



---

# ‘HOW DO PRIMARY SCHOOLS ENACT POLICY? THE CASE OF ASSESSMENT WITHOUT LEVELS’

---

For the Degree of Doctor of Philosophy in Education.



JANUARY 18, 2025  
SHAILEN POPAT  
Brasenose College, University of Oxford

## Thesis Abstract

This thesis conveys the importance of understanding how school leaders engage in sensemaking to comprehend why a policy has been enacted in a particular way. Through interviews with headteachers, teachers, school governors, parents, and pupils, that were conducted in 2017/18 and in 2019, this thesis has sought to document differences and synthesise similarities on the enactment of the policy of Assessment without Levels (AwL) in six primary schools in England and to treat this as an indicative case of how schools respond to and enact a significant policy that reforms aspects of schooling such as curricula, pedagogy and assessment. This thesis presents how AwL was enacted in these schools both in-terms of the new progress capture and feedback processes, and also in-terms of changes to pedagogy to enact Assessment for Learning (AfL) and mastery learning. Where there are similarities between the systems, analysis discusses why the national environment of schooling in England may make that so; and where there are differences, discussion on stakeholder influences, headteacher values, and other local factors is offered. What emerges is that the enactment of AwL tells the story of many of the challenges of improving educational outcomes for pupils that schools across England were likely to be encountering during that period. This thesis concludes that the enactment of AwL was successful in the six participant schools and a heuristic as to why that was is offered which credits the clarity of the communication of the policy intentions, coupled with the latitude for headteachers to design the policy and for teachers to believe in the changes, together with strong national accountability systems that ensure standardisation. Finally, very detailed recommendations for future research on topics such as feedback, performance indicators, mastery learning, and the narrowing of the curriculum are offered.

## **Acknowledgements**

Shiva, thanks for constant support and assistance.

Mum, you are a diamond and I'm glad that you are around to see me get to Oxford and become a DPhil.

Velda, thanks for believing in me and for being pivotal in my career change and indeed it has been life-changing.

To all those at Oxford University who supported my studies.

To all those at Oxford University and the UK Government who enabled me to overcome my visual disability to be able to study and graduate through providing funds and facilities.

To Jason and my colleagues at University of Birmingham who supported my studies by giving me time and feedback.

Thank you.

## Table of Contents

<b>Acknowledgements</b> .....	2
<b>Chapter 1: Introduction to the Thesis</b> .....	6
<b>Chapter 2: Literature Review</b> .....	21
<b>Chapter 3: Research Design and Methods</b> .....	46
Chapter 4: The Types of AwL systems being used .....	83
<b>Chapter 5: AwL Internal Moderation as the enactment of an Audit Culture</b> .....	112
<b>Chapter 6: Findings and Discussion on the Enactment of Assessment for Learning through the Enactment of Assessment without Levels</b> .....	126
<b>Chapter 7: The Headteacher as Coordinator of Management Approaches towards Stakeholders</b> .....	165
<b>Chapter 8: Insights into the Wider Paradigm of Policy Enactment in Schools</b> .....	192
<b>Chapter 9: Conclusions and Recommendations for Future Research</b> .....	214
References .....	231

## List of Figures

- Figure 1: Reflection Questions provided by CAWL (2015)
- Figure 2: Hook's SOLO Taxonomy presented in NCTL (2014)
- Figure 3: An example of a Learning Progression Guide (Alonzo and Steedle 2009:16-17)
- Figure 4: An example of Learning Intentions and Success Criteria provided by NCTL (2014)
- Figure 5: NCTLs (2014) suggested points-based scale for AwL
- Figure 6: The Case and Units of Analysis
- Figure 7: The Research Process
- Figure 8: Development Process of Phase 1 Interview Questions<sup>1</sup>
- Figure 9: A Standardised Network Map based on Antonucci (1986)
- Figure 10: Wall Display on SchoolQualitative's Values outside the Year 6 classroom
- Figure 11: A Wall Display in the Assembly Hall of some of SchoolPAT's Values
- Figure 12: A Wall Display of SchoolPAT's Motto
- Figure 13: A Banner Advertising SchoolPAT's Ofsted Judgement
- Figure 14: Pupils in SchoolQualitative Year 6 Group Discussions Showing the Researcher how they self-assess and self-evaluate their own work
- Figure 15: Pupils in SchoolPAT Showing their Corrections in Response to Feedback
- Figure 16: The Continuous Reflexive Process of Induction and Deduction in this DPhil Study
- Figure 17: A Pupil's progress through a Year Group in the Passports system
- Figure 18: SchoolPoints AwL System Categories
- Figure 19: Assessment Principles behind the Continuum
- Figure 20: The Process used in the Continuum
- Figure 21: Example of Continuum 'I can' statements within English Writing Points 4, 5, 6 and 7
- Figure 22: The 2018 Year 6 KPIs for Reading published by the National Association of Head Teachers (NAHT)
- Figure 23: Policy Recommendations under the Coalition Government (2010-2015) have combined with Policies under the New Labour Government (1997-2010)
- Figure 24: The Overload of Incongruent Criteria on Teachers Cognition
- Figures 25 & 26: Head-Continuum's Policy Interpretation & Enactment Stage Maps
- Figures 27 & 28: Head-Qualitative Policy Interpretation and Enactment Stage Maps
- Figures 29 & 30: Head-SPT Policy Interpretation & Enactment Stage Maps
- Figures 31 & 32: Head-PAT Policy Interpretation Stage Map
- Figures 33 & 34: Head-Points Policy Interpretation and Enactment Maps
- Figures 35 & 36: Head-Passports Policy Interpretation & Enactment Stage Maps
- Figure 37: SchoolContinuum Stakeholder Influence Chart

- Figure 38: SchoolQualitative Stakeholder Influence Chart
- Figure 39: SchoolSPT Stakeholder Influence Chart
- Figure 40: SchoolPAT Stakeholder Influence Chart
- Figure 41: SchoolPoints Stakeholder Influence Chart
- Figure 42: SchoolPassports Stakeholder Influence Chart
- Figure 43: A Typology of Policy Work Roles in the Translation of a Policy
- Figure 44: Management Approaches of School Stakeholders during Policy Enactment  
(Adapted from Chase et al.'s (2000) management approaches towards stakeholders)
- Figure 45: The interaction between school stakeholder perspectives
- Figure 46: Frames of Sensemaking in a School
- Figure 47: The Four Keys to the Successful Policy Implementation of AwL

### **List of Tables**

- Table 1: An example of a National Curriculum Level Descriptor for Number and Algebra  
(DCSF et al. 2010:30)
- Table 2: National Curriculum Level Descriptors (DfE 2010:31)
- Table 3: Phase 1 Units of Analysis and Ciphers
- Table 4: Interviewee Roles and Ciphers in SchoolQualitative
- Table 5: Interviewee Roles, and Ciphers in SchoolPAT
- Table 6: SchoolQualitative Data Generation Sites
- Table 7: SchoolPAT Data Generation Sites
- Table 8: Themes from literature and Phase 1 data Informing Phase 2 Interviews
- Table 9: Examples of Thematic Codes Categorised by a Research Question
- Table 10: Similarities and Differences between AwL Systems
- Table 11: Summary Table of all schools Interpretation Maps
- Table 12: Summary Table of the Influence of the Stakeholders During the Enactment Phase
- Table 13: Aggregated Influence of the Stakeholders in each school
- Table 14: Ball et al. (2011:626) Typology of Policy Actors and Policy Work in a School

## Chapter 1: Introduction to the Thesis

### 1.1 Policy sensemaking and enactment in primary schools

This thesis conveys the importance of understanding how school leaders engage in sensemaking to comprehend why a policy has been enacted in a particular way. The sensemaking process conveys the underlying interpretations, values, and contextual factors that influence policy decisions. By examining the interactions and collaborative efforts involved in sensemaking, readers gain insights into the rationale behind policy enactment. This understanding informs future policy development, ensuring that policies are more effectively aligned with the needs and realities of the school community. The sensemaking process conveys an inherently interpretive nature, involving the construction of meaning from complex and sometimes ambiguous policy directives. School leaders bring their own experiences, beliefs, and values to the table when interpreting policies. These personal and professional backgrounds shape how they understand and implement policy mandates. By understanding the sensemaking process, stakeholders uncover these underlying interpretations and values, providing a clearer picture of why certain decisions are made (Beard 2025).

Context conveys a significant role in policy sensemaking. The unique characteristics of a school, including its demographic makeup, community values, and existing practices, all influence how policies are interpreted and enacted. Understanding the contextual factors that influence policy decisions helps stakeholders appreciate the complexities involved in policy enactment and the need for tailored approaches (Rycroft-Malone et al. 2022). Policy sensemaking is a collaborative activity involving interactions among various stakeholders within the school community. Teachers, administrators, parents, and even students contribute to the sensemaking process. These interactions help to construct a shared understanding of policy directives and facilitate collective action. By examining these collaborative efforts, readers gain insights into the dynamics of policy implementation and the roles different actors play in shaping outcomes (Donaldson et al. 2021). This understanding highlights the importance of fostering a collaborative culture within schools to support effective policy enactment.

A deep understanding of the sensemaking process informs future policy development. Policymakers use insights gained from examining how current policies are interpreted and enacted to design more effective and contextually relevant policies. This approach ensures that policies are not only well-intentioned but also practical and adaptable to the diverse

realities of schools. One of the key benefits of understanding policy sensemaking conveys the ability to align policies more closely with the needs and realities of the school community. When policymakers and school leaders have a clear understanding of how policies are interpreted and enacted, they can identify gaps and areas for improvement. This alignment ensures that policies are not only enacted effectively but also resonate with the values and priorities of the school community (Hay 2025). Understanding the sensemaking process enhances enactment by providing insights into potential obstacles and enablers. For instance, if school leaders perceive a policy as misaligned with their school's values or context, they may be less motivated to implement it fully. By addressing these perceptions and involving leaders in the policy development process, policymakers can increase buy-in and support for new initiatives. This collaborative approach leads to more successful and sustainable policy implementation.

How primary schools enact policy is under-researched and at the time of starting this DPhil in 2016, there were no studies of how they respond when confronted with a significant change in curricula and assessment policy. Coburn (2005) used sociological theories of sensemaking to investigate how principals in two California elementary schools influenced enactment of changing reading policy. They found that principals influence teachers' enactment by shaping access to policy ideas, participating in the social process of interpretation and adaptation, and creating substantively different conditions for teacher learning in schools. They argue that these factors are influenced by principals' understandings about reading instruction and teacher learning. Hu et al. (2011) investigated how three primary school English teachers in mainland China interpreted, negotiated, and represented the recent task-based language teaching innovation in their classroom practice. They found that working environments, teachers' beliefs about language teaching, their understanding (or lack thereof) of the intended curriculum, their perceptions of the students' learning needs, prevalent instructional practices found in their teaching contexts, availability of resources, support (or lack thereof) for the adoption of the innovation all interacted with each other, which influenced their interpretation and construction of the pedagogical approach in their classrooms. Braun and Maguire (2018) explored teachers' experiences of assessment policy enactments in two London primary schools in 2015 under the previous National Curriculum when level descriptors were being used. Their first observation was that much of teacher sensemaking centred on second-guessing what version of standards they should focus upon as they felt that national policy goalposts were constantly shifting. Their second observation was that some teachers resented the target of students making 'more than expected progress' as it implied that achieving expected progress was a failure. Thirdly, that teachers were being forced to target interventions on groups of learners as defined by

data collection such as gender or ethnicity rather than tailoring provision to the individual pupil based on a class teacher's judgement. Participant teachers expressed a feeling that trust in their judgements was often supplanted by the data. Their fourth observation was that the focus upon the data meant that pupils were referred to as their level of attainment, such as, 'this pupil is a level 3', and therefore, there was a culture of labelling. Their fifth observation was that national testing dominated teaching and that the fear of students attaining poor results could result in dubious reporting practises. One of the contributions of this thesis is to update our understanding of how teachers feel about assessment policy enactments when they have been given greater autonomy to design the approach that is used within their context. This thesis also follows a particular policy being interpreted and enacted rather than documenting experiences conveyed by teachers about assessment policies more generally.

I have extensive experience as a local authority manager within school support services and has also operated his own non-profit youth services company. These roles have necessitated frequent interpretation and enactment of national and organizational policies, fuelling his interest in the factors influencing policy interpretation and enactment. I had observed that the intention behind a policy is often diluted or altered during enactment, and the enacted policy becomes the accepted version, even if it significantly diverges from the original intent. This phenomenon is particularly pronounced when enactment occurs over a large geographical area and in multiple settings. Each setting that I managed had dominant actors, resource priorities, working practices, and cultures that significantly influenced how a policy was interpreted and enacted. I became aware of AwL during my Master's Degree in Educational Assessment at the University of Warwick, which included modules on Policy and Resources, Issues in Assessment, and Assessment for Learning. My dissertation examined whether the rationale for the removal of level descriptors, as posited in Chapter 8 of the Report by the Expert Panel for National Curriculum Review 2011 was considered valid by teachers and parents in a Free school in Derby in the first few months of AwL (Popat 2015). The Panel had argued that levels, being best-fit measures, obscured teachers' ability to specifically identify what a pupil knew and could do, thereby reducing the quality of Assessment for Learning approaches and the accuracy of interventions. They also suggested that removing levels would raise teachers' expectations for all students, address social differentiation, and prevent pupils from labelling themselves (e.g., "I'm a Level 4"). The Panel contended that reporting in levels was overly burdensome for teachers and unclear for parents and pupils (DfE, 2011). The teachers in the researcher's study had differing views from The Expert Panel, perceiving that AwL descriptors such as emerging, developing, and secure, were also best-fit measures like levels. They asserted that with time and good

interpretative skills, teachers have always been able to use assessment data to target learning support, and that levels had no bearing on Assessment for Learning practices. These teachers acknowledged that levels might have been burdensome for some, but they believed that an AwL system would be equally demanding. Rather than removing levels, they felt that social differentiation could be better addressed by examining ability grouping, and that empowering high expectations would require training staff in a 'growth mindset' (Dweck 2006). There was scepticism about whether removing levels would prevent children from labeling themselves, as pupils often seek a number or grade. Of greater concern to these teachers was the reporting of pupils being "ready to progress" only in key elements of core subjects, which they felt would perpetuate and possibly extend a narrowing of the curriculum as teachers prioritize these aspects at the expense of others. The parent focus group found levels obscure, but they also found the new AwL system to be similarly unclear. One parent mentioned testing her child at home online to discern his knowledge and skills because the school's feedback was obscure. This DPhil project aimed to scale up the Master's study and investigate what had happened in several primary schools in the subsequent years as AwL was embedded.

