT mood swings, very up and down. Negative mindset and giving up before started. Sat crying and would not accept help or support. Frequently stated he has mental health issues. Work refusal for whole day, rude to adults	Writing – TB very reluctant to check and self edit, instead chose to play with a dictionary. DS also similar attitude (both U chn) TA spent lots of time on cold write – gave up her break voluntary to do this.	Maths Y5-behind on teaching strategy as don't know x tables. Enjoying xtables games. 2 lots of learn its. Reading- listened to readers. Fluency improving. Noticeable in Lower attainers. Using

		<p>but other children accepting and got on.</p> <p>Writing – TW, EM, DS,TB,BB not self-editing accurately or using time productively. AS, TA keen to make improvements. Very eager to edit and improve work. Class showing more confidence using subordinating/coordinating conjunctions. TW, EM heavily supported in non-independent tasks – though didn't use any of the guidance.</p>	<p>Maths – MC proudly announced green on energiser all week. This prompted other chn to comment on positives. Class displaying good attitude to maths and working hard. TA shared a tale from older sibling in high schl – no point in primary schl. Other chn annoyed and dismissed the comment.</p>	<p>schl library and 7 achieved first reward. Writing- cold write setting description. Grammar work needed. Text maps. Enjoyed and good job. Vocab for changing mood. CT good at magpieing.</p>
Thoughts?	<p>Attitudes to errors?</p> <p>Can't do attitude more 'popular' in foundation subjects?</p>	<p>B/G differences in reading?</p> <p>MA boys – short cutting, why?</p>	<p>Attitude to errors – is believe secure in a method. Possibly boys? Needing reassurance – different attitude when no support Proud when achieving.</p>	<p>Attitude to maths is behind?</p>

Jan	AP	KM	EM	JW
WB 8.1.18	<p>Police sketch – BD more academic than creative. Cried and said can't do</p>	<p>Girls appear to have returned with appositive attitude Keen to find out</p>	<p>Wow day went well, chn researched new topic at home.</p>	<p>Reading – test again (same as Sept). 6 passed, 2 in Sept. 3 more within 2</p>

	<p>art. Spoke to him about perseverance and determination. In end felt proud not given up.</p>	<p>when they would be doing more learning towards SATs and the topics this term.</p>	<p>Writing- CT refused which escalated rapidly. Others enjoyed letter of complaint about gifts received.</p> <p>Reading – Mock reading test. Most scores improved and chn didn't appear to mind too much. Some y6 not listening to time prompts and advice - not completing the paper.</p> <p>Maths- worked well until Weds then forgot everything by Friday.</p>	<p>marks, 2 in Sept. Most improved by 4 marks. KE worry, lost 3 marks since Sept. Mum concerned wasn't being challenged. Raised in Nov when KE struggled with GD work.</p> <p>Writing – Cold write. Writing more but varies in quality. Hook (courtroom) and came up with good arguments. TB, KE,ZO answered questions well in character. Class worked well in pairs. Maths y5 –assessment completed. Most 4/5/6 from 22. Equivalent fractions – struggled to draw own but got concept. Will give diagrams next week.</p>
WB 15.1.18	<p>Alcatraz sketch – BD tackled it more confidently. Said he enjoyed it.</p> <p>SH checks behaviour chart each morning. Became</p>	<p>Attitudes to learning sulkier. More noticeable in year 6 – SH, EM,RS,CT,TB</p>	<p>Maths – given method for multiplying decimals. Most took on board but showed clear gaps in understanding that needed addressing. OJ</p>	<p>Reading – OJ remembered pioneer from text in Oct. R. reading went really well. Gd understanding of quite tricky NF text. Some struggled with summary,</p>



	<p>clear eats into her time and feeling rushed. As a consequence, energiser scores suffering. Spoke to her about resigning until after SATs. She agreed. KS had several days off. Subsequent lessons on x decimals problematic. Stayed for intervention – understood and felt confident at end. OJ- reluctant to show working out, even when asked. Gave reasons why important (listed them) and accepted this</p>		<p>refused to show working out as didn't see the value in it. More able chn worked through very quickly. SH not consistent enough in energiser. Thinking of giving up monitor job for more time. Arithmetic test- every high scores. Writing- informal/formal. Found informal much easier as didn't have formal vocab. RS very much disliked checking his work for errors and developed negative attitude. Topic- Alcatraz outline. Step by step but struggles with instructions.</p>	<p>worked on this. Majority got over 4/7 on ind based on cold reading. Assessment – all but 2 at exp, 2 GD. KE very upset as one mark off passing/ this has happened before. Looked at her conciseness and will get there. Will look at 3 mark questions next week. Writing- Features of formality. Fewer clerical errors and spelling errors than last term. BR used word of the day. Sentence structure continuing to improve. BD upset after not knowing what to write. Recapped using model. Maths y5 – diagrams helping. Showed way of changing fractions using process. Some struggled so lots of modelling.</p>
--	--	--	---	--

				<p>LH, KG,FM good understanding of a tricky GD challenge.</p> <p>Y6 – struggled with Thurs energiser and 2 step prob.</p>
WB 22.1.18	<p>TA – test scores generally expected. Conveyed to staff and friends feels panicky when given tests, especially reading. She explained worried about reading as knows an adult can't read to her. Listed lots of anxieties.</p> <p>SH – energisers better after resigning from job. Also doing online SAT tests at home.</p> <p>Class- lots of work on division. RS – do we have to show working or can we guess!</p> <p>CGT – English. Slower than rest. Lots of copying either model text or other pupil ideas. To monitor if confidence or laziness.</p>	<p>English – DS, AS,TB,EL not stretching themselves and lack of positive attitude to learning. No resilience and give up easily.</p>	<p>Topic – Alcatraz shading. RS, TB did very well and demonstrated perseverance.</p> <p>Reading – Avoiding 3 mark questions. DS told another pupil wasn't coming on Thursdays as didn't like 3-mark questions on that day's reading task. All chn reminded will help them and will develop their confidence. Chn now saying test style questions not a test!</p> <p>Energisers – given tough one this week and told if did well would get dibble. Good number of chn rose to challenge and all correct. Arithmetic test now 12 with 14 marks so</p>	<p>Reading – Gd engagement on NF text on cholera in London 1800. BR asked about more tests. Explained format of sessions. BR just wants to listen to class story as he said not good at questions. Looked at 3-mark questions and like model. Lots of 3 marks on ind task. Pupils wanted to mark test. All but 1 got exp, 1 GD. Most getting 2/3 on 3-mark questions.</p> <p>Writing – many struggled to remain formal. Colloquialisms used. Gd additional lesson on maintaining formality. Watched video on subjunctive to help</p>

	<p>Worked with CGT and will continue to monitor. Reading SAT questions. KS got bogged down on 3 mark questions and little teary. Left and moved on. To monitor for actual test.</p>		<p>double. Chn coped well and good number got full marks. Maths – partner work and enjoyed the problem – lots of discussions. Writing- independent formal piece. AS (shd be GD) struggled to produce anything that made sense.</p>	<p>understanding. Now using subjunctive. Maths Y5 – brilliant understanding of improper fractions after visual approach. AB, CWB both had misconceptions so 1:1. Successful after. All enjoying loop game. Reduced original time by third.</p>
WB 29.1.18	<p>KS in maths went blank and usually confident in this subject. Beginning to stress. Had tears and signs of frustration. Art- Few 'I can't' before even knew task. Discussed this and all prepared to have a go.</p>	<p>SATs practice- TA, CT,RS,AS stressed before and in sessions. TA shows visually signs such as tapping, pulling faces and the pace slows down leaving lots of unanswered questions. Handwriting improvements for BB, PG,TW.</p>	<p>Topic – graffiti art. Enthusiastic but when drafting own were negative. MR vocal in frustration. CT refused. EM did really well and proud of efforts. Writing – cold write. CT completed no problems. Redrafting, y5 needed lots of guidance. Reading – Still focusing on 3 mark questions. RS (shd be GD) very defeatist and reluctant to have a go. AS, TB slow and not finishing.</p>	<p>Reading – cont with 3-mark questions. Looked at finding evidence then making point. After modelling, all pupils got the 3-mark question. Most passed weekly assessment. Maths y5 – features of shapes, enjoyed pinboard but struggled to sort diagrams. Writing – Redrafts of letters very accurate and good level of formality. Cold writes showing good structure but not much</p>

			<p>MR, SH giving them a good go.</p> <p>Maths – new partners for problem solving. TW decided to work independently of partner OJ, even though good friends.</p> <p>TN finding step up in maths hard. Appears to accept he gets them wrong but happy when gets one correct.</p> <p>Multiplying decimals. Advised chd not to use column method but vocal in disagreement. Allowed them to try their way. Quickly realised too tricky and reverted to my method.</p>	<p>depth. Enjoyed hook and able to recount events from new text.</p>
Thoughts?	<p>Worried over pressure of SATs?</p> <p>Some showing perseverance – why?</p> <p>Determination to be better – giving up job</p>	<p>LA/MA attitudes to learning.</p>	<p>Attitude to reading?</p> <p>Pupils with low self-esteem struggling –CT?</p> <p>DS and reading.</p> <p>Attitude to errors – is believe secure in a</p>	<p>Reading improving- linked to own attitude and achieving reading awards?</p> <p>KE reading? Other pupil starting to worry in</p>

			<p>method. Possibly boys/CT          –RS gives up in core not          foundation.          Determination to be          better – giving up job          New methods – will they          now see benefit?</p>	<p>reading – not reading but          what can write?          First upset in writing –          why? So far least          threatening as no test?          Maths Y5– not getting          upset when don't          understand.</p>
--	--	--	--	---

## Appendix 4:

### Participant Information Sheet

#### Title of the Study

An examination of teacher perspectives of assessment and testing regimes within primary schools.

#### Introduction

I would like to invite you to participate in a research study as part of a EdD. Before you decide whether you would like to participate, It is vitally important that you fully understand why the research is being conducted and what it will involve. Please take time to read the following information carefully and discuss with your colleagues or head teacher if you wish. Should you wish to discuss anything that is not clear or you require any further information, then contact details are provided at the end of this sheet. Please take your time to decide whether or not to participate in this research.

#### Background and purpose of the study

There have been vast amounts of research into the necessity and validity of formative assessment within primary schools. This research is not therefore, focussing on the process of SAT assessments and the reasons for them but instead on the impact of testing and assessment from the perspectives of those enforcing them – the teachers. With a new National Curriculum, removal of levels and new SATs framework recently introduced (2015), research on the impact of this system through the eyes of those enforcing the changes is vital, if the teaching profession are to influence policy development and also future practice within the classroom.

The teaching profession are one that constantly seeks to improve its own practice and this research will aid future practitioners and senior leaders understand the impact of assessment and testing of both the pupils it serves and the staff fulfilling the statutory requirements. Therefore, the aim of this research is to explore the impact of assessment and testing in primary school from a teacher perspective; to look at the impact on the teachers themselves **and** their understanding of how the results impact on the students.

#### Am I a suitable participant for this study?

I am seeking teachers and support staff within year 6. There is no requirement to have been teaching for a set period of time, or working in a supporting role for a set time period. Participants are welcome from newly qualified teachers and experienced teachers. However, trainee teachers are not suitable participants due to the limited experience of formative assessments (SATs) they may have. Participants are not expected to have taught for any given length of time in year 6 but must be teaching within this year group during the 2017/18 academic year.

#### Do I have to take part?

It is your decision whether or not to take part. Should you choose to participate, then this will need to be confirmed on the consent form enclosed alongside the information sheet. If you decide to participate you are still free to withdraw at any time and without given reason. Should you not wish to participate, then this can also be conformed on the consent form.

#### What will happen if I agree to take part?

I will arrange a convenient time to introduce myself and answer any questions you may have following consenting to participate.

#### What will I have to do?

Participants will be interviewed to answer questions regarding their experiences of assessment and testing within this academic year. The questions will be semi-structured, in that base questions will be asked and any further clarification will be sought. The interviews will be face to face and answers given will be recorded.