In their exploration of policy enactment in secondary schools, Maguire et al. (2015) argued that viewing policy implementation as a series of enactments over time enables a nuanced examination of the continual dynamic process of enacting a policy (Maguire et al. 2015). This thesis utilises the case of the enactment of Assessment without Levels (AwL) as a vehicle for providing exemplars of how AwL was enacted in six primary schools together with an exploration of why choices were made in each school context at two different points in the enactment process – the initial stage of interpretation and after the AwL system had been enacted. Yin (1989) argued that 'the case study method has a specific advantage when 'a "how" or "why" question is being asked about a contemporary set of events, over which the investigator has little or no control' (Yin 1989:20). According to Yin (1989), such "how" and "why" questions 'deal with operational links needing to be traced over time, rather than mere frequencies or incidence' (Yin 1989:18). The position taken in this thesis is that a case study method exploring a single policy will provide more depth and detail on the enactment process. Pal (2005) argues that this is because situations being studied in single case studies will have a level of complexity and a density of explanation that is not usually characteristic of many other methodological approaches. Through 14 in-depth interviews with headteachers and class teachers and 10 group discussions with pupils, parents and school governors, this thesis explores how six primary schools enacted AwL between 2014 and 2019. In this thesis, the case of the enactment of the policy of AwL in primary schools in England is treated as a critical case that provides insight into the wider topic of how schools

enact a new policy (Flyvberg 2006). The units of analysis within the case are the participant schools (Teddlie and Yu 2007), and each of these units have stakeholders such as headteachers, class-teachers and pupils which are the sub-units of analysis.

The first two research questions explored in this study were defined from the start and address the “how” and “why” aspects mentioned above:

- 1) How did primary schools enact AwL? For example, how do schools capture pupil progress and how has it impacted upon pedagogy?
- 2) Why did they enact AwL that way? For example, what were the factors guiding decision making and which stakeholders influenced enactment and how?

These two research questions informed data collection questions during both Phase 1 and Phase 2. The third research question was used to add sophistication to the analysis:

- 3) What does the data tell us about primary school policy enactment more generally? For example, from the data, can one glean insights into the wider paradigm of how primary schools respond to a significant new policy?

The incorporation of an analytical question within a research framework serves as a pivotal mechanism for theoretical integration. This approach enables the researcher to bridge empirical findings with broader theoretical constructs, thereby enriching the scholarly contribution of the study. By systematically connecting data to established theories, the analytical question facilitates a deeper understanding of the phenomena under investigation (Heale and Noble 2019). This integration not only validates the empirical data but also advances theoretical discourse, offering new insights and reinforcing existing paradigms (Battistone et al. 2023). Focusing on an analytical question also fosters insightful interpretation of the data. It prompts the researcher to engage in critical thinking, allowing for the identification of underlying patterns and relationships that might otherwise remain obscured. This analytical lens encourages a sophisticated examination of the findings, moving beyond surface-level descriptions to uncover the nuanced dynamics at play. Such an approach enhances the interpretive depth of the research, providing a more comprehensive and meaningful analysis (Olmos-Vega et al. 2023). Moreover, the inclusion of an analytical question adds an essential layer of rigour to the research process. It ensures that the analysis transcends mere description, embracing a more interpretive and evaluative stance.

This rigour is crucial for understanding the complex interplay of factors influencing the research outcomes. By adopting an analytical question, the researcher is better equipped to elucidate the underlying mechanisms and implications of the findings, thereby contributing to a more robust and credible body of knowledge (Greckhamer and Cilesiz 2014).

## 1.2 The Policy of Assessment without Levels (AwL)

Since the introduction of the National Curriculum in 1988, until 2015, level descriptors were used by teachers to determine a pupil's attainment by judging which description best fits the pupil's performance. When doing so, each description was considered alongside descriptions for adjacent levels. Table 1 outlines the descriptions for Levels 3 and 4 published by the Department for Children, Schools and Families (DCSF) et al. in 2010 under the section titled 'Number and Algebra'. This national rubric of levels was removed in 2014 and not replaced, creating a situation known as Assessment without Levels (AwL).

Table 1: An example of a National Curriculum Level Descriptor for Number and Algebra

Level 4	Pupils use their understanding of place value to mentally multiply and divide whole numbers by 10 or 100. When solving number problems, they use a range of mental methods of computation with the four operations, including mental recall of multiplication facts up to 10 x 10 and quick derivation of corresponding division facts. They select efficient strategies for addition, subtraction, multiplication and division. They recognise approximate proportions of a whole and use simple fractions and percentages to describe these. They begin to use simple formulae expressed in words.
---------	--

The source of AwL is Chapter 8 of the 2011 National Curriculum Review Report by the Expert Panel. The Panel concluded that the level-based system hinders teachers from clearly understanding and identifying students' specific weaknesses or misunderstandings. They also noted that the current generalised reporting using levels hides the fact that too many students fail to grasp essential parts of the curriculum needed for their next educational phase (Popat 2015). The assessment system had traditionally reflected the psychometric tradition 'that saw ability as an innate trait that could be measured' (McCormick and Murphy 2008:10). This measurement approach differentiated students on the basis of their ability located on a normal distribution curve that was supposedly representative of society. They were streamed into ability groups and taught different curricula or separated into academic and vocational pathways. The Panel were concerned that levels led to pupils

labelling and limiting their self-belief and also resulting in teachers not keeping high expectations for all and detrimental to social inclusion. The Panel observed that in jurisdictions that perform well in international measures such as South Korea and Finland, teachers had high expectations for all students and didn't feel that that was case in all schools in England. They refer to the work of psychologist, Carol Dweck (2006) who promoted a 'growth mindset' in which teachers believe that all pupils can achieve if provided with correct support. The Panel identified that levels encouraged undue pace and progression onto more difficult work while pupils still had gaps in their knowledge or understanding. Under levels the definition of progress was movement through the rubric and pressure to achieve such movement would influence teaching decisions (Braun and Maguire 2018) Levels relied on 'dubious assumptions' (Earle 2019:14) that learning is linear with children moving in a predictable way along the continuum each year (Stobart 2008), and so teachers, pupils and parents were focusing on movement rather than mastery of concepts. Under a mastery learning approach, pupils are required to demonstrate mastery of the learning from each unit before being allowed to move on to the next, with the assumption that all pupils will achieve this level of mastery if they are appropriately supported, 'Some may take longer and need more help, but all will get there in the end' (DfE 2015:17). The new National Curriculum introduced in England in 2015 is premised on this kind of understanding of mastery. It claims to focus upon deep, secure learning for all, with extension activities on the same topic for more able students rather than rapid movement to new content (DfE 2015). As National Curriculum Levels were considered to be incompatible with such a curriculum, schools were provided with the freedom to create their own assessment systems which was deemed compatible with the wider policy agenda of increasing school autonomy. The intention to remove level descriptors came as a surprise to the profession when it was announced in (National Association of Headteachers 2014) and the decision not to provide a national rubric created uncertainty as to what the inspectorate Ofsted would be expecting to observe. This prompted the Department for Education (DfE) to establish a Commission for Assessment without Levels in early 2015 which reported that school autonomy was one of the aims behind AwL and therefore, a centralised rubric would not be provided or recommended products would not be listed and schools had the autonomy to create their own AwL systems or purchase ones that they deemed to be best suited to the needs of their pupils and not be burdensome to teachers as levels were (CAWL 2015). The autonomy to shape in-house assessment systems combined with a new curriculum based on mastery learning significantly changed pedagogy and assessment in primary schools in England and so using this as the case study to explore policy enactment produces insights that are helpful in updating understanding of policy enactment in schools.

### 1.3 Structure of this Thesis

This Introduction is Chapter 1 and Chapter 2 presents a literature review organised in six parts. The first part outlines the changing landscape of schooling over the past 50 years, providing context for the current enactment of Assessment without Levels (AwL). It notes that schools had autonomy before 1988, but the policy environment has changed significantly since 2015. The second part discusses the policies underpinning AwL and their removal to facilitate mastery learning. The third part explores learning theories associated with mastery learning. The fourth part describes AwL approaches known before Phase 1 data collection (Nov 2017-Jan 2018). The fifth part explains the policy enactment conceptual framework by Maguire et al. (2015). The sixth part details three studies that informed the research: Poet et al. (2018), commissioned by the Department of Education, reviewed AwL approaches in schools; Thrupp et al. (2018) examined the enactment of Kiwi standards in New Zealand primary schools and has parallels with this study; and the House of Commons Education Committee Report on Primary Assessment (2017) criticised the implementation of the new primary assessment system, highlighting delays, poor communication, and the negative impact on teaching. It recommended better support for teachers, a balanced approach to assessments, and measures to reduce high-stakes testing, urging a focus on a broad and balanced curriculum. Chapter 3 proceeds to detail the research design and approaches to data collection. The research design involves two phases. Phase 1 focuses on headteachers' narratives explored through interviews which were thematically analysed to understand how schools respond to the new policy of Assessment without Levels (AwL). Phase 2 aims to explore how multiple actors within schools respond to and enact the new policy. This phase captures multiple perspectives through in-depth interviews and group discussions with teachers, pupils, parents, and school governors. The selection of participants was guided by role homogeneity and purposive sampling. The data from both phases were analysed with reference to existing scholarship and theories were developed to address the research questions.

Chapter 4 explores the various Assessment without Levels (AwL) systems implemented in six primary schools, each offering unique insights into their effectiveness and challenges. The School Pupil Tracker (SPT) is an electronic package available since 2014, designed to help schools develop their own assessment systems. It supports planning, assessment recording, tracking, analysis, and report writing, aligning with National Curriculum objectives. SPT allows schools to customize Key Performance Indicators (KPIs) and uses both Detailed and Broad Tracking methods. Detailed Tracking assesses a child's attainment level using

National Curriculum records, while Broad Tracking relies on teachers' professional judgment or test outcomes. Together, the Short-Term and Long-Term Grids provide a comprehensive view of student performance, with personalized elements like diaries, targets, and progress cards enriching student profiles. The Pupil Asset Tracker is another electronic package designed to monitor student progress and attainment across various frameworks. It uses National Curriculum Age Bands and Point-in-Time Assessment (PITA) to provide clear metrics for assessing student abilities. The system's flexibility allows customization of assessment criteria and the weighting of statuses like 'Beginning,' 'Developing,' and 'Embedded.' The Tracker's color-coding feature helps identify students not making expected progress, facilitating accountability and targeted discussions with teachers. The Passports system was developed in 2014 when educational levels were removed. It focuses on core subjects like Reading, Writing, and Maths, using Performance Indicators and KPIs for each year group. Teachers compile evidence from pupil work to support assessment judgments, integrating assessment into their teaching while ensuring a manageable workload. The system includes Progress in Reading Assessment (PIRA) and Progress in Understanding Mathematics (PUMA) tests to support teacher judgments. The Points system, based on Clive Davies' materials, uses spreadsheets with curriculum statements for each year group. Teachers identify KPIs necessary for a pupil to be ready for the next year-group's curriculum, requiring all KPIs to be met. The system includes a nine-point progress scale and emphasizes mastery statements for quick learners. Moderation meetings with partner schools ensure consistent expectations. The Continuum system, developed since 2014, measures progress and attainment based on an assessment continuum. It focuses on a growth mindset approach, allowing pupils to progress at their own pace. The system breaks down reading into strands with curriculum statements for assessment, considering handwriting, sentence construction, composition, and spelling. Teachers track progress termly, aiming for significant improvements in pupil performance. Finally, the Qualitative system adopts a qualitative approach involving in-class conversations with pupils and written explanations to parents. Teachers use daily classroom interactions to assess pupils, focusing on progress rather than tracking. Pupil attainment is tracked through curriculum statements, with progress categorized as 'Good,' 'Slow,' or 'Excellent.' Moderation ensures consistent understanding and accurate assessment.

Chapter 5 explores the approaches to internal moderation and argues that being a trusted professional after the enactment of AwL refers to the triangulation of a teacher's judgement with evidence from tests, workbooks and senior teachers' perspectives prior to inputting progress data into the school's AwL system. Apple (2004) argued that such approaches require the constant production of evidence that tasks are being performed both efficiently

and correctly, a concept he referred to as an audit culture which seems to be the approach adopted in AwL moderation. The concept of internal institutionalized active inquiry involves making a teacher's work visible for documentation, scrutiny, evaluation, and enhancement, serving as internal quality assurance. This approach fosters questions and internal disagreements, crucial for quality improvement. Schools in the study used external tests combined with teacher experience and evidence from pupils' workbooks to track progress. The aim behind AwL is to restore autonomy to schools and increase teachers' sense of being trusted professionals, although this involves triangulating teacher judgments with evidence from tests and senior teachers' perspectives. This reflects debates around the audit culture, where trust in professionals is balanced with control systems. e schools adopted a social moderation approach, which ensures consistent teacher judgment through collaborative efforts. This approach, seen as a sociocultural learning experience, improves teacher judgment and fosters a community of practice where teachers learn from each other. Under AwL (Assessment without Levels), moderation serves three purposes: ensuring valid, reliable, and accurate teacher judgments; providing professional learning to improve teaching, learning, and assessment practices; and offering quality assurance by aligning judgments within and between schools. This is achieved using National SATs Frameworks in Years 2 and 6 and discussing examples of work. Effective relationships between schools, Multi-Academy Trusts (MATs), and among teachers are crucial for understanding and implementing moderation. Inter-school moderation under AwL is based on externally provided criteria, with teachers from different schools leading moderation activities in turns. This practice combines a focus on National Curriculum standards with socially created meanings. Within a school, moderation centres around learning objectives and success criteria set for each year group, with additional objectives tailored to individual pupils.

Chapter 6 details the findings from Phase 2 of the data collection where the researcher spent two days in two schools to observe classroom interactions and interview teachers, pupils, parents, and governors. The data indicates that many Assessment for Learning (AfL) practices were being implemented. Teachers appeared committed to the core principles of AfL, aiming to foster meaningful reflections on learning and develop self-regulated learners. They seemed accountable and focused on student success, with the alignment of the AwL system to curriculum statements aiding this effort. In both schools, teachers encouraged students to share their thoughts and lead discussions, with feedback rooted in dialogues between teachers and students or among students themselves. These conversations helped address prior knowledge and misconceptions, forming a foundation for understanding concepts and improving skills. This integration of classroom activities, homework, conversations, and feedback exemplifies successful AfL implementation.