Participants will not have to collect any data or evidence to support their views in regards to the questions.

#### How often will I have to participate?

Participants will be interviewed twice. Once at the start of the school year and towards the end of the school year. This will allow for the teacher perspectives on assessment and testing generally to be analysed and then also for the perspectives on any impact the results may have had on the pupils. At each interview you will agree that what has recorded is accurate.

You will also be required to keep a reflective diary. This will be a very brief description of your observations (ideally weekly) and thoughts and should take no more than 15-20 minutes per week. It may note changes towards learning, successes or failures and any general observations such as pupil anxiety.

All time to complete any such work (both interviews and reflective diary) will be covered in order to not impact on your own time and will also be arranged to have the least impact on the pupils.

#### What are the possible benefits of taking part in this study?

This research is being undertaken to try and understand the impact that assessment and testing are having on the teaching profession. It will give teacher perspectives on what has become a major part of not just day to day teaching but also government policy. League tables, which are based on published results, inform future policies. Yet very little research has taken place on the professionals trying to achieve these results or the pupils who are given them. Once results are published it is the teaching profession that are left accountable and the pupils that live with their results. In participating in this research, it is an opportunity for teachers to be heard and therefore inform future practice, even if only within their own setting.

On a personal note, it may inform your own teaching style. The research may highlight patterns that until an analysis is undertaken do not become known. For example: is there a shift in learning attitudes at set times of the year? Are certain groups more compliant than others? Is resilience a whole class issue or just particular groups? A copy of patterns identified, alongside the reflective diaries, will be returned to you at the end of the research.

#### Will taking part in this study be confidential?

All information collected will be kept strictly confidential, other than to me and my university supervisors. School names and participant names will not be used in the research. A copy of your signed consent form will be kept, alongside interview responses and kept securely in a

locked filing cabinet and a password protected computer. Following the completion of the research all paper records will be shredded and electronic files deleted.

What will happen to the results?

Once the research is completed, it will be submitted as part of a doctoral thesis. The study findings will be clarified with participants before submission, to ensure an accurate representation of teacher perspectives is given. The study findings will also be made available to interested parties.

Contact details

Ethics Co-ordinator

Jean Laight  
School of Education and Childhood  
Leeds Beckett University  
Carnegie Hal  
Headingley Campus  
LS6 3QS  
[J.Laight@leedsbeckett.ac.uk](mailto:J.Laight@leedsbeckett.ac.uk)  
0113 812 6224

Director of Studies

Dr Nick Sutcliffe  
School of Education and Childhood  
Leeds Beckett University  
Carnegie Hal  
Headingley Campus  
LS6 3QS  
[n.sutcliffe@leedsbeckett.ac.uk](mailto:n.sutcliffe@leedsbeckett.ac.uk)  
0113 812 1771

EdD Supervisor

Sarah Swann  
School of Education and Childhood  
Leeds Beckett University  
Carnegie Hal  
Headingley Campus  
LS6 3QS  
[s.swann@leedsbeckett.ac.uk](mailto:s.swann@leedsbeckett.ac.uk)  
0113 812 4629

My details removed to ensure confidentiality for the participants.



## Appendix 5

### Example of Participant Clarification Notes

<u>WB 27.11.17</u> <u>Thoughts</u>	Attitudes to errors? Can't do attitude more 'popular' in foundation subjects?	B/G differences in reading? MA boys – short cutting, why?	Attitude to errors – if believe secure in a method. Poss boys? Needing reassurance – different attitude when no support Proud when achieving.	Attitude to maths is behind?
Clarification	Would you agree that pupils are less resilient in foundation subjects?	Do you think there is a difference between the attitudes towards reading in boys and girls? Do you think there are nay attitude differences between the boys and girls?	Do you think there is a different attitude towards working between the boys and girls?	Is there a different attitude towards maths by the pupils compared to other subjects?
Discussion notes	Depends on the subject. Art yes but not really topic or PE. It could just be the unit of work. I wouldn't say it was a gender issue more who they are sat with issue.	Not really, it's more the actual text for reading. Some boys hate every book I offer but so does K and she's female.	Not generally but it's quite top heavy and I have a table just of boys.	It's more black and white I suppose so those who are unsure take longer to get there. Nobody hates it but it's the area we've worked on as a school so it's hard to tell. Check with me next time when we move topics!

## Bibliography

- Ainscow, M., Dyson, A. and Goldrick, S. (2011) *Developing Equitable Education Systems*. New York, NY: Routledge.
- Arthur, J., Waring, M., Coe, M. and Hedges L. (eds)(2012) *Research Methods and Methodologies in Education*. London: Sage.
- Baker, M. (2010) <https://www.bbc.co.uk/news/10094248> {Internet} (Accessed 5.7.2029)
- Bleeker, M., Sherman, J.F. and Nedelkopoulou, E. (eds.) (2015) *Performance and Phenomenology: Traditions and Transformations*. United Kingdom: Routledge.
- Bourdieu, P. (1977) *Cultural Reproduction and Social Reproduction* in Karabel, J. and Halsey, A (eds) *Power and Ideology in Education*. New York; Oxford University Press.
- Bordieu, P. (1986) *The forms of Capital* in Robinson, J (eds) *Handbook of Theory and Research for the Sociology of Education*. New York; Greenwood press.
- Bourdieu, P. and Wacquant, L. (1992) *An Invitation to Reflective Sociology*. Chicago: Chicago University Press.
- Bourdieu, P. (1993) *Socilology in Question*. London: Sage.
- Brenda, A. (1994) *In Defence of Choice in Education*. London: Halsted
- Bridges, D. and Smith, R. (eds) (2007) *Philosophy, methodology and Educational Research*. Oxford: Blackwell Publishing.
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*, 22(6), pp 723–742
- Bronfenbrenner, U. (Ed.). (2005). *Making human beings human: Bioecological perspectives on human development*. Sage Publications Ltd.
- Claxton, G. (1998) *Hare brain, Tortoise Mind: Why Intelligence Increases When You Think Less*. London: Fourth Estate.
- Christensen, G. (2023). Three Concepts of Power: Faucault, Bourdieu, and Habermas. *Power and Education*, 0, pp1-14.
- Crotty, M.J. (1998) *The Foundations of Social Research: Meaning and Perspective in the Research Process*. Los Angeles, CA: Sage Publications.
- Denzin, N. K. & Y. S. Lincoln (Eds.) *Handbook of Qualitative Research*. Thousand Oaks, CA:Sage.
- Earl, L and Katz, S. (2008) *Getting to the Core of Learning. Using Assessment for Self-Monitoring and Self-Regulation* in Swaffield, S. (ed) (2008) *Unlocking Assessment. Understanding for Reflection and Application*. Oxon: Routledge. pp90-99
- Fullen, M. (2001) *Leading in a Culture of Change*. San Francisco: Jossey-Bass.