Teachers explained how they adapted their teaching methods and planned additional support to ensure all students grasped key concepts and skills, using real-time interpretations of student learning and homework assessments. These skills in interpreting and adjusting for all students reflect the principle of 'high expectations for all' combined with AfL's responsiveness to student needs. Students were given opportunities to discuss feedback with their teachers and time to act on it during lessons and homework. They expressed excitement and engagement with their work and teacher feedback, showing an understanding of their personal learning goals and clarity on how to achieve them. There was no passivity; students demonstrated a positive attitude and active role in their learning, indicating that teachers had fostered autonomy and responsibility. Student involvement was meaningful and beneficial for future learning, with encouragement to self-assess, emphasizing shared responsibility for learning and assessment. Notably, students referred to qualitative feedback rather than test scores or grades. School leaders played a crucial role in implementing AfL by establishing a school-wide culture with a clear vision and expectations for AfL use. Teachers received support in using AfL-related strategies, principles of good feedback, and effective questioning, including instructional resources, materials, and preparation time for AfL activities. Collaboration was also vital, with teachers engaging in discussions about formative assessment and teaching and working together to solve shared problems. This collaboration was particularly emphasized in SchoolQualitative, where teachers participated in communities of practice for long-term learning of AfL techniques and problem-solving. A positive school climate based on trust, mutual respect, and cooperation underpinned this, giving teachers the confidence to try new approaches. Teachers were trusted to create assessments that authentically reflected classroom learning, while also being supported by senior teachers.

Chapter 7 presents the analysis of stakeholder influence in the initial interpretation of the Assessment with Levels (AwL) policy and 2-3 years after its enactment, co-produced with the school's headteacher. Stakeholders include not only leaders and teachers but also pupils, who are considered policy actors in translating and enacting policy in school. Section A aims to clarify the complexity of the policy enactment process and identify the stakeholders involved. The AwL policy serves as a case study to illustrate the processes of (re-)interpretation, (re-)translation, and (re-)enactment in a primary school. Data was generated using a hierarchical mapping technique during interviews with six headteachers between October 2017 and January 2018. This technique helped understand how headteachers created policy spaces and established mobility for the AwL policy by engaging with school stakeholders. Headteachers placed stakeholders on a map according to their influence during the interpretation stage and after policy enactment. This mapping was part

of a semi-structured qualitative interview using a thinking-aloud approach, creating both visual and verbal data. The analysis reveals that headteachers faced uncertainty in interpreting what AwL should entail, leading them to rely on expert stakeholders such as themselves and subject leaders. It also shows that headteachers have the power to invite and uninvite stakeholders and coordinate them. A reflection on this dynamic of headteacher coordination is presented. The second part of the chapter uses Ball et al.'s (2012) framework to understand the types of policy roles and work different stakeholders play. Headteachers must distribute leadership, as teachers and other stakeholders engage in various policy work. The chapter reflects on Ball et al.'s (2012) typology of policy actors and policy work, suggesting some alterations based on the data from this thesis.

Chapter 8 argues that by focusing on the enactment of the AwL policy as a single case study, with six primary schools as sub-units, the research offers a nuanced exploration of the policy, the schools, and key themes in the reform of pedagogy and assessment in England during the early 21st century. This design combines the depth of a single case study with the comparative benefits of a multi-case study, allowing for a deeper understanding and exploration of theoretical relationships. The study's design preserves the benefits of a single case study while gaining insights from multiple schools enabling a contrast of similarities and differences in data across contexts, aiding in the exploration of the research questions. The thesis highlights individual voices and specific details, presenting observable interconnections within real-life school settings. The qualitative approach provides a rich description of AwL and AfL, immersing the reader in the technical details and experiences surrounding the policy's enactment. This thesis incorporates a longitudinal element by collecting data at two points in the enactment process. This approach captures how experiences and perspectives evolve over time, providing insights into the transitions from initial interpretation to more embedded stages of policy implementation. The study explores how individuals adapt to changes and assesses the success of new policies. The enactment of AwL is a critical case, representing broader trends in primary schools in England between 2014 and 2019. The case study also serves as a paradigmatic example, illustrating how AwL reformed assessment approaches and embedded AfL into pedagogy nationwide. The thesis examines the paradigm of school practices and stakeholders' worldviews. The study emphasizes the importance of understanding habitual practices to determine a paradigm, noting that these practices reflect the true nature of educational processes. The research also considers the constructed perceptions and ideas of stakeholders, particularly headteachers, which guide their understanding of reality and influence their professional actions. This thesis provides a comprehensive exploration of the enactment of AwL, offering valuable insights into the reform of assessment and pedagogy in primary schools. It

combines detailed qualitative analysis with a broader understanding of educational paradigms, contributing to the field of educational research. The chapter proceeds to reflect on the successful implementation of AwL and suggests four factors that have enabled this to happen: 1) Clear communication of policy intentions; 2) Latitude for the Headteacher; 3) Teachers observing benefits for pupils; 4) Strong National Accountability Processes. Although much has been written about sensemaking and also that it is situated, less has been written on the specifics of the influences within those situations and these are presented in Chapter 8. Even where scholars have explored the influences on sensemaking none seem to have considered whether these influences can be categorised into frames. Using this concept of the frame and an inductive approach to analysing interview data, the findings below are presented in three frames of sensemaking: social, values, and accountability. The social influences come from multiple actors some of whom are local to the school, such as nearby schools with whom the participant school has had a long running relationship. The Commission on Assessment without Levels Report (CAWL 2015) encouraged teachers to use their values and principles on the purposes of assessment as a starting point in the process of AwL sensemaking and some of these are presented. The third is the accountability frame which argues that although AwL has allowed schools to exercise a high level of autonomy, it is tempered by engagement in accountability processes and also thoughts about what accountability processes may require.

Chapter 9 of the thesis offers a comprehensive set of recommendations for future research on various educational strategies and their impacts. It emphasizes the importance of exploring continuous dialogue and verbal feedback in classroom settings, particularly focusing on Schools PATs and Qualitative's AwL approaches. Future studies should investigate how these methods affect pupil engagement and learning outcomes, especially for those with Special Educational Needs and Disabilities (SEND). Additionally, the benefits of interactive classroom discussions on critical thinking, communication skills, and collaborative learning should be explored. The concept of the 'danger zone,' where pupils are encouraged to embrace challenges, warrants examination for its effects on motivation and resilience. Constructive feedback's role in fostering a growth mindset and the impact of teaching GRIT (Growth mindset, Resilience, Independence, and Teamwork) on student development are also crucial areas for research. The dual-book system for main work and skills practice, along with self-evaluation sheets, supports reflective learning and requires further investigation. The balance between independent and collaborative learning, and the effectiveness of verbal feedback groups in enhancing student engagement and understanding, should be explored. The chapter also recommends further research on the implementation of Key Performance Indicators (KPIs) in primary schools, focusing on their

impact on teacher workload and well-being. Studies should investigate the optimal level of KPI detail to support effective teaching without causing burnout. Additionally, research is needed to evaluate the effectiveness of different attainment thresholds and their impact on pupil progress. The direct influence of KPIs on student outcomes, engagement, and well-being should be explored, along with the perspectives of teachers, parents, and students. Long-term studies are essential to understand the sustained impact of KPIs on student progress and institutional performance. Further research on mastery-based learning is recommended, emphasizing its potential to allow students to progress at their own pace, advancing only upon mastering the material. Studies should explore how personalized, student-centred learning and formative assessments impact student progress and self-efficacy. The role of standards-based grading in accurately reflecting student understanding and providing targeted support warrants investigation. Research should focus on developing curriculum design principles that balance individualized pacing with standardized curricula, and on evaluating teacher training programs for effective mastery learning implementation. Additionally, the impact of various formative assessments on student learning should be examined. Equity and access are crucial, necessitating studies on supporting disadvantaged students. Longitudinal research is needed to understand the long-term effects of mastery learning on academic and career success. The integration of technology in mastery learning environments and its impact on student engagement and outcomes should be explored. Finally, identifying best practices for implementing mastery learning across different educational contexts is essential. The chapter also recommends further research on whether curriculum opportunities have been widened in primary schools, particularly in response to the Education Inspection Framework (EIF) introduced in 2019. The EIF emphasizes a broad and balanced curriculum, reducing the focus on standardized testing and highlighting the importance of diverse subjects. The chapter recommends investigating the impact of granting schools more autonomy in curriculum design and the role of school leaders in driving innovation. It also suggests examining the professional development needs of teachers to deliver a comprehensive curriculum. Longitudinal studies are recommended to understand the long-term effects of a narrowed curriculum on students' critical thinking, creativity, and preparedness for secondary education. Comparative studies across different countries can provide insights into effective curriculum practices. Additionally, the chapter recommends exploring the impact of curriculum breadth on student and teacher well-being, the integration of arts, humanities, and physical education into daily schedules, and the effectiveness of alternative assessment methods. Policy analysis is suggested to identify strengths and weaknesses in current curriculum policies and propose recommendations for reform. Further research on the relationships between headteachers and school Governing Bodies (GB) is recommended, focusing on their influence on the implementation of

Assessment without Levels (AwL) systems. The chapter highlights the impact of governor expertise on school leadership and decision-making processes. It suggests investigating the training and development needs of governors to support headteachers effectively. The dynamics of role boundaries between headteachers and governors are also examined, with research needed to explore how these boundaries affect collaboration and policy implementation. The involvement of governors in designing and implementing assessment systems is another critical area, with studies recommended to identify best practices for engaging governors. Conflict resolution strategies between headteachers and governors should be explored to enhance school leadership and governance. The chapter also recommends investigating the influence of external pressures, such as local authority recommendations and Ofsted inspections, on headteacher-governor relationships. Finally, case studies of schools with strong headteacher-governor relationships are suggested to identify best practices and key factors for successful partnerships. Lastly, the chapter recommends further research on the interactive visual method based on Antonucci's (1986) hierarchical mapping technique, used to access headteachers' memories on policy interpretation and enactment. The chapter suggests investigating the cognitive processes involved in using visual methods, such as how visual aids trigger memory recall and help participants organize their thoughts. Comparing visual methods with traditional interview techniques is recommended to determine differences in data quality and participant engagement. Further research should explore the long-term impact of visual methods on policy implementation and leadership practices through longitudinal studies. Additionally, the application of visual methods in diverse educational contexts and other sectors, such as healthcare and business, should be examined. Understanding how visual methods influence leaders' ongoing interpretation and enactment of policies can provide valuable insights into effective leadership and organizational change. Finally, the chapter recommends exploring different types of visual methods to enhance participant engagement and data quality in research interviews.

## Chapter 2: Literature Review

This literature review has six parts. The first explains the changing landscape of schooling in the 50 years prior to AwL to assist the reader's understanding of the debates that have formed the environment within which AwL is being enacted. AwL existed previously as schools had had autonomy prior to 1988 to determine their own curricula and how progress was captured but the policy and cultural environment was significantly different to 2015-2019 when the current version of AwL was enacted. The second part explains the policies underpinning AwL with an explanation of the rationale for their removal to assist in the enactment of mastery learning, and the third part outlines some of the learning theories that are considered to be part of a mastery learning approach. The fourth part of this review describes some of the AwL approaches that were known prior to Phase 1 data collection which was conducted between November 2017 and January 2018. Part five explains the policy enactment conceptual framework adopted in this thesis which was developed by Maguire et al. (2015), and part six details two studies that added to the researchers understanding prior to Phase 2 data collection which was conducted in March 2019. The first of these studies authored by Poet et al. (2018) was commissioned by the Department of Education to provide an overview of what AwL approaches are being used in both secondary and primary schools, and the second is by Thrupp et al. (2018) which, although not specifically on AwL, details the enactment of Kiwi standards in six New Zealand primary schools between 2010-2013 and has parallels with this DPhil study.

### 2.1 The Landscape of Schooling in England from 1965-2015

The school system in England is built on a complex blend of trust and autonomy, alongside mistrust and control, which permeates the curriculum, pedagogy, structures, and accountability systems. It is worth noting that AwL existed before 1988 although it would not have been termed as such because National Curriculum Levels did not exist, and the educational landscape was significantly different in terms of structures, relationships, teacher identity, and control. In the 1970s and 1980s, primary teachers enjoyed considerable autonomy over the curriculum and pedagogy. Former primary headteacher, Tony Eade (2015), reminisced that, aside from breaks, assemblies, and hall times, teachers had the freedom to organize their timetables as they saw fit. He noted that many opted for an integrated day, avoiding the division of the timetable into separate subjects and minimizing detailed lesson planning. Children were encouraged to develop at their own pace through a variety of experiences, as there were no external tests, and some teachers chose not to test

at all. Additionally, Her Majesty's Inspectorate (HMI) did not evaluate data (Eaude 2015). Schools faced minimal external pressure, such as league tables, and experienced almost no structural changes. There was no national curriculum, no parental choice, no systematic performance monitoring, and no publication of exam results beyond what schools chose to disclose. Schools were part of robust local authority structures that determined catchment areas, decided which children attended which schools, and allocated funds, specifying expenditures on teachers' salaries, repairs, books, and other items. Eaude (2015) crystallised the atmosphere within the teaching profession at the time, 'It now seems remarkable the extent to which teachers and schools were trusted and given autonomy, how untouched by external events most teachers were' (Eaude 2015:44).

Wilby (2013) explained that Margaret Thatcher took away from her four years as Education Minister (1970-74), 'an abiding hatred of its culture: of "self-righteously socialist" civil servants, of academics who "pounded" every "decent value" out of students' minds, of "trendy" teachers who didn't inculcate the three Rs, of local authorities she couldn't control.' (Wilby 2013 in Guardian Newspaper). When she was elected Prime Minister in 1979, she implemented policies to control the education system by linking funding to the pupil, allowing parents to choose their school. This approach encouraged schools to compete for students, creating a semi-quasi-market (West 2015). The idea was that schools would either improve due to competitive pressures and the adoption of best practices, or they would close, with students being reassigned to better-performing (popular) schools. The underlying assumption of the marketisation of education was that increased competitive incentives would lead to higher education quality. Independence from local authorities was achieved with the introduction of Local Management of Schools, which permitted schools to manage their own budgets. Primary schools tried to maintain the role of the local authority; however, the latter's approach became more inspectoral (Eaude 2018).

The National Curriculum introduced in 1988 brought in very detailed, subject-specific requirements, and Standard Assessment Tests (SATs) were implemented in English, mathematics, and science for seven and eleven-year-olds, combining tests and Individual Teacher Assessments. Level descriptors were established as a national measure of progress and attainment, requiring teachers to adapt their teaching methods to meet the new curriculum and testing demands aimed at raising attainment standards. Results from national tests were compiled and published, and the Office for Standards in Education (Ofsted) replaced HMI, examining progress and attainment data in their inspection reports, which were also made public. Eaude (2015) noted that this inspection regime altered

teaching and learning, as headteachers paid more attention to meeting the demands of inspection teams that emphasized written policies. Ball (2003) referred to this as 'performativity,' where Ofsted judgments were heavily influenced by pupil data assessments, making the stakes of assessment very high.