- Fullen, M. (2008) *The Six Secrets of Change*. San Francisco: Jossey-Bass.
- Fullen, M. (2011) *Change Leaders: Learning to do What Matters Most*. San Francisco: Jossey-Bass
- .
- Gipps, C.V. and Murphy, P. (1994) *A Fair Test? Assessment, Achievement and Equity*. Philadelphia: McGraw-Hill Companies,
- Gomm, R., Hammersley, M. and Foster, P. (eds.) (2000) *Case Study Method: Key Issues, Key Texts*. Thousand Oaks, CA: SAGE Publications.
- Greenfield, T. (2002) *Research Methods for Postgraduates*. 2nd edn. Arnold: London.
- Halstead, M. (ed) (1994) *Parental Choice and Educational Principles, Policy and Practice*. London: Kogan Page.
- Hayton, J. (2015) *PhD: An Uncommon Guide to Research, Writing & PhD life*. United Kingdom: James Hayton PhD.
- Hirsch, E. D. (2016) *Why Knowledge Matters*. Cambridge, Mass.: Harvard Education Press.
- Howard, S., Dryden, J and Johnson, B. (1999) *Childhood Resilience: Review and Critique of Literature*. Oxford Review of Education. 25:3. Pp307-321.
- Richardson, V. (2003) *Constructivist Pedagogy* Teachers College Record 105(9 pp1623-1640.
- Rugg, G. and Petre, M. (2006) *A Gentle Guide to Research Methods*. Maidenhead, England: McGraw Hill/Open University Press.
- Sharron, H. and Coulter, M. (1996) *Changing Children's Minds: Feuerstein's Revolution in the Teaching of Intelligence*. 3rd edn. Birmingham: Sharron Publishing Co.
- Simmons, H. (2009) *Case Study Research in Practice*. London: Sage
- Sokolowski, R. (1999) *Introduction to Phenomenology*. 7th edn. Cambridge, UK: Cambridge University Press.
- Tahira, M., Yousuf, M. and Saboor, A. (2021) Role of Headteachers as an Instructional Leader. *Global Education Studies Review*, 6(2), pp114-123.
- Tajfel, H. (1978) *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relationships*. London: Academic Press.
- Tajfel, H. (ed) (1982) *Social Identity and Intergroup relations*. Cambridge: Cambridge University Press.
- Winthrop, R., Morris, E. and Qargha G (2023) *Revisiting Education Systems Transformation: Why Understanding the 5 forms of Power is Essential*. Washington: Brookings Institution.
- Zuccollo, J., Dias, J., Jimenez, E. and Braakmaan, N., (2023) *The Influence of Headteachers on Their Schools*. UK: Education Policy Institute

## Reference List

- Almond, P.C. (1994) *Heaven and Hell in Enlightenment England*. Cambridge: Cambridge University Press.
- Ball, S (1981) *Beachside comprehensive. A Case Study of Secondary Schooling*. Cambridge University Press.
- Ball, S. (1994) *Education Reform: A Critical and Post-structural Approach*. Buckingham: Open University Press.
- Ball, S. (2006) *Education, Policy and Social Class. The Selected Works of Stephen Ball*. Oxon: Routledge.
- Ball, S. (2012) *New Policy Networks and the Neo-liberal Imaginary*. Oxon: Routledge.
- Ball, S. (2013) *The Education Debate. 2<sup>nd</sup> edn*. Bristol: Policy press.
- Bassey, M. (2008) *Case Study Research in Educational Settings*. Berkshire: Open University Press.
- Beal, A. (2018). <https://www.wired.com/story/growth-mindset-education-psychological-theory-children-mirage/> [Internet] (accessed 31.8.2020).
- Biesta, G.J.J. and Professor Gert J. J. Biesta (2014) *The Beautiful Risk of Education*. Boulder: Paradigm Publishers.
- Black, P.J. (1998) *Testing, Friend or Foe? The Theory and Practice of Assessment and Testing*. Washington: Routledge Falmer.
- Bourdieu, P. (1990) *In Other Words: Essays Towards a Reflective Society*. Cambridge: Stanford University Press.
- Bradbury, A. (2019) *Pressure, Anxiety and Collateral Damage. The Headteacher's Verdict On SATs*. London: UCL.
- Brawn, V and Clarke, V. (2016). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3 77-101
- Broadfoot, P. (1995) *Performance Assessment in Perspective: International trends and Current English Experience* in Torrance, H. (ed) (1995) *Evaluating Authentic Assessment*. Buckingham: Open University Press. Pp9-34
- Bronfenbrenner, U. (1979) *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge: Harvard University Press.
- Brown, S (1991) *Assessment a Changing Practice* in Horton, T. (ed.) (1991) *Assessment Debates (curriculum and learning)*. London: Hodder & Stoughton in association with the Open University. Pp5-10.
- Burkhardt, H. & Schoenfeld, A. H (2003) Improving educational Research: Toward a More Useful, More Influential, and Better Funded Enterprise. *Educational Researcher*, 32(9), Pp3-14
- Burton, D. Bartlett, S. (2005) *Partitioner Research for Teachers*. London: Sage

- Burton, N. Brundett, M. Jones, M. (2008) *Doing Your Education Research Project*. London: Sage.
- Bushby (2018) [<https://www.independent.co.uk/news/education/education-news/primary-school-cheating-tests-sats-exam-teachers-pupil-stress-a8656831.html>] [internet] (accessed 1.5.2019)
- Carr, J. (2019) <https://schoolsweek.co.uk/five-year-teacher-retention-rate-worsens-and-5-other-school-workforce-findings/> {Internet} (Accessed 20.5.2024)
- Chandler, D. (2014) *Resilience: The Governance of Complexity*. United Kingdom: Routledge.
- Children's Society Report 2020 <https://www.childrenssociety.org.uk/sites/default/files/2020-11/Good-Childhood-Report-2020.pdf> [Internet] (Accessed 10.1.21),
- Christensen, J. (2016) *A Critical Reflection of Bronfenbrenner's Development Ecology Model*. Problems of education in the 21<sup>st</sup> Century. Vol 69. Pp 22-28.
- Claxton, G., Chambers, M.R., Powell, G. and Lucas, B. (2011) *The Learning Powered School: Pioneering 21st century education*. Bristol: TLO.
- Cochran-Smith, M. & Lytle, S. (2009) *Relationships of Knowledge and Practice: Teacher Learning in Communities*. Review of Research in Education, Vol 24. Pp.249-305.
- Cochran-Smith, M. & Lytle, S. (2009) *Inquiry as Stance: Practitioner Research for the Next Generation*. Practitioners Inquiry. London: Teachers College Press.
- Coe, R. (2012) *Nature of Educational Research (2012)* in Arthur, J., Waring, M., Coe, M. and Hedges L. (eds)(2012) *Research Methods and Methodologies in Education*. London: Sage. Pp9-20.
- Conner, C. (1991) *Assessment and Testing in the Primary School*. Hampshire: Falmer press.
- Costa, A. (2008) *Habits of Mind: Strategies for Disciplined Choice Making*. The Systems Thinker. Vol 19 (10). Pp 2-5.
- Crooke, A. (2011) *Music Therapy, Social Policy and Ecological Models: A Located Example of Music in Australian Schools*. <https://voices.no/index.php/voices/article/view/2279/2034> [Internet] (Accessed 3.6.2022)
- Davis, A.J. (1998) *The Limits of Educational Assessment (journal of philosophy of education)*. Oxford: Blackwell Publishers.
- Daniel, B., Wassell, Campbell, I. and Wassell, S. (2002) *The Early Years: Assessing and Promoting Resilience in Vulnerable Children: V. 1*. 3rd edn. Philadelphia: Jessica Kingsley Publishers.
- Dawson, C. (2009) *Introduction to Research Methods A Practical Guide for Anyone Undertaking a Research Project*. 4<sup>th</sup> edn, Oxford, How to Books Ltd
- Dewey, J. (1938) *Experience and Education*. New York: Macmillan Company.
- Department for Education, (2010) *The Importance of Teaching: Schools White Paper 2010*. <https://www.gov.uk/government/publications/the-importance-of-teaching-the-schools-white-paper-2010>. [Internet] (Accessed 8.1.2011)

- Department for Education (2013) <https://www.gov.uk/government/news/teach-first-to-recruit-more-top-graduates> [Internet] (Accessed 4.8.2018)
- Department for Education, (2015) *Final Report of the Commission on Assessment without Levels*. <https://www.gov.uk/government/publications/the-importance-of-teaching-the-schools-white-paper-2010> [Internet] (Accessed 5.1.2016)
- Department for Education, (2016 ) Educational Excellence Everywhere: Schools White Paper 2016. <https://www.gov.uk/government/publications/educational-excellence-everywhere> [Internet] (accessed 1.4.2016).
- Department for Education, (2018) Primary School Accountability Guide. <https://dera.ioe.ac.uk/id/eprint/32127/> [Internet] (Accessed 3.2.2019)
- Department for Education (2021) <https://www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education/physical-health-and-mental-wellbeing-primary-and-secondary> [Internet] (Accessed 1.10.2021)
- Department for Education, (2022) [https://assets.publishing.service.gov.uk/media/62e7a429d3bf7f75af0923f3/Executive\\_Summary.pdf](https://assets.publishing.service.gov.uk/media/62e7a429d3bf7f75af0923f3/Executive_Summary.pdf) [Internet] (Accessed 10.4.2022)
- Department for Education, (2023) *Working Lives of Teachers and Leaders, Wave 1 Research Report*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1148571/Working\\_lives\\_of\\_teachers\\_and\\_leaders\\_-\\_wave\\_1\\_-\\_core\\_report.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1148571/Working_lives_of_teachers_and_leaders_-_wave_1_-_core_report.pdf) [Internet] (Accessed 1.5.23)
- Donmoyer, R. (2000) *Generalizability and the Single Case Study* In Gomm, R., Hammersley, M. and Foster, P. (eds.) (2000) *Case Study Method: Key Issues, Key Texts*. Thousand Oaks, CA: SAGE Publications. Pp48-65
- Drill, K. Miller, S. Behrstock-Sheratt, E. (2012). Teachers' Perspectives on Educational Research. *Brock Education Vol 23 (1) pp3-17*
- Drummond, M. (2008) *Assessment and Values. A Close and Necessary Relationship* in Swaffield, S. (ed) (2008) *Unlocking Assessment. Understanding for Reflection and Application*. Oxon: Routledge. pp11-19.
- Dweck, C.S. (2006) *Mindset. The New Psychology of Success*. London: Random House.
- Dweck, C.S. (2012) *Mindset: How You Can Fulfil Your Potential*. London: Little, Brown Book Group.
- Dweck, C.S. (2019) *The Choice to Make a Difference*. Association of Psychological Science, Vol 14, Issue 1, pp 21-25.
- Dweck, C.S. and Carol, D.S. (2000) *Self-theories: Their Role in Motivation, Personality, and Development*. Philadelphia, PA: Psychology Press.
- Edmonds, R. (1982) *Programs of School Improvement: An Overview* in Education Leadership. Vol 40. PP4-11.
- Eisner, E. & Peshkin, A. (1990) *Qualitative Inquiry in Education: The Continuing Debate*. Columbia: Teachers College Publications.