The New Labour government (1997-2010) maintained the National Curriculum and level descriptors, along with accountability measures such as Ofsted, SATs, and the publication of league tables, continuing a deficit model in their approach to schooling. Oates (2015) remarked that until then, the concept of failing schools, headteachers, teachers, and pupils did not exist. He argued that New Labour's education policy focused on identifying failure and then implementing strategies to address it. For instance, schools deemed to be failing could have their governing structure altered to become autonomous entities known as academies. Thus, autonomy was being reintroduced into the school system, but it was based on the identification of failure. The New Labour government believed primary teachers were pedagogically deficient, leading to the introduction of the National Literacy and Numeracy Strategies in 1998 and 1999, respectively, which prescribed not only content but also teaching methods and national standards for pupils (Eaude 2015). The Thatcher government had started the process of increasing choice through Grant Maintained Schools, which were independent of local authorities, but it was New Labour that began privatizing provision and financing, particularly for schools deemed to be failing (Oates 2015). Sahlgren (2014) explained that privatisation can only succeed if the right competitive incentives exist within schools, as merely replacing public owners with private ones is unlikely to bring about positive change. Therefore, New Labour allowed academy sponsors to freely enter and exit the market and encouraged schools to scale up and capitalize on economies of scale. The New Labour government (1997-2010) invested significantly in developing tools for tracking pupil progress and attainment to help teachers use the results of statutory national tests for monitoring, prediction, and target setting. Pilot studies were conducted under the 'Making Good Progress' (MGP) project, and a set of materials called Assessing Pupils' Progress (APP) was published, which used Assessment Focus (AF). APP aimed to provide a more detailed assessment framework than levels, with one to three bullet points for each AF identifying what teachers should look for in pupils' classroom work (DCSF 2009). Teachers were in the process of implementing this approach when the Coalition government came to power and initiated several reviews of education.

A change of government in 2010 prevented the recommendations of the Independent Review of the Primary Curriculum (Department of Children, Schools and Families 2010) from being implemented. This review argued that the primary national curriculum should place

less emphasis on subject areas and more on areas of learning and development. It drew on examples from several jurisdictions, such as Finland, which had moved away from emphasizing knowledge at the time it reached the summit of the international education league tables, aiming to increase autonomy. It described the Finnish position as follows: 'Core content is described as activities and skills, rather than detailed subject-based content. This places the onus on the municipality, and more importantly on the school, to develop their curriculum to meet learners' needs as well as national expectations' (Department of Children, Schools and Families 2010:116). The OECD had reviewed this international test data found that successful school systems spend large amounts of money on education and tend to prioritise teachers' pay and status (OECD 2010a). They also help children to apply knowledge rather than just recall and repeat it and prioritise improving the provision for low performing students (OECD 2010).

The Coalition government's (2010-15) and Conservative governments' (2015-2024) reforms to curriculum and assessment stemmed from their opposition to the above perspectives. They argued that the curriculum should focus on the delivery of knowledge - an approach they claimed existed within fee-paying private schools - rather than a skills focus, which they perceived as detrimental to learning. This type of comment has connotations of Wiener's (1981) perception of an educational system with an aristocratic style of culture. For example, Schools Minister Nick Gibb subscribed to the arguments made by E.D. Hirsch, Jr. who is the founder of the Core Knowledge Foundation in the USA and based on Hirsch's research Gibb argued that for 'real social justice and for social mobility to occur, all pupils must have access to the rigorous curricula that characterise our world-renowned independent schools' (Gibb 2017). The government seemed to be attempting to reduce the focus on higher-order thinking skills such as problem-solving and critical thinking, which form part of what has become known as 21st-century skills (Gibb 2016). As Gibb (2016) explained,

'It is an unfortunate fact, however, that many modern conceptions of education either ignore the importance of knowledge, or actively deride it. During the 1960s, it became fashionable amongst educationists to dismiss the accumulation of knowledge as a joyless anachronism: rote learning of unconnected facts, inflicted upon bored and unwilling pupils. School curricula were increasingly rewritten to focus not upon subject content, but upon skills and dispositions...[which] had a marked effect on the quality of education that generations of children have received in Britain. For me, the crowning glory of this dumbing down was the 2007 rewrite of the national curriculum, which systematically expunged any mention of subject content, replacing it with

references to “processes”, “concepts”, and with an overlay of “personal, learning and thinking skills” such as “independent learning” and “learning to learn” (Gibb 2016).

Gibb was dismissive of those who disagreed with him, such as Sir Ken Robinson, whose views on creativity within education he called ‘superficially appealing... [but] profoundly wrong’ (Gibb 2016). He also derided suggestions that education should be led by curiosity as a new version of the ‘disastrous child-centred approach of yesterday’ (Gibb 2015a). He criticized those who encouraged independent learning from the internet, citing evidence from the OECD which showed that increased internet use in schools does not lead to higher academic outcomes (OECD 2016). Gibb described all criticisms of his position as ‘anti-knowledge’ and ‘anti-evidence’, and instead placed his faith in the work based on ‘robust scientific evidence’ (Gibb 2016) of cognitive scientists such as Daniel Willingham. Willingham had expressed scepticism over the 21st-century skills agenda based on what he perceived as a neglect of domain knowledge, arguing that 21st-century skills were a relatively small part of what students needed to learn (Rotherham and Willingham 2009). A problem for Gibb’s contention is that Willingham (2017) later admitted he was wrong and accepted the findings of McGrew et al. (2016) that a content-free strategy could make a big difference in student assessment of website accuracy. Willingham (2017) acknowledged that his assertion that 21st-century skills would be a small part of what a student needs to know was inaccurate given the centrality of the internet to education. The OECD developed a project called ‘The Future of Education and Skills: OECD Education 2030 Framework’ which aimed to change the traditional disciplinary curriculum to incorporate 21st-century skills (OECD 2016a).

Gibb justified his position on the basis that England suffered what he referred to as a ‘PISA shock’ in 2007, when it was ranked 14th out of 57 countries in science, 17th in reading, and 24th in mathematics (Gibb 2015a). He described the subsequent consistency of mean scores as evidence of the failure of the 21st-century skills agenda. However, data is often contradictory, and the Trends in Mathematics and Science Study (TIMSS) 2015 found that pupils in England performed, on average, significantly above the international mean in maths and science in both years 5 and 9, and that there was evidence that the system had improved over the last 20 years (Greany 2016). Also, PISA only measures some fifteen-year-olds, and only in reading and not writing, and the surveys leave out many things like education for entrepreneurial skills, civic responsibility and participation, social cohesion, independent thinking, justice and enhanced equity, critical thinking, and justice (Rogers 2014). Thus, to argue that PISA justifies a change to the primary school curriculum seems to

be a use of international test data for policy legitimation (Lawn and Grek 2012). Gibb (2017a) spoke at the launch of a pamphlet titled *The Question of Knowledge* produced by the Association of School and College Leaders (ASCL) and Parents and Teachers for Excellence (PTE), which describes itself as a new movement to promote reforms within the education system. At the launch, Gibb emphasized that, 'pedagogy, assessment, tracking and qualifications must lead on from us developing further our understanding of what makes a pupil knowledgeable' (Gibb 2017a).

Coupled with the knowledge-focused curriculum, schools were encouraged to become academies which operate under direct funding agreements with central government freeing them from Local Authority (LA) control. This resulted in new forms of collaborative school partnerships such as chains, federations and teaching school alliances (Sandals and Bryant 2014). Allowing each school to interpret AwL is influenced by this landscape as the Education Select Committee commented, 'School autonomy and increased flexibility is one aim of the removal of levels' (House of Commons Education Committee Primary Assessment 2017:13). Michael Gove, Secretary of State of Education (2010-13) aimed to create a self-improving schools system (Hargreaves 2014), 'guided by the powers of ambition, competition and choice' (Dennis 2016:1). These ideological principles of 'a free-market libertarian conservative' (Coiffait 2015:140) underpin the atomisation of governance structures and therefore influence that manner in-which the removal of levels was presented to schools as an opportunity to exercise their own assessment principles.

## 2.2 Policies Underpinning AwL

Since the introduction of the National Curriculum in 1988, level descriptors had been used by teachers to determine a pupil's attainment by judging which description best fits the pupil's performance. When doing so, each description was considered alongside descriptions for adjacent levels. Table 2 outlines the descriptions for Levels 1-3 published by the Department for Education (DfE) in 2010 under the section titled 'Using and applying mathematics':

Table 2: National Curriculum Level Descriptors, 'Using and Applying Mathematics' (DfE 2010:31)

Level 1	Pupils use mathematics as an integral part of classroom activities. They represent their work with objects or pictures and discuss it. They recognise and use a simple pattern or relationship.
Level 2	Pupils select the mathematics they use in some classroom activities. They discuss their work using mathematical language and are beginning to represent it using symbols and simple diagrams. They explain why an answer is correct.
Level 3	Pupils try different approaches and find ways of overcoming difficulties that arise when they are solving problems. They are beginning to organise their work and check results. Pupils discuss their mathematical work and are beginning to explain their thinking. They use and interpret mathematical symbols and diagrams. Pupils show that they understand a general statement by finding particular examples that match it.

The source of AwL is Chapter 8 of the Report of the Expert Panel for the National Curriculum Review 2011 (The Panel) who concluded that the system of levels,

“weakens teachers clear understanding and identification of pupils’ specific weaknesses or misunderstandings... that current ‘generalised’ reporting using levels obscures the fact that too great a proportion of pupils fail to attain elements of the curriculum that are vital for the next phase of their education.”  
(Expert Panel Report 2011:50)

The Panel identified that levels were not consistent with a mastery approach because they encouraged undue pace and progression onto more difficult work while pupils still had gaps in their knowledge or understanding. Mastery denotes a focus on achieving a deeper understanding of fewer topics through problem-solving, questioning and encouraging deep thinking about a topic. Mastery learning is a specific approach in which learning is broken down into discrete units and presented in logical order. Pupils are required to demonstrate mastery of the learning from each unit before being allowed to move on to the next, with the assumption that all pupils will achieve this level of mastery if they are appropriately supported, ‘Some may take longer and need more help, but all will get there in the end’ (Expert Panel Report 2015:17). Thus, a mastery approach incorporates a belief that all children can achieve a high standard, and that the purpose of assessment is not differentiation but ensuring all children have grasped content (Expert Panel Report 2015). The new National Curriculum introduced in England in 2015 was premised on this kind of understanding of mastery. It claimed to focus upon deep, secure learning for all, with

extension activities on the same topic for more able students rather than rapid movement to new content. There is some evidence in grey literature that this mastery approach is influencing how schools are interpreting and enacting AwL. In the Times Educational Supplement (TES) Online Toolkit, schoolteacher Jenni Willis described her experience of leading her school in the development of an AwL system. The fundamental principle that they decided upon was that mastery of fewer concepts was more important than a shallow understanding of a broad curriculum. As Willis comments, 'Less but better was our mantra' (Willis 2016). To illustrate some of the types of AwL systems that teachers may have adopted, the next part of this literature review provides an overview of some popular AwL approaches based upon a mastery approach. Following instruction, pupils undertake formative assessment that shows what they have learned well and what they still need to work on, and identifies activities to help them do this followed by a parallel assessment. The results of international assessments (such as TIMSS and PISA) have shown jurisdictions using such mastery approaches demonstrate excellent performance in mathematics (OECD 2014). Some research has also demonstrated the effectiveness of programmes that follow the mastery approach (Wiley et al. 2009), and in a report from August 2017, the UK Education Endowment Foundation (EEF) found it to be particularly effective when pupils work in groups or teams and take responsibility for supporting each other's progress (particularly for lower attaining students). However, they also posit that mastery learning may also be more effective when used as an occasional or additional teaching strategy as the impact decreases for longer programmes of over 12 weeks or so (Education Endowment Foundation 2017).

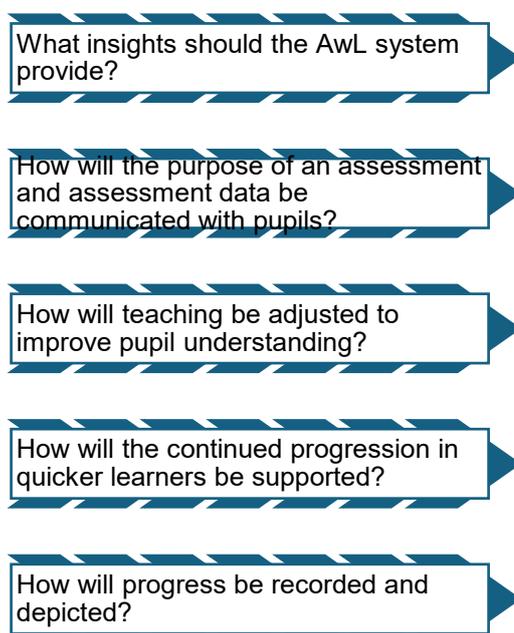
In September 2014, the National College of Teaching and Learning (NCTL) published research conducted by 34 teaching school alliances on alternative assessment approaches to levels. The research identified three main priorities: developing tools to support individual progress, capturing and recording progress, and using technology to track attainment. In terms of assessment tools to support individual progress, the report highlighted the use of visible learning approaches, which emphasised making teaching and learning explicit to increase student achievement. Additionally, the revised Bloom's Taxonomy was employed as a two-dimensional framework to classify objectives, activities, and assessments. Mastery statements, such as 'I can' grids, were used to involve students in formative assessment, while curriculum progression objectives provided detailed objectives for core subjects to align assessment closely with the new curriculum. Various feedback methods were also discussed to help students understand how to improve their work. To capture progress, the report discussed the use of SOLO Taxonomy grids to capture progressive development in understanding and skills. The revised Bloom's Taxonomy was applied to create tools like

progress ladders and command word posters. Mastery statements ensured progress was on a consistent trajectory, and marking grids, such as 'entering, developed, secure,' were used to track progress. Summative testing, including frequent low-stakes tests, informed progress, and the practice of working backwards from GCSE expectations was used to form expectations from Year 7 upwards. The importance of moderation within and between schools was also emphasised. The development of software for tracking progress was another key area explored in the report. Various technological options were considered, with SIMS emerging as a popular tracking tool that met the necessary criteria for assessment software. The National College of Teaching and Leadership Report (2014) also discussed the outcomes and impact of these assessment approaches on schools. It noted an increase in professional dialogue among teachers and school leaders, leading to a greater understanding of the purposes of assessment. The involvement of pupils in the assessment process was emphasized, as was the need for subject-specific assessment models. Summative assessment, particularly low-stakes testing, was seen as a motivating way to ensure progress. Schools experienced a shift in assessment culture, focusing more on student progress and understanding, and more teachers became involved in research, appreciating its significance. The report made several recommendations. It called for a shift in the nature, range, and purposes of assessment, the organization of national conferences and seminars to develop assessment expertise, the development of peer-reviewed commercial tracking systems, financial support for alliances to develop assessment practices, and the encouragement of teachers to be research-active.

Scepticism from many headteachers about the likelihood that the government would persist in their intention to not provide a national standardised system prompted the Department for Education to appoint a Commission of Assessment without Levels (CAWL) which reported in 2015 and emphasised that levels would not be replaced by a national system and encouraged headteachers to exercise autonomy while providing some scaffolding in the form of reflection questions (Figure 1). The first question asked headteachers to consider what insights the assessment should provide to discern the security of the students' knowledge and understanding of the subject matter, concept, or skill and whether they can apply that with a growing independence. To ensure inclusivity for pupils with Special Educational Needs and Disabilities (SEND), thought should be given on how they can demonstrate their learning. Secondly, headteachers were asked to reflect upon how they would communicate the purpose of an assessment and also the assessment data with pupils so that they know their next steps in learning. Thirdly, headteachers were asked to encourage teachers to reflect on how they could adjust teaching to improve pupil understanding and to consider whether some pupils need extra teaching or practice.

Fourthly, headteachers were asked to consider how to support progression where learning is secure in quicker learners as they should not be taught outside of the topic being taught to the rest of the class. This might involve assessing whether pupils can apply the concept in different contexts or to explore a concept in greater depth before moving on to new material. Fifth, headteachers were asked to consider how to record and depict progress. This chapter captures how AwL was interpreted by presenting the narratives that headteachers used to retrospectively organise the events and understandings that underpinned policy enactment decisions.

Figure 1: Reflection Questions provided by CAWL (2015)



The Expert Panel Report (2011), National College of Teaching and Leadership (NCTL) Report (2014), and CAWL Report (2015) collectively shaped teachers' understanding and choices regarding Assessment without Levels (AwL) systems. The Expert Panel Report criticised the levels system for not supporting deep learning and highlighted the need for a mastery approach, which focuses on thorough understanding and problem-solving. This report laid the foundation for the shift towards mastery-based assessments. The NCTL Report (2014) provided practical insights from teaching school alliance schools on alternative assessment approaches. It emphasised tools to support individual progress, visible learning, and the use of technology for tracking attainment. This report offered concrete examples and strategies for implementing mastery-based assessments, reinforcing

the principles outlined by the Expert Panel. The CAWL Report (2015) further supported this transition by emphasising headteacher autonomy and reflective practices. It provided guidance on how to ensure assessments accurately reflect students' understanding and progress, encouraging a more personalised and inclusive approach. Together, these reports informed teachers about the benefits of mastery-based assessments and provided practical frameworks and tools for implementation. This combined understanding helped teachers choose AwL systems that focus on deep learning, individual progress, and effective use of technology.

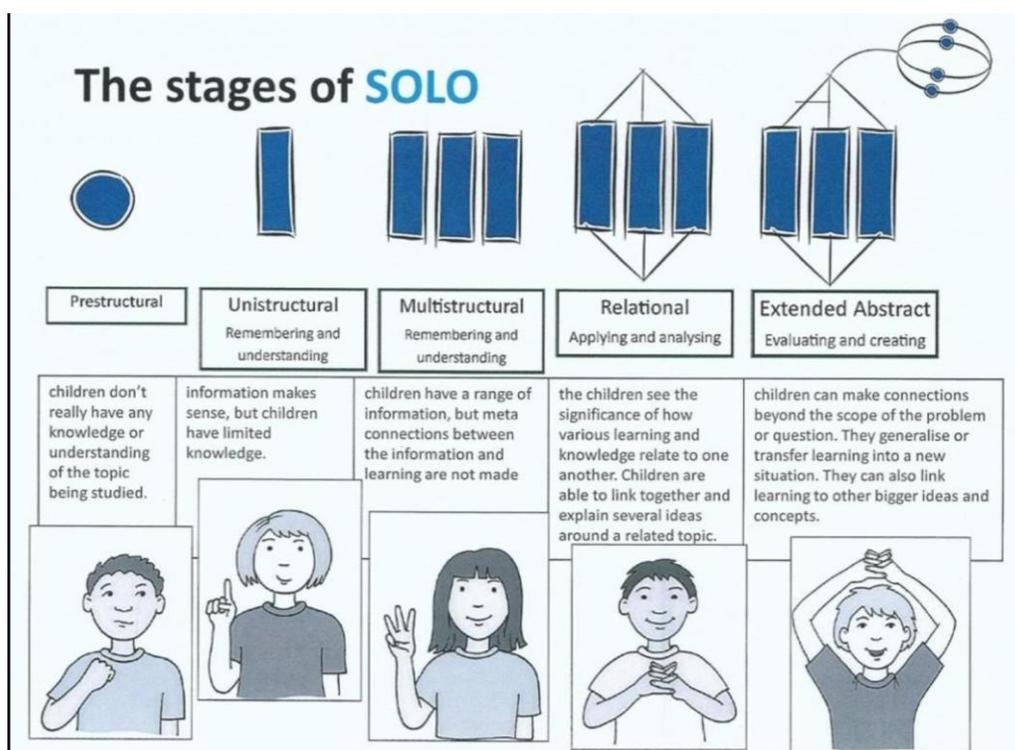
### 2.3 What was known about AwL systems prior to Phase 1 Data Collection

Brown and Wilson (2011) found that most assessment approaches examined in their study lacked an explicit model of cognition or a theory about how students represent knowledge and develop competence in a subject domain. They argue that without a model of cognition, assessment designers, including class teachers, are handicapped by a largely implicit knowledge of how understanding develops and no clear guidance on how to create meaningful assessments. Andrade (2013) explains that a careful study of the learning theory and assessment can inform the design of assessment processes that are grounded on research as to how students represent knowledge and develop competence, while also providing an interpretative lens on the intended and unintended consequences of assessment. In the USA, the National Research Council (NRC) Report 'Knowing what Students Know' (2001) described three necessary components of a valid assessment system:

- 'a model of student cognition and learning in the domain
- a set of beliefs about the kinds of observations that will provide evidence of student's competencies
- an interpretation process for making sense of the evidence' (National Research Council 2001:44).

This was considered by the Teaching Schools that participated in National College of Teaching and Leadership, 'Beyond Levels' research, and they used the SOLO Taxonomy as supported by the materials created by Hook (2011, 2012, 2013) for the New Zealand context (Figure 2).

Figure 2: Hook's SOLO Taxonomy presented in National College of Teaching and Leadership Report (2014)



Mastery approaches are underpinned by cumulation, which it is argued maximises student achievement by teaching all students everything (Alexander 2004). Cumulation requires the teacher to match teaching to the learner while respecting the validation of the subject being taught. The teacher has to provide an individually tailored response to a pupil's answer that advances learning (Alexander 2004). As one of the members of The Panel, Dylan Wiliam (2011) explains,

'When the assessment is cumulative, there is no incentive for students (and teachers) to adopt a shallow approach, because if material is forgotten, it has to be learned again, because it is going to be assessed again. The important thing is that the assessment system, as far as possible, creates positive incentives for teachers to teach well, and for students to study well' (Wiliam 2011:3).

Allal (2010) noted that every theory of learning includes a mechanism of regulation of learners' thinking, affect, and behaviour. He explained that the regulation of learning involves four main processes: (1) goal setting, (2) the monitoring of progress toward the goal, (3) interpretation of feedback derived from monitoring, and (4) adjustment of goal-directed

action including, perhaps, redefining the goal itself (Allal 2010)<sup>2</sup>. Cumulative approaches may incorporate various learning theories and below is a brief outline of three popular ones: learning progressions; learning Intentions and success criteria; and vertical scaling.

Learning Progression is a model of successively more sophisticated ways of thinking about a topic with more detailed steps of intermediate learning and can also indicate common pre- and misconceptions students have about a topic (Alonzo and Steedle 2009). They provide teachers with a goal for summative assessment, indicate a sequence of activities for instruction (Figure 3), and can inform the design of formative assessment processes that provide indicators of students' understanding (Corcoran, Mosher and Rogat 2009).

Figure 3: An example of a Learning Progression Guide (Alonzo and Steedle 2009:16-17)

<p>Level 2: The student believes that motion implies a force in the direction of motion and that non-motion implies no force. Conversely, the student believes that force implies motion in the direction of the force.</p> <p>Level 2A: The force acting on an object could be the initial force (which is carried with the object and may dissipate over time).</p> <p>No Force: If no force is acting upon an object, it is not moving.</p> <p>Motion: If an object is moving, a force is acting upon it.</p> <p>No Motion: If an object is not moving, no force is acting upon it.</p> <p>Common Errors:</p> <ul style="list-style-type: none"><li>• If there is no motion, there are no forces acting.</li><li>• When an object is moving, there is a force in the direction of its motion.<ul style="list-style-type: none"><li>◦ 2A: This motion could be the force that put the object into motion initially.</li><li>◦ 2A: This object may come to rest because the force it carries with it has been used up.</li></ul></li></ul>
---

In contrast with learning progressions which tend to be broad, learning intentions/objectives and success criteria specifically describe the qualities of excellent student work. These can be for a particular lesson or for a longer period of work such as a key stage. In his review of 800 meta-analyses, Hattie (2009) concluded that successful learning is improved by providing clear learning intentions that describe the skills, knowledge, attitudes and values that the student needs to learn. He also found that sharing success criteria that help

---

<sup>2</sup> Depending on the data collected, a discussion on metacognition may be necessary



arbitrary definition by teachers who would require a level of expertise to determine this.

McBride and Wise (2000) suggest a focus on the following questions to establish validity:

- How does the content to be mastered at each succeeding grade relate to and extend the content to be mastered at the prior grade?
- How does the assessment of content mastery compare from one grade to the next?
- Is there a logical way to assess mastery of the cumulative content?

Figure 5: The National College and Teaching and Leadership Report (2014) suggested points-based scale for AwL

Year	Emerging			Expected			Exceeding		
N	1			2			3		
R	4			5			6		
1	7	8	9	10	11	12	13	14	15
2	10	11	12	13	14	15	16	17	18
3	13	14	15	16	17	18	19	20	21
4	16	17	18	19	20	21	22	23	24
5	19	20	21	22	23	24	25	26	27
6	22	23	24	25	26	27	28	29	30

In May 2015, deputy headteacher, Times Educational Supplement columnist, and blogger, Michael Tidd, used Twitter and his blog to explore what systems were in use by teachers who followed him. Approximately one-third used Target Tracker or CapitaSIMS.net with only 6% stating that they had designed their own system. A significant proportion had not transitioned to an AwL system at that time. Target Tracker and CapitaSIMS.net were prominent in Tidd's (2015) study as they had been in existence for some years prior to AwL and covered the entire curriculum. Target Tracker was developed by Essex County Council in response to teachers requesting assistance in the process of assessing, benchmarking and targeting pupils against National Curriculum guidelines. Staff entered data termly, matching pupils' performance against a library of 'I can' statements. Tracking grids and graphs were provided that show how a pupil's attainment compares to national standards. Progress could be tracked over an extended period showing a pupil's strongest year and when they made most improvement, as well as whole class data and teachers could also use the software to produce pupil reports for parents. CapitaSIMS.net assessment is provided by business services company Capita Plc. The package gave teachers access to the entire national curriculum programmes of study including both the statutory and non-statutory guidance. Teachers could also record formative assessments against any statement of the National Curriculum or provide summative judgements for pupils to indicate which phase of the curriculum they are working within for any subject or strand area. Instant

analysis was displayed on the data entry screen so teachers could view pupil and class progress towards end of year expectations (SIMS Assessment 2017).

In 2014, the National College for Teaching and Leadership (NCTL) Report highlighted several electronic Assessment without Levels (AwL) approaches that were being utilized by Teaching Schools Alliances (TSA). These schools explored a variety of technological options to enhance their assessment practices. Some tools were designed to capture formative assessments and share them with parents, while others focused on tracking student progress. Bishop Challoner TSA brought attention to four software applications for AwL. PiXle software emerged as the most popular due to its ability to group subjects into domains and align with the new national curriculum grading arrangements, making it particularly attractive because of its consistency with GCSE standards. EDlounge software, which featured 'I can...' statements, was well-received by primary schools but found lacking for secondary schools that required more technical and subject-specific terminology. Flightpath software, although showing potential, was considered too prescriptive in defining progress trajectories. Ultimately, all schools adopted SIMS for tracking, as it met the alliance's criteria for benefit, compatibility, simplicity, and trialability. Wigmore TSA utilized SIMS to develop a system for recording pupil attainment, reflecting the new curriculum's increased expectations. Primary Excellence TSA employed technology for both formative and summative assessments, using tools like dictaphones and flip cameras to capture pupil reflections and sound files to document student opinions. Silk Alliance collaborated with an associate partner school to develop a cloud-based assessment application for iOS platforms, aligned with national curriculum expectations. Lightwoods TSA created bespoke assessment software to capture evidence of children's progress, communicate with parents, and plan lessons using skills ladders. The e-profiles generated by this software provided a detailed picture of assessment, showcasing the learning process and journey. These profiles allowed for internal and external moderation and could be shared with parents via learning platforms, enabling children to reflect on and assess their own learning.

The overarching purposes of formative assessment technologies are (1) systematic monitoring of student progress to inform instructional decisions; (2) identification of misconceptions that may interfere with student learning; (3) rapid feedback to students, teachers, and others; and (4) information about student learning needs during instruction (Russell 2010). Computer-based assessment programs integrate the management of learning, curricular resources, embedded assessments, and detailed student-level and aggregate-level reporting of strengths and weaknesses. There are many products on the market and selecting the correct one for each context can be complex. There is often an

assumption that technologies of mass assessment are a scientifically neutral and reliable means of measuring learning or judging capacities for future learning (Filer 2014).