- Eggleston, J. (1991) *School Examinations; Some Sociological Issues* in Horton, T. (ed.) (1991) *Assessment Debates (curriculum and learning)*. London: Hodder & Stoughton in association with the Open University. Pp58- 65.
- Elliot, J. (2007) *Educational Research as a form of Democratic Rationality* in Bridges, D. and Smith, R. (eds) (2007) *Philosophy, methodology and Educational Research*. Oxford: Blackwell Publishing. Pp149-156
- Engler, K. (2007) A look at how the ecological systems theory may be inadequate. A Capstone Project. Winona State University.
- Evers, C. and Wu, E.H. (2007) *Generalising from Single Case Studies: Epistemological Reflections* in Bridges, D. and Smith, R. (eds) (2007) *Philosophy, methodology and Educational Research*. Oxford: Blackwell Publishing.
- Ferretti, N., Ganley, C. & Kofler, M. (2019) *Predicting children's school grades: Unique and interactive effects of parental beliefs and child inattention/hyperactivity symptoms* in British Journal of Development Psychology. Vol 27 (2). Pp 300-307.
- Filler, A. (1994) *Teacher Assessment: A Sociological Perspective* in Hutchison, D. and Schagen, I. (eds.) (1994) *How Reliable is National Curriculum Assessment?* Slough: National Foundation for Educational Research. Pp45-50.
- Fleming, J. & Ledogar, R. (2008) *Resilience, an Evolving Concept: A Review of Literature Relevant to Aboriginal Research*. Pimatisiwin, Vol 6 (2). Pp 7-23.
- Garcia-Crespo, F. (2021) *Academic Resilience in European Countries: The role of teachers, families and student profiles*. American Psychological Association. Vol 16 (7) Pp 1-20.
- Gibbs, A. (2012) *Focus Groups and Group interviews* in Arthur, J., Waring, M., Coe, M. and Hedges L. (eds)(2012) *Research Methods and Methodologies in Education*. London: Sage. Pp185-189
- Gilliard, D. (2018) *Education in England. The History of Our Schools*. <https://education-uk.org/history/index.html> [Internet] [Accessed 15.5.18]
- Gipps, C. and Stobart, G. (1993) *Assessment: A Teachers' Guide to the Issues*. 2nd edn. London: Hodder & Stoughton Educational Division.
- Gipps, C. (1994) *Beyond Testing. A Theory of Educational Assessment*. London: Falmer press.
- Glaserfeld, E. (1989) *Cognition, Construction of Knowledge and Teaching*. Synthese, Vol 80 (1) pp 121-140.
- Guba, E. & Lincoln, Y. (1982) *Epistemological and Methodological Bases of Naturalistic Inquiry*. Educational Communication and Technology, Vol 30 (4). Pp 233-252.
- Gray, D.E. (2013) *Doing Research in the Real World*. 3rd edn. London: SAGE Publications.
- Greenfield, T. (2002) *Research Methods for Graduates*. London: Hodder-Arnold.
- Hamilton, L., Corbett-Whittier, C. and Fowler, Z. (2013) *Using Case Study in Education Research*. London: SAGE Publications.
- Hannon, P. (1990) *Parents' and Teachers' Perspectives on Preschool Literacy Development*. British Educational Research Journal. Vol 16 (3). Pp 259-272.

- Hanushek, E. Piciunik, M. Wiederhold, S. (2019) *The Value of Smart Teachers*. Journal of Human Resources, Vol 54(4) pp857-899.
- Haralambos, M. & Holborn, M. (2013) *Sociology. Themes and Perspectives*. London: Collins.
- Hart, A., Blincow, D. and Tomas H. (2007) *Resilient Therapy, Working with Children and Families*. London: Routledge.
- Hart, B. & Risley, T. (1990) *The Early Catastrophe. The 30 Million Word Gap*. Psychology Education. The American Educator. Vol 27. Pp 4-17.
- Harlen, W. (2008) *Trusting Teachers Judgement* in Swaffield, S. (ed) (2008) *Unlocking Assessment. Understanding for Reflection and Application*. Oxon: Routledge. pp139-150
- Harlen, W. (1993) *Teaching and Learning Primary Science*. Virginia: Paul Chapman Publishing.
- Hart, S., Dixon, A., Drummond, M.J. and McIntyre, D. (2004) *Learning without Limits*. Maidenhead: Open University Press
- Holt, J. (1969) *The Underachieving School*. London: Pitman.
- Holweck, T., Netolicky, D. & Campbell, P. (2021) *Defining and Exploring Pracademia: Identity, community and Engagement*. Journal of Professional Capital and Community, Vol7 (1). Pp 6-25.
- Hunnybun, G. (2012) *Understanding Resilience in Children; Implications for Social Work Practice*. MSc, Leeds Metropolitan University.
- House Of Commons Library (2017)  
<https://publications.parliament.uk/pa/cm201617/cmselect/cmeduc/682/68202.html>  
[Internet] (Accessed 28.4.2018).
- House of Commons Library (2020) *Grammar School Statistics*  
<https://commonslibrary.parliament.uk/research-briefings/sn01398/>. [Internet]  
(Accessed 8.7.2020)
- House of Commons NC review (2021)  
<https://commonslibrary.parliament.uk/research-briefings/sn06798/> [Internet] (Accessed 10.9.2021)
- Horton, T. (ed.) (1991) *Assessment Debates (curriculum and learning)*. London: Hodder & Stoughton in association with the Open University.
- Howard, S., Dryden, J. & Johnson, B. (1999). Childhood resilience: Review and critique of literature. Oxford Review of Education, 25(3), Pp307–323
- Hutchison, D. and Schagen, I. (eds.) (1994) *How Reliable is National Curriculum Assessment?* Slough: National Foundation for Educational Research.
- James, M. (2008) *Assessment and learning* in Swaffield, S. (ed) (2008) *Unlocking Assessment. Understanding for Reflection and Application*. Oxon: Routledge. pp20-30.
- Jerrim, J. (2021) *National tests and the wellbeing of primary school pupils: new evidence from the UK*. Assessment in Education: Principles, Policy and Practice. Vol 28, Issue 5-6 pp 507-544.