## 2.5 Conceptual Framework – Policy Interpretation and Enactment

This DPhil study will explore ‘the ways in which policy options are explored’ (Forester 2012:486). Maguire et al. (2015) explain that ‘policy enactment is a process of social, cultural and emotional construction and interpretation’ (Maguire et al. 2015:485). Thus, it is essential to examine this process of interpretation to understand both how a policy has been enacted and why it was enacted that way. Often what is being enacted is not the actual policy, but throughout the implementation process there are likely to be interpretations of interpretations (Rizvi and Kemmis 1987). These interpretations of interpretations are undertaken at all levels, from senior management meetings to classroom interaction. It is important to keep in mind that space to interpret also varies from policy to policy and sometimes, from person to person. Understanding this illustrates the complicated relationship between making policy and practising policy in complex situated contexts like schools (Colebatch 2006). Timperley and Parr (2009) argue that making sense of the ‘policy change messages’ is critical to implementation (Timperley and Parr 2009:137). They use Spillane’s (2004) sense-making framework which regards the influences upon policy enactment as a series of cues, with government policy being just one of those cues. In this framework, the interpreter, whether an individual actor or organisational actor, must firstly notice a policy, then engage with it to understand it, followed by developing a plan of how it could be applied to their context, ‘... a policy’s messages about changing behaviour are not inert ideas that reside in the legislation or policy regulations and that are transmitted unaltered into minds’ (Spillane 2004:76).

Timperley and Parr (2009) emphasise the complexity of enacting policy within a school as classroom activities are comprised of policy cues, the need to deliver curriculum content and effective pedagogy. Maguire et al. (2015) developed an approach while examining secondary schools in England in-which they argue that policy in schools is not merely implemented but is rather enacted through a series of interpretations and actions by multiple actors over a period of time. As Forester (2012) explains, enactments emanate from the ‘micro-politics of policy practices through the diverse accounts of situated and entangled practitioners themselves’ (Forester 2012:23). Policymaking occurs at all layers of the system as the policy intent is re-translated by actors at each layer. Interpretation is often a ‘messy process in which policy actors negotiate and create social capital to enable them to enact various policies’ (Maguire et al. 2015:485). Ozga (2000) explains that policy making involves ‘negotiation, contestation or struggle between different groups who may lie outside the

formal machinery of official policymaking' (Ozga 2000:113). Maguire et al. (2015) underline this point rather poetically by pointing out that often the 'other moments' in the process of policy are equally important to the formal process of policy implementation (Maguire et al. 2015:485). In some cases, non-teachers such as Trust leaders, business managers or governors etc may interpret and enact policies more often than teachers.

The value given to a policy may also depend upon the will of a dominant actor within the setting. Maguire et al. (2015) found that some teachers preferred not to interpret policy, such as junior and newly qualified teachers who exhibit what they label as policy dependency in which these teachers display high levels of compliance with what senior members of staff instruct them to do. Time, space, positionality, commitments, experience, loyalties, allegiances, disciplinary commitments and in-school relations, often influence how policies are interpreted and enacted and who interprets and enacts them (Maguire et al. 2015). Human beings often seek to reduce uncertainty and create conditions of equilibrium (Heifetz and Linsky 2002). A teacher may insist that the assessment process is best represented according to their experience (Langer 2002). Even if change is adopted, it may not be genuinely believed in by the teacher or may be enacted in a manner that matches the beliefs of teachers rather than according to the policy intent. When personal theories do tally with policy intent, the likelihood of implementation is enhanced, but there may be a superficiality of understanding which could result in interpretations that are substantially different from those intended (Timperley and Parr 2009). Ball et al. (2012) illustrated how teachers often negotiate tensions with developmental policies that are closer to their professional identity, and imperative policies which they perceive hinder their daily practice. Spillane (2004) terms the conflict of priorities that often determines policy enactment as 'internal representations' versus 'external representations' (Spillane 2004:181). These tensions can 'generate a dissonant professional identity, where a coherent sense of direction is often lost, and a superficial implementation replaces what should be a deep change in pedagogy' (Hopfenbeck et al. 2015:46). Studies in the implementation of Assessment for Learning (AfL) have found that teachers need time to develop assessment skills and that the time demands for deep learning can be at odds with administrative requirements (Hopfenbeck et al. 2015). Maguire et al. (2015) found that the visibility of high stakes policies being implemented resulted in other policies being given little or no attention. Research around progress measurement often focuses upon evaluating the success of the metric and of the policy implementation rather than considering a broader perspective even when this emerges as a crucial aspect of the enquiry (Florez-Petour 2015). Through exploring the interpretation and enactment process, this DPhil study seeks to consider this broader perspective.

After data was generated, it became necessary to revisit the academic literature and add to the conceptual framework. Although much has been written about sensemaking and also that it is situated, less has been written on the specifics of the influences within those situations. Even where scholars have explored the influences on sensemaking none seem to have considered whether these influences can be categorised into frames. Zuckerman (2019) discusses how collective action frames in social movement literature emerge from the process of negotiating shared meanings during group sensemaking. Snow and Benford (1992) describe a frame as an interpretative schema that encompasses various elements from one's current or past environment, such as objects, situations, events, experiences, and sequences of actions, to clarify one's ability to act. They mention that frames originate from master frames, which are broader societal narratives, like the injustice of workers not receiving a share of a company's profits. Gahan and Pekarek (2013) state that frames, as cognitive schemas, not only emphasize aspects of a social situation but also express them as grievances and necessary solutions. Snow and Benford (1988) outline three interconnected purposes of framing: diagnostic, prognostic, and motivational framing, which involve identifying a problem, explaining its cause, and proposing a solution. These processes are often contested, especially in strategic action fields with multiple counter-frames, leading Snow et al. (1986) to assert that frames must align.

The concept of internal institutionalized active inquiry involves making a teacher's work visible for documentation, scrutiny, evaluation, and enhancement, serving as internal quality assurance. This approach fosters questions and internal disagreements, crucial for quality improvement. Schools in the study used external tests combined with teacher experience and evidence from pupils' workbooks to track progress. The aim behind AwL is to restore autonomy to schools and increase teachers' sense of being trusted professionals, although this involves triangulating teacher judgments with evidence from tests and senior teachers' perspectives. This reflects debates around the audit culture, where trust in professionals is balanced with control systems. The schools adopted a social moderation approach, which ensures consistent teacher judgment through collaborative efforts.

Following the implementation of AwL, professional trust now involves cross-referencing a teacher's judgment with test results, workbooks, and insights from senior teachers before recording progress data in the AwL system. This situation echoes the discussions from the 1990s and 2000s about the public sector's adoption of New Public Management, which emphasizes management and measurement (Hood, 1991). Apple (2001) suggested that these methods require constant evidence production to show that tasks are done efficiently and correctly, a concept he called "audit culture", which seems to be the approach used in

AwL moderation. This raises the question of whether trust in teachers has been entirely replaced by control systems (Becker, 1996); can a teacher be autonomous without true autonomy? Groundwater-Smith and Sachs (2002) argued that an audit culture compels professionals to set aside their personal judgment and self-regulate by adhering to external directives. Bailey (1984) highlighted the importance of professionals using their discretionary powers as independent practitioners, rather than just following orders from headteachers, inspectors, or governments. This process suggests a lack of trust in professionals' ability to perform their duties properly. The challenge for headteachers is that trust always involves some risk, and educational attainment is high stakes for all school stakeholders (Tolo et al., 2019).

The AwL policy serves as a case study to illustrate the processes of (re-)interpretation, (re-)translation, and (re-)enactment in a primary school. This perspective helped an exploration of how headteachers created policy spaces and established mobility for the AwL policy by engaging with school stakeholders. Novak (2009) also introduces the concept of invitational leadership, which involves leaders fostering a welcoming and collaborative atmosphere that encourages active participation from all stakeholders (Novak 2019). These principles are applied within educational institutions to cultivate inviting classrooms and school cultures, where students, teachers, and parents feel valued and engaged (Novak 2019).

Novak's contributions underscore the significance of fostering a positive and inclusive educational setting, which can lead to enhanced outcomes for all participants (Novak 2019). Ball (2012) identifies several key policy roles that are crucial in this process. Policy entrepreneurs are those who actively promote and advocate for specific policies within the school setting. These individuals are often innovative and driven by a strong commitment to particular educational reforms. Policy implementers, on the other hand, are responsible for putting policies into practice. They interpret and adapt policy directives to fit the specific context of their school, often dealing with practical challenges and constraints. Policy critics play a vital role by critically assessing and questioning the policies being implemented. They highlight potential issues, unintended consequences, and areas for improvement, providing a necessary counterbalance to the enthusiasm of policy entrepreneurs. Policy mediators facilitate communication and negotiation between different stakeholders involved in the policy process. They help to bridge gaps and resolve conflicts, ensuring that policies are enacted smoothly. Ball's analysis emphasises the dynamic and contested nature of policy enactment, where these roles interact and sometimes conflict, reflecting the complexities of translating abstract policy formulations into practical strategies within specific educational contexts. This nuanced understanding of policy work underscores the importance of

examining how various actors within schools contribute to the shaping and implementation of educational policies (Ball 2012).

## 2.6 Research on Assessment Reform Informing Phase 2 Data Collection

An exemplar of using the Maguire et al. (2015) enactment framework to explore responses to new assessment policy in primary schools is documented in Thrupp et al. (2018). The Research, Analysis and Insight into National Standards (RAINS) Project was undertaken in five primary schools (Kanuka, Magenta, Cicada, Seagull, Juniper), and one middle school (Huaia) in New Zealand between 2011–2013 by the authors to explore the impact of the new assessment arrangements called Kiwi Standards on the culture of primary schools. It sought to address the following research questions:

- 1) How are Boards of Trustees (BOT), senior leadership teams (SLT) and teachers in different school contexts enacting the Kiwi Standards policy?
- 2) To what extent is performativity apparent in these enactments of policy?
- 3) How does the evidence on policy enactments and performativity in relation to New Zealand's Kiwi Standards compare to the international evidence?
- 4) What lessons are there from the research for policy and for practice in schools?

(Thrupp 2018:88)

Kiwi Standards had been created with a large degree of latitude for interpretation and so the RAINS study involved case study research to examine the ways in which the BOT, the SLT and individual teachers in each school were enacting policy, as well as understanding the responses of children and parents. Semi-structured interviews and other digitally recorded and unrecorded discussions formed the mainstay of data collection and there was also observation of classrooms and wall displays, and meetings and collection of relevant school documents and student data. Together with providing a snapshot, the researchers attempted to examine changes in enactment over time and provide rich descriptions of how schools were enacting the Kiwi Standards. Case studies of two low, two mid and two high Socio-economic Status (SES) schools, with diversity in terms of rural/urban, primary/intermediate and ethnicity, including at least one school serving a high Māori intake. Schools that were known to have reputational, recruitment or other known problems were excluded from the sample, as were those schools who had declared opposition to the Kiwi standards.

Thrupp et al. (2018) found that schools' enactments of the Kiwi Standards were coloured by both their social or situated contexts (such as intake and community characteristics) and their organisational contexts (such as school type, stability of staffing and leadership, and

recent approaches to curricula, pedagogy, assessment and professional development). The researchers concluded that the responses of the schools to the Kiwi Standards reflected school context and the involvement and support of the SLT. Schools also varied in the extent to which staff discussion around the Kiwi Standards and Kiwi Standards-related areas (such as testing or reporting) was apparent to the researchers. Staff training was also variable and depended upon a school's knowledge and resources. Kiwi standards provided a 4-point scale but the application of this differed across the schools (ibid 185). Some schools felt that they needed to go outside of the standards to judge pupils accurately which suggested limitations of the framework (ibid 186). Teachers most often raised that it was unclear what it meant to be Working At or Working Towards a particular curriculum level because they found them to be broad and open to interpretation, and end of year expectations varied in important ways too. The researchers found that this led to differences in the At judgement, and there was also variation of specificity of standards within each level which seemed to increase variation between both subjects and schools. The researchers also found that schools in the lower socio-economic areas were also (generally) a little less demanding or specific about what they required of children in terms of end-of-Year At expectations. Researchers observed that many teachers had classroom wall displays that showed pupil's relative positioning of themselves and their classmates against curriculum levels and/or the Kiwi Standards. Under Kiwi Standards, teachers are required to grade in Overall Teacher Judgement (OTJ) and the balance between types of evidence for informing OTJs was another point of variation depending on teacher experience. Most of the schools had a record of achievement through portfolios or similar which conveyed evidence of achievement against national norms as well as learning samples that could be used as evidence in relation to Kiwi Standards by individual teachers. Moderation within clusters of local schools was one activity that often showed up such variations in interpretations between s

During the process of data collection for this thesis, the National Foundation for Educational Research (NFER) were commissioned by the Department for Education (DfE) to gather evidence on the types and range of non-statutory assessment approaches in use in primary and secondary schools following the removal of levels. They conducted 118 semi-structured qualitative telephone interviews (42 with senior leaders and 76 with teachers) in 42 primary and secondary schools in England between October and November 2017 (Poet et al. 2018). The data revealed that teachers and senior leaders were focusing more on formative assessment since the introduction of Assessment without Levels (AwL). This was reported to be enabling teachers to differentiate activities, refine their planning and provide support more effectively. Poet et al. (2018) observed four main components to schools' non-statutory assessment approaches – formative assessment, summative assessment, moderation and

tracking and reporting systems. Primary schools tended to use external tests and moderation more than secondary schools. Non-statutory assessment tended to focus on the core subjects of maths, English and science. Assessment of non-core subjects was, comparatively speaking, less frequent and rigorous in primary schools, whereas secondary schools tended to use a similar approach across all subjects. Some schools also assessed pupils' wider skills, attitudes and behaviour.

Poet et al. (2018) found that although there was a great deal of overlap, no two schools were using exactly the same descriptive categories to assess pupils' performance. Some were using the same number of categories, but with different names. Schools' non-statutory assessment approaches were more curriculum-led and effective than before AwL. All participants reported that their school had spent a great deal of time identifying and implementing their new approach. Most interviewees reported that teachers were now spending a similar amount of time on assessment as before AwL with a few reporting an increase in workload due to the requirements of their school's new approach. Schools were usually involved in collaboration with other schools, initially to discuss the merits of different assessment approaches and subsequently to provide ongoing information and moderation (especially among primary schools). Generally, interviewees said the best way to share good practice on non-statutory assessment was through collaboration with other schools and/or online resources. Interviewees reported both positive and negative impacts of AwL on teachers' confidence in making summative judgements of pupils' attainment. Some were confident and felt that their schools' approaches were working well, especially those who had taken part in training and moderation. Others said that the lack of national standards for non-statutory assessment had undermined teachers' confidence.

Some interviewees reported to Poet et al. (2018) that pupils now had a deeper understanding of topics, rather than focusing on progressing as quickly as possible through a series of levels. However, only a minority of interviewees felt that their school's assessment approach worked well for pupils with special educational needs (SEN). Others said that their school's approach did not adequately recognise the small steps of progress made by pupils with SEN. Schools used a variety of methods to communicate assessment information to pupils with some interviewees reporting a positive impact on pupils' understanding of what they needed to do in order to make progress, particularly where the pupil had set their own learning objectives. Generally, interviewees felt that the shift to providing pupils with more detailed feedback was helping to reduce the 'labelling' effect of levels. Some participants reported that some parents welcomed the more descriptive

information about their pupil's progress, whereas others said that parents had understood the previous system of levels and did not currently understand their new approach.