- Johnston, P. & Costello, P. (2009) *Principles for Literacy Assessment*. Reading Research Quarterly. Vol.40 (2). Pp 256-267.
- Kaiser, K. (2009) *Protecting Confidentiality* In Gubruim, J. Holstein, J. Marvash, A. and McKinney, K. *Sage Handbook of Interview Research*. London: Sage
- Keddie, N. (1973) *The Myth of Cultural Deprivation*. London: Penguin.
- Kilpatrick, W. (1925) *Foundations of Method: Informal Talks on Teaching*. New York: Macmillan.
- Kohn, A. (2000) *The Case Against Standardized Testing: Raising the Scores, Ruining the Schools*. Portsmouth: Heinemann.
- Kvale, S. (1996) *Interviews: An introduction to Qualitative Research Interviewing*. London: Sage.
- Lucas, B., Claxton, G., Spencer, E. and Costa, A.L. (2013) *Expansive Education. Teaching Learners for the Real World*. Maidenhead, England: Open University Press.
- MacBlaine, S. (2018) Bronfenbrenner: Children's Learning in a Wider context.
- Marlow, R. Norwich, B., Ukoumunne, O., Hansford, L., Sharkey, S and Ford, T. (2014) A Comparison of Teacher Assessment (APP) with Standardised Test in Literacy and Numeracy. *Assessment in Education: Principles, Policies and Practice* 21 (4) pp 412-426.
- McAllister, P. (2018) A Teachers' Perspective on What's Wrong with Our Schools. *Cato Journal*, Vol. 38, No. 1 (Winter 2018) .Pp225-245.
- Milne, C. (2017) <https://fullfact.org/education/teachers-first-class-degrees/> {Internet} (Accessed 5.5.2020)
- Mohaupt, S. (2009) *Resilience and Social exclusion*. Social Policy and Society. 8 pp63-71.
- Murphy, R and Torrance, H.(1988) *The Changing Face of Educational Assessment*. Milton Keynes, United Kingdom: Open University Press.
- O'Grady, S. (2018) <https://www.express.co.uk/news/uk/959663/sats-2018-half-pupils-worried-embarrassed-badly-results> [Internet] (Accessed 1.9.2019)
- O'Leary, P. Tsui M (2022) *Lived Experience: A Constant Companion for the Social Work Relationship*. Internal Social Work, Sage Journals. Vol 65 (6) pp1075-1077.
- Openakker, R. (2006) Advantages and Disadvantages of four interview techniques in Qualitative Research. *Forum Quality Social Research*, Vol 3, No7, Article 11.
- Nice, R. (1997) *Outline of a Theory of Practice. Pierre Bordieu*. Cambridge: Cambridge University Press.
- Mason, J. (2009) *Qualitative Researching*. 2<sup>nd</sup> edn. London: Sage.
- Miles, M. B. & Huberman, A.M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook* (2<sup>nd</sup> Ed). Thousand Oaks, CA:Sage.
- Oliver, P. (2014) *Writing Your Thesis*, 3<sup>rd</sup> edn. Sage Study Skills. London: Sage.

- Opie, C. (ed) (2008) *Doing Educational Research*. London: Sage.
- Panda, A. (2014), *Bringing academic and corporate worlds closer: we need pracademics*, *Management and Labour Studies*, Vol. 39 No. 2, pp. 140-15
- Piaget, J. (1936) *The Origins of Intelligence in Children*. London: Intl Universities Press.
- Pollitt, A. (1994) *Measuring and Evaluating Reliability in NC Assessment* in Hutchison, D. and Schagen, I. (eds.) (1994) *How Reliable is National Curriculum Assessment?* Slough: National Foundation for Educational Research.
- Punch, K. (2011) *Introduction to Research Methods in Education*. London: Sage.
- Prilleltersky, I. and Prilleltersky, O. (2005) *Beyond resilience: Blending Wellness and Liberation in the Helping Professions* in Ungar, M. (ed) *Handbook of Working with Children and Youth*. London: Sage. Pp88-103.
- Report of the Assessment Review Group (2017) *Redressing the Balance*.  
<https://earlyexcellence.com/wp-content/uploads/2017/04/Redressing-the-Balance-Assessment-Review-Group-report.pdf> [Internet] (Accessed 12.5.2018)
- Reay, D. & Wiliam, D. (1999) '*I'll be a nothing*': structure, agency and the construction of identity through assessment, *British Educational Research Journal*, 25:3, PP 343-354.
- Reid, K. (2019) *Managing School Attendance: Successful Intervention Strategies for Reducing Truancy*. London: Routledge.
- Rhodewalt, F. (1994) *Conceptions of Ability, Achievement Goals and individual Differences in self-handicapping behaviour: On the Application of Implicit Theories*. *Journal of Personality*, 62, pp67-85.
- Rhodewalt, F. and Davison, J. (1986) *Self-Handicapping and Failure*. *Basic and Applied Social Psychology*, 7 pp307-322.
- Richardson, G. (2002) *The Methatheory of Resilience and Resiliency*. *Journal of Clinical Psychology*, Vol 58 (3). Pp 307-321.
- Robinson, M. (2014) *The Feeling Child. Laying the Foundations of Confidence and Resilience*. Oxon: Routledge.
- Rosen, M. Williams, E. (2013) <https://www.historyworkshop.org.uk/education/secondary-modern/> [internet] (Accessed 8.7.2020).
- Rosenthal, R. & Jacobson, L. (1968) *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt.
- Rutakumwa, R. Mugisha, J. Bernays, S. Kabumga, E. Tumwekwase, G. Mbonye, M and Seeley, J. (2019) *Conducting in-depth Interviews with and Without Voice Recorders: A Comparative Analysis*. *Qualitative Research* Vol 20 (5) pp565-581.
- Sainsbury, M. (1994) *The Structure of National Curriculum Assessment* in Hutchison, D. and Schagen, I. (eds.) (1994) *How Reliable is National Curriculum Assessment?* Slough: National Foundation for Educational Research. Pp2-12.