The House of Commons Education Committee published a report on primary school assessment in 2017 which highlighted several critical issues and recommendations regarding the changes introduced in 2016 to National Tests in Year 2 and Year 6. The Committee felt that the Standards and Testing Agency (STA) had poorly implemented the new assessment system, citing delays in guidance and the leaking of test papers online, which caused significant disruption in schools. The STA's failure to provide timely and clear information left schools with insufficient time to prepare for the new tests. The Committee recommended longer lead-in times for future changes and improved communication processes to mitigate the negative impacts of constant change. The design of the new tests, particularly for reading and writing, faced criticism. The focus on technical aspects like grammar and spelling over creativity and composition in writing assessments was questioned. The Committee suggested making spelling, punctuation, and grammar tests non-statutory at Key Stage 2 and urged the government to reconsider the balance between technical skills and overall writing quality. It also highlighted the need for a thorough evaluation of the reliability of teacher assessment judgments. The removal of national curriculum levels without adequate support for schools led to the adoption of ineffective assessment systems. The Committee emphasised the importance of Continuous Professional Development (CPD) for teachers, particularly in assessment training. It called for the government to provide high-quality advice and guidance to schools, including a more developed 'item bank' of assessment tools and case studies.

The Committee concluded that the high-stakes nature of the assessment system negatively impacted teaching and learning, leading to a narrowing of the curriculum and 'teaching to the test.' The Committee noted that the pressure to achieve results in English and maths often came at the expense of other subjects like science and the arts. It recommended that Ofsted ensure a broad and balanced curriculum is reported in every primary school inspection and that performance tables include a rolling three-year average of Key Stage 2 results to reduce the focus on individual year's results. The Committee acknowledged the inherent link between statutory assessment and school accountability. It suggested that the high-stakes system did not necessarily improve teaching and learning at primary school. The introduction of a baseline measure in reception to better track pupil progress was discussed, with the Committee urging caution to avoid shifting negative consequences to early years. It recommended a thorough evaluation of the potential harmful impacts of such a measure, involving early years experts and practitioners. To combat the negative impacts of high-

stakes assessment, the Committee proposed several measures. These included publishing a rolling three-year average of Key Stage 2 results, ensuring Ofsted reports on a broad and balanced curriculum, and supporting a culture of wellbeing among staff and pupils. The Committee also called for a review of how Ofsted inspectors use Key Stage 2 data and suggested that expected standards should be raised over a longer period to give schools time to adjust. In summary, the Education Committee's report highlighted significant issues with the implementation, design, and impact of the new primary assessment system. It called for better support and training for teachers, a more balanced approach to assessment, and measures to reduce the high-stakes nature of the current system.

## Chapter 3: Research Design and Methods

### 3.1 Introduction

As detailed in the introduction, how primary schools interpret new policies and enact them is not well documented in the research literature. Through 14 in-depth interviews with headteachers and class teachers and 10 group discussions with pupils, parents and school governors, this thesis explores how the policy of Assessment without Levels (AwL) was enacted, why it was enacted that way, and through this an insight into the wider topic of policy enactment in primary schools is gained. This thesis aims to increase understanding of how primary schools enact new policies by exploring how six primary schools enacted AwL between 2014 and 2019. Meyer (2001) contests that case studies are particularly useful for responding to how and why questions about a contemporary set of events. In this thesis, the case is the enactment of the policy of Assessment without Levels (AwL) in primary schools in England, and this is treated as a critical case that provides insight into the wider topic of how primary schools enact a new policy (Flyvberg (2006). The units of analysis within the case are the participant schools (Teddlie and Yu 2007), and each of these have stakeholders such as headteachers, class-teachers and pupils which are the sub-units of analysis. The three research questions explored in this study are:

- 1) How did primary schools enact AwL?
- 2) Why did they enact AwL that way?
- 3) What does the data tell us about primary school policy enactment more generally?

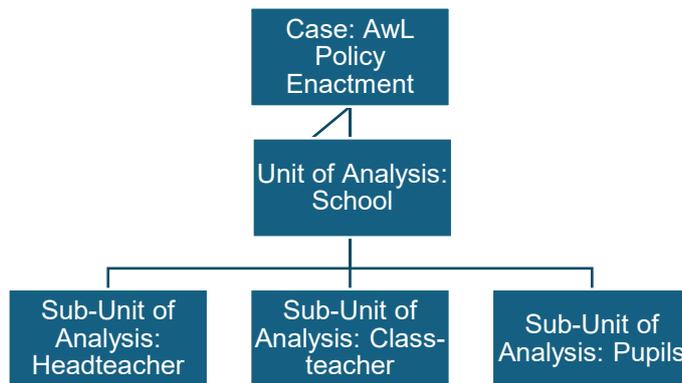
This was a part-time DPhil study conducted over 8 years with several breaks in-between and data was collected in two phases. The research design and literature review were formed in 2016 and 2017 and after the Transfer of Status meeting in October 2017, Phase 1 data was produced through interviews with headteachers in six primary schools conducted between November 2017 and January 2018. Phase 2 data was generated through interviews with teachers, pupils, parents and governors in two primary schools in March 2019 and analysis and theorising occurred from 2022 after the Covid-19 pandemic.

### 3.2 Research Design

Gummesson (1988:76) argues that an advantage of case study research is the opportunity for a holistic view of the process 'within its total environment' or as Yanow (2014) argues, a policy cannot be fully understood through objective examination but through 'the situated context of the "knower" producing it' Yanow (2014:10). Pal (2005) explains that case studies

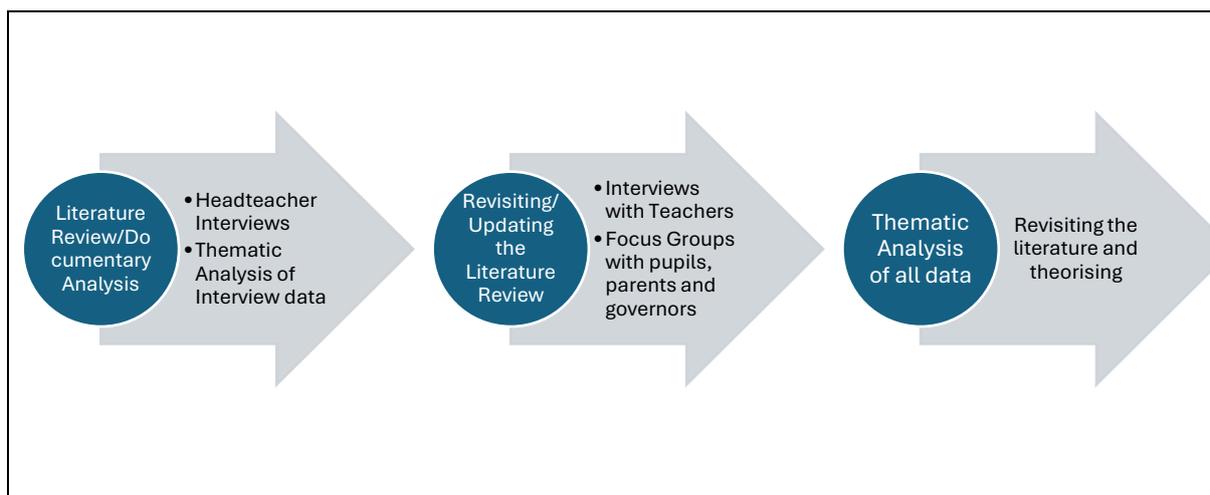
can be significant in two ways. First, cases can be exemplars of a more general phenomena, and secondly is what Yin (1989) calls, analytical generalization, where the researcher uses data from case studies to inform theory. This thesis seeks to utilise the case of the enactment of AwL as a vehicle for providing exemplars of how an assessment policy is enacted in a primary school in England and also exemplars of why choices were made in each school context that shapes enactment in a certain way. Using this data, a reflection upon the concepts and models provided in the literature on policy sensemaking and enactment will be provided. Yin (1989) argues that the case study method has a specific advantage when ‘a “how” or “why” question is being asked about a contemporary set of events, over which the investigator has little or no control’ (Yin 1989:20). According to Yin (1989), such how and why questions ‘deal with operational links needing to be traced over time, rather than mere frequencies or incidence’ (Yin 1989:18). Pal (2005) argues that this is because situations being studied in cases will have a level of complexity and a density of explanation that is not usually characteristic of many other methodological approaches. Yin (1989) defines a case study approach ‘investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used’ (Yin 1989:23). Feagin et al. (1991) emphasise this, “Since the case study seeks to capture people as they experience their natural, everyday circumstances... [It] can permit the researcher to examine not only the complex of life in which people are implicated but also the impact on beliefs and decisions of the complex web of social interaction” (Feagin et al. 1991:8–9). In this thesis both enactment and context are considered to be always evolving and changing, and this thesis provides snapshots of enactment in two different years within the continuous process of enactment (Maguire et al. 2015). Multiple Perspective Interviews (MPIs), that is, separate interviews conducted with participants in the same situational context (Vogl et al. 2019), produce rich detail and provide triangulation.

Figure 6: The Case and Units of Analysis



This is a part-time DPhil study starting with documentary analysis of policy documents from which the research questions inductively emerged which then guided a literature review on the basis of which Phase 1 interview questions were created and these interviews were conducted between October 2017 and January 2018. In October 2018, Phase 1 interviews were analysed thematically and the literature was reviewed again for any relevant updates and an approach to Phase 2 interviews was developed and all six primary schools were contacted with a request to allow me conduct further data collection with two schools agreeing. Subsequently, all of the data from both phases of data collection was analysed with reference to scholarship to answer the research questions (Figure 7).

Figure 7: The Research Process



### 3.3 Participant Selection

Phase 1 concentrated on Headteachers' narratives as these would provide the most comprehensive perspectives of how a school responds to the new policy of AwL (Drew and Guilleman 2014). The population was framed purposively by selecting the location and school type (primary) on the county council website, and a simple random sampling approach was adopted as to the actual schools studied (Drew and Guilleman 2014). A list of 44 schools was generated and all headteachers were emailed with information about the research project with the six that responded becoming units of analysis (Table 3). The headteachers and schools are identified in this thesis by the AwL system that they used (Table 3).<sup>3</sup>

Table 3: Phase 1 Units of Analysis and Ciphers

<b>AwL System</b>	<b>School Identifier</b>	<b>Headteacher Cipher</b>
School Pupil Tracker	SchoolSPT	Head-SPT
Points	SchoolPoints	Head-Points
Passports	SchoolPassports	Head-Passports
Continuum	SchoolContinuum	Head-Continuum
Pupil Asset Tracker	SchoolPAT	Head-PAT
Qualitative	SchoolQualitative	Head-Qualitative

**SchoolQualitative** Church of England Primary School is a one-form entry primary school situated in a designated area of outstanding natural beauty and boasts excellent facilities, including two playgrounds and a large playing field, and the school kitchen offers fresh, home-cooked lunches, ensuring that students receive nutritious meals daily. With 235 pupils in total, SchoolQualitative has an annual intake of 30 pupils for its Reception year. Approximately 7.1% of the students are eligible for free school meals, indicating a range of economic backgrounds. SchoolQualitative is part of a local partnership of schools, with most pupils progressing to a nearby secondary school. This partnership ensures a smooth transition for students as they advance in their educational journey. The active Parent-Teacher Association (PTA) plays a vital role in the school, organising various social events

---

<sup>3</sup> The school and headteacher cipher links with the AwL system used. For example, the school that used School Pupils Tracker (SPT) is labelled SchoolSPT and the headteacher is Head-SPT.

and fundraising activities that significantly benefit the children. SchoolQualitative has consistently maintained a "Good" rating since 2015. The most recent inspection in June 2024 rated the school "Good" across all categories, including quality of education, behaviour and attitudes, personal development, leadership and management, and early years provision. The report highlighted the school's well-planned and sequenced curriculum, strong subject knowledge among staff, and high expectations for all pupils, including those with special educational needs and disabilities (SEND). Previous inspections in April 2023 and October 2017 also confirmed the school's "Good" rating, noting improvements in the curriculum and teaching methods. SchoolQualitative offers a range of extracurricular activities and educational visits, fostering a well-rounded education. The school's commitment to educating the 'whole child' is evident in its efforts to develop the individual interests and talents of its students. This holistic approach to education ensures that children not only excel academically but also grow into well-rounded individuals.

**SchoolPAT** caters for children aged 2 to 11 in a deprived area of a large town. The school has a capacity of around 175 students but currently enrolls approximately 221. The school has a strong commitment to supporting students with special educational needs, ensuring that every child receives the attention and resources they need to thrive. The socio-economic background of the pupils is diverse. Approximately 22.3% of the students are eligible for free school meals, indicating a range of economic backgrounds. In terms of academic performance, SchoolPAT has consistently demonstrated excellence. The school's curriculum is broad and balanced, with a particular emphasis on STEAM subjects—Science, Technology, Engineering, Arts, and Mathematics. The school's commitment to high standards is reflected in its Ofsted inspection results. In September 2018, SchoolPAT was rated "Outstanding" in all areas. The inspection report highlighted the school's strong leadership, high-quality teaching, and excellent student outcomes. The inclusive environment and effective support for students with special educational needs were particularly praised. Prior to this, the school had been rated "Good," showing a trajectory of continuous improvement. Beyond academics, SchoolPAT offers a rich array of extracurricular activities. Students can participate in various sports, music programs, and clubs, contributing to a well-rounded education. The school's dedication to physical education has been recognized with the School Games Gold Mark Award for the 2023/24 academic year. Community engagement is another key aspect of the school's identity. The active Parent-Teacher Association (PTA) plays a significant role in organizing events and fundraising activities, fostering a strong sense of community and collaboration.

**SchoolPoints** is a large community primary school in a medium-sized city. The school's mission is to provide a creative and stimulating curriculum, ensuring that every child has the opportunity to excel both academically and personally. There are 631 pupils and approximately 7.1% of the students are eligible for free school meals. The school has been rated by Ofsted as Outstanding in 2015 in its overall effectiveness. Pupils praised the quality of education provided by the school due to a well-designed and effectively implemented curriculum that leads to high levels of pupil achievement. Inspectors observed that pupils exhibited excellent behaviour and positive attitudes towards learning, contributing to a conducive learning environment. They said that the school also excels in supporting pupils' personal development, offering exceptional programmes and initiatives that foster growth and well-being. They concluded that the leadership and management team at the school is highly effective, consistently driving school improvement and maintaining high standards across all areas. Community engagement is a key aspect of SchoolPoints' identity. The school actively involves parents and the local community through various events and initiatives. The strong Parent-Teacher Association (PTA) plays a significant role in supporting school activities and fundraising efforts, fostering a sense of collaboration and community spirit. In addition to its academic offerings, SchoolPoints provides a wide range of extracurricular activities. Students have the opportunity to participate in sports, music, and various clubs, allowing them to develop their skills and interests beyond the classroom. These activities are designed to complement the academic curriculum and promote a well-rounded education. In January 2025, Ofsted graded the school as being Good in Quality of education, and Outstanding in Behaviour and attitudes, Personal development, Leadership and management, and Early years provision.