- Saleebey, D. (1996) *The Strengths Perspective in Social Work Practice: Extensions and Cautions*. Social Work Journal, 41:3 Pp 296-305.
- Savin-Baden, M and Howell-Major, C. (2013) *Qualitative Research. The Essential Guide to Theory and Practice*: Oxen: Routledge.
- Schoenfeld, A.H. (2009) instructional Research and the Improvement of Practice. In Bransford, N., Vye, d. Stypek, L. Gomex, & D. Lam (Eds), *The Role of Research in Educational improvement*. (pp.161-188). Cambridge, MA: Harvard Education Press.
- Seligman, M. (1967) *Learned Helplessness: Theory and Evidence*. Journal of Experimental Psychology. Vol 105 (1). Pp 3-46.
- Severs, J. (2019) <https://www.tes.com/magazine/archive/growth-mindset-where-did-it-go-wrong> [internet] (Accessed 8.3.2021)
- Sisk, Victoria F. Burgoyne, Alexander P., Jingze Sun, Butler, Jennifer L. McNamara, Brooke. *To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses*. *Psychological Science*, 2018; 29 (4): 549
- Skedsmo, G. and Huber, A. (2022) *Data Driven Approaches to Education Governance and Their Implications*. Educational Assessment, Evaluation and Accountability 34:1-4.
- Stake, R.E (1994) *Case Studies* in N. K. Denzin & Y. S. Lincoln (Eds.) *Handbook of Qualitative Research* (pp236-247). Thousand Oaks, CA: Sage.
- Stake, R.E. (1995) *The Art of Case Study Research*. 11th edn. Thousand Oaks, Ca: Sage Publications.
- Stenhouse, L. (1984) *An Introduction to Curriculum Research and Development*. London: Heinemann.
- Styron, J. and Styron, R. (2012) *Teaching to the Test: A Controversial Issue in Quantitate Measurement*. Journal of Systemics, Cybernetics and Informatics, Volume 10.
- Sugrue, C (ed) (2008) *The Future of Educational Change and International Perspectives*. Oxon: Routledge
- Sumner, R. (1987) *The Role of Testing in Schools*. Windsor: NFER.
- Sutherland, M.B. (1988) *Theory of Education*. 2nd edn. New York: Longman.
- Swaffield, S. (ed) (2008) *Unlocking Assessment. Understanding for Reflection and Application*. Oxon: Routledge.
- Swanborn, P.G. (2010) *Case Study Research: What, Why and How?* Los Angeles: SAGE Publications.
- Tambini, J. (2017) <https://www.express.co.uk/news/uk/842559/GCSEs-v-O-levels-exams-getting-easier-results-day-gcse-exam> {Internet} (Accessed 1.9.2019)
- Tregaskis, S. (1995) *Into the Woods* Human Ecology, Spring 2015, Vol 43 (1). Pp 12-18.

- Travers, M. (2001) *Qualitative Research Through Case Studies*. Thousand Oaks, CA: SAGE Publications.
- Thomas, G. (2016) *How to do Your Case Study*. 2<sup>nd</sup> edn, United Kingdom: SAGE Publications.
- Torrance, H. (ed) (1995) *Evaluating Authentic Assessment*. Buckingham: Open University Press.
- Torrance, H., Pryor, J.H. and Torrance, H. (1998) *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. Philadelphia: Open University Press.
- Torrance, H. and Murphy, R (1991) *The Need for Change* in Horton, T. (ed.) (1991) *Assessment Debates (curriculum and learning)*. London: Hodder & Stoughton in association with the Open University. Pp12-20
- Turner, C. (2017) <https://www.telegraph.co.uk/education/2017/04/30/sats-risk-giving-children-mental-health-issues-education-select/> [internet] (Accessed 13.8.2018).
- Turner, J.C. (1882) *Towards a cognitive redefinition of the social group* in Tajfel, H. (ed) (1982) *Social Identity and Intergroup relations*. Cambridge: Cambridge University Press.
- Ward-Schofield, J (2000) *Increasing the Generalizability of Qualitative Research* in Gomm, R., Hammersley, M. and Foster, P. (eds.) (2000) *Case Study Method: Key Issues, Key Texts*. Thousand Oaks, CA: SAGE Publications. Pp71-74.
- Weall 2017 <https://www.theguardian.com/education/2017/may/01/sats-primary-school-children-suffering-stress-exam-time> [internet] (accessed 13.8.2018).
- Weiner, B. (2000). *Intrapersonal and Interpersonal Theories of Motivation from an Attributional Perspective*. *Educational Psychology Review*, 12, Pp 1-14.
- Wertsch, J.V. ( 2015) *Making Human Beings Human: Biological Perspectives on Human Development led by Urie Bronfenbrenner*. *British Journal of Developmental Psychology*. Vol 23 (1). Pp 143-152.
- West, A. (1994) *Choosing Schools – The Consumers’ Perspective* in Halstead, M. (ed) (1994) *Parental Choice and Educational Principles, Policy and Practice*. London: Kogan Page. Pp109-120.
- Wellington, J.J. (2000) *Educational Research: Contemporary Issues and Practical Approaches*. London: Continuum International Publishing Group.
- William, D. (2008) *Quality in Assessment* in Swaffield, S. (ed) (2008) *Unlocking Assessment. Understanding for Reflection and Application*. Oxon: Routledge. pp120-138
- William, D. Thompson, M. (2007) Tight but Loose: A Conceptual Framework for scaling Up School Reforms. Paper presented at the American Educational Research Association (AERA) between 9<sup>th</sup> April 2007 and April 13<sup>th</sup> in Chicago, Illinois
- Williams, J. (2013) *Towards a Standards Based Curriculum. A Toolkit for Primary Schools in England*. London: Create Space.
- Winch, C. & Gingell, J. (2004) *Philosophy and Educational Policy*. London: Routledge.

Yin, R. K. (1994). *Case Study Research*. Newbury Park, CA: Sage.

Yin, R.K. (2009) *Case Study Research: Design and Methods (applied social research methods)*. 4th edn. Thousand Oaks, CA: Sage Publications.