**SchoolPassports** is a one-form community primary school in an urban area of a medium-sized city. The school places a high emphasis on values such as friendship, trust, hope, justice, and forgiveness. These values are deeply integrated into the daily life of the school, fostering a positive and respectful atmosphere. SchoolPassports has a capacity of 210 students but currently enrolls approximately 240. For the 2024/25 academic year, the school received 137 total preferences for places, with 60 first preferences offered. Approximately 11.7% of the students are eligible for free school meals and the school is committed to supporting all students, regardless of their socio-economic status. The school has consistently demonstrated strong performance in Ofsted inspections. In February 2023, it was rated "Good" overall, with inspectors highlighting strong leadership, effective teaching, and the well-being of students. Previous inspections in March 2017 and November 2012 also rated the school as "Good," showcasing its consistent performance over the years. The curriculum at SchoolPassports is broad and balanced, designed to engage and challenge

students. The school aims to inspire a love of learning and help students achieve their full potential. In addition to its academic offerings, the school provides a variety of extracurricular activities, including sports, music, and clubs, enriching students' educational experience and promoting well-rounded development. Community engagement is a key aspect of SchoolPassports' identity. The school actively involves parents and the local community through various events and initiatives. The active Parent-Teacher Association (PTA) supports school activities and fundraising efforts, fostering a strong sense of community and collaboration.

**SchoolContinuum** is a large Church of England primary school in a medium-sized city. The school is built on core values such as moral courage, tolerance, respect, and compassion. It has a capacity of 480 students and currently enrolls approximately 475. For the 2024/25 academic year, the school received 137 total preferences for places, with 60 first preferences offered. Approximately 25% of the children have Special Educational Needs (SEN), 45% speak English as an Additional Language (EAL), and on average, 30% are disadvantaged and therefore eligible for the Pupil Premium Grant (PPG). The school has consistently demonstrated strong performance in Ofsted inspections. In September 2023, it was rated "Good" overall, with inspectors highlighting strong leadership, effective teaching, and the well-being of students. Previous inspections in March 2018 and March 2013 also rated the school as "Good," showcasing its consistent performance over the years. In addition to its academic offerings, the school provides a variety of extracurricular activities, including sports, music, and clubs, enriching students' educational experience and promoting well-rounded development. Community engagement is a key aspect of SchoolContinuum's identity. The school actively involves parents and the local community through various events and initiatives. The active Parent-Teacher Association (PTA) supports school activities and fundraising efforts, fostering a strong sense of community and collaboration.

**SchoolSPT** is a one-form primary school serving children aged 4 to 11 on the edge of a medium-sized city and is part of a Multi-Academy Trust formed by local schools. The school emphasizes core values such as respect, compassion, and responsibility. SchoolSPT has a capacity of approximately 140 students and currently has a full enrolment. For the 2024/25 academic year, the school received 45 total preferences for places, with 17 first preferences offered. The socio-economic background of the pupils is generally above average. Most children start school with overall attainment that is above typical for their age. Nearly 6% percent of pupils are from ethnic minority heritages, mainly from mixed White and Black African families. There are no pupils who speak English as an additional language, and no

pupils are entitled to free school meals. About five percent of students have special educational needs (SEN) for learning, behavioural, or physical difficulties, which is well below the national average. The school has consistently demonstrated strong performance in Ofsted inspections. In March 2018, it was rated "Good" overall, with inspectors highlighting strong leadership, effective teaching, and the well-being of students. Previous inspections have also maintained a "Good" rating, reflecting the school's commitment to continuous improvement. The curriculum at SchoolSPT is broad and balanced, designed to engage and challenge students. The school aims to inspire a love of learning and help students achieve their full potential. In addition to its academic offerings, the school provides a variety of extracurricular activities, including sports, music, and clubs, enriching students' educational experience and promoting well-rounded development. Community engagement is a key aspect of SchoolSPT's identity. The school actively involves parents and the local community through various events and initiatives. The active Parent-Teacher Association (PTA) supports school activities and fundraising efforts, fostering a strong sense of community and collaboration.

One year after Phase 1 of data generation, all six headteachers were approached to participate in Phase 2 with Head-Points, Head-Qualitative and Head-PAT agreeing, although at a late stage, Head-Points withdrew due to other commitments. Phase 2 sought to learn how multiple actors within a school respond to and enact a new policy (Maguire et al. 2015). Single case studies can be criticised for not providing multiple perspectives that triangulate each-other (Meyer 2001, Davies et al. 2001) and so attention was paid to ensuring that there were sufficient units and sub-units of analysis together with a mix of in-depth interviews and group discussions to provide depth and also triangulation (Davies et al. 2001). Phase 2 selection was driven by role homogeneity with the visual method from Phase 1 guiding the selection of categories (roles) of participants with a purposive approach being used to define inclusion criteria for teachers, pupils and parents from Year 4 and Year 6 together with school governors. Heads Qual and PAT agreed to be interviewed again, and as gatekeepers they facilitated recruitment of other participants. One teacher from Year 4 and one from Year 6 were selected in each school. These years were chosen because Year 6 teachers have experience of using school progress data provided by Years 3,4 and 5 while preparing pupils for National Curriculum Tests. Phase 1 data revealed that inter-school moderation in Year 6 is conducted using National Curriculum Tests frameworks rather than AwL systems and so it was hoped that that could be explored. To provide a contrasting perspective, it was decided that Year 4 teachers and pupils would be interviewed as these pupils were old enough to be able to make meaningful contributions and the teachers would have experience of their school's AwL systems without concerns of performance in external tests. In some schools

like SchoolPAT, Year 5 and 6 pupils are taught together and so it would be the Year 6 teacher teaching them, and therefore, this Year group would be explored already. Perspectives from teachers in different Year groups and of differing length of teaching experience adds greater depth, detail and on occasions contrasts to headteacher narratives. Some spontaneous sampling occurred during emplacement within the school as informal conversations with a Year 5 teacher in SchoolQualitative and a Year 3 teacher in SchoolPAT suggested that it would be useful to conduct a formal recorded interview with them (Drew et al. 2014).

The pupil selection strategy was based on experience homogeneity, that is, pupils in the same class would have a similar experience of assessments and so know how things are done by their teachers. Pupils were considered as experts in the receipt of feedback for progress within their class and workbooks. Gibson (2012) suggested that pupils would be more comfortable speaking and would stimulate each-other's experiences and so in each school, one discussion group with Year 4 pupils and one discussion group with Year 6 pupils was planned as pupils of this age group are old enough to be able to reflect and express their thoughts although they vary greatly in their maturity (Gibson 2012). The class-teacher was asked to select two pupils working Below, two working At, and two working Above expectations so that the perspectives of those with differing levels of attainment are explored. Limiting each group discussion to six participants should allow all participants time to contribute and provide sufficient data as some pupils are more confident and articulate and other may be shy. Class-teachers are best-placed to be the gatekeepers and safeguard pupils and know the attainment of the pupils and their mental readiness to participate. A drawback could be that the class-teacher's bias and presumptions about pupils may influence selection, and their explanation of the research could result in pupils having a different understanding of the study to the intention of the researcher.

Experience homogeneity was also the main consideration for parent selection as the reporting processes to parents would be a school-wide policy with some variations for to meet the needs of some pupils. In particular, there was a desire to understand whether parents had clarity about the strengths and weakness in their child's knowledge and skills and next steps of learning so that they can help them at home which was a stated aim of AwL. The school offices were asked to send emails and notices to all parents of pupils in Years 4 and 6 and give those interested a participant information letter and consent form. Parents will trust invitations from the school but are very unlikely to participate if a stranger contacts them. A drawback is that only parents that engage with the school and/or read school letters and newsletters will be aware of the research and so the sample would not be

representative but rather it would be indicative. Also, there would be a self-selection bias although it was hoped that parents with children in the same class know each other and motivate each other to participate.

Role homogeneity was the main aim behind school governor selection although school governing bodies comprise parents, diocese or academy representatives, teachers and other expert volunteers and each perform different duties to hold the school accountable. They are also enhanced by the reputation of their school and so governors may be cautious to speak honestly or to comment on topics that fall under the remit of school leaders or other governors. Headteachers are the gatekeepers, and it was hoped that all governors would be invited to participate but Head-Qualitative invited two governors whom they were comfortable me speaking with and Head-PAT did not permit the group discussion.<sup>4</sup> Therefore, the one group discussion with the two governors in was not triangulated but the data is useful for providing further context to the understanding of AwL policy enactment in SchoolQualitative and when combined with Phase 1 data from the six headteacher interviews does provide some interesting insights on the dynamics between school leaders and the Governing Body (GB) which could be explore in future research. Ciphers for all participants were created by combining the participant role with the school (Tables 4 & 5).

---

<sup>4</sup> School governor group discussions were planned partly because Head-Points had invited me to meet the whole governing body during their regular meetings but then they had to withdraw from the study.

Table 4: Interviewee Roles and Ciphers in SchoolQualitative

<b>Role</b>	<b>Cipher</b>
Headteacher	Head- Qualitative
Year 6 Teacher	SQ-Y6T
Year 5 Teacher	SQ-Y5T
Year 4 Teacher	SQ-Y4T
School Governor A	SQ-SGA
School Governor B	SQ-SGB
Year 4 Pupil Group discussions	Data aggregated
Year 6 Pupil Group discussions	Data aggregated
Parent A	SQ-PA
Parent B	SQ-PB

Table 5: Interviewee Roles, and Ciphers in SchoolPAT

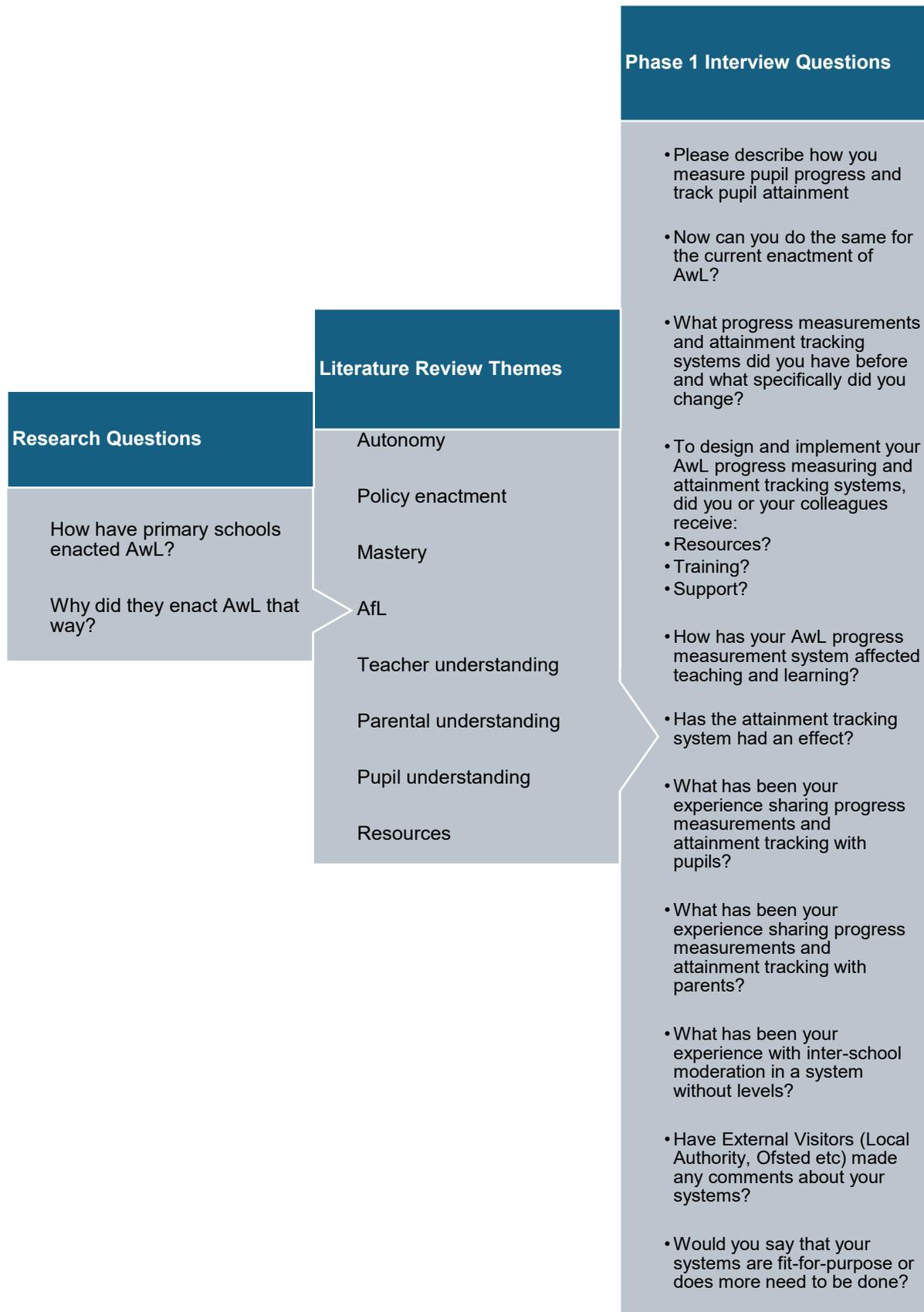
<b>Role</b>	<b>Cipher</b>
Headteacher	HPAT
Year 5/6 Teacher	SPAT-Y6T
Year 4/5 Teacher	SPAT-Y4T
Year 3 Teacher	SPAT-Y3T
Year 4 Pupil Focus Group	Data aggregated
Year 6 Pupil Focus Group	Data aggregated
Parent Focus Group	Data aggregated

### 3.4 Approach to Phase 1 Interviewing

A thematic analysis of the literature review and documentary analysis, together with the first two research questions that explored both “how” AwL was enacted and “why” it was enacted

that way, guided the creation of the interview questions (Figure 8) which were piloted twice and adjusted according to participant feedback (Schreiber 2012).

Figure 8: Development Process of Phase 1 Interview Questions<sup>5</sup>



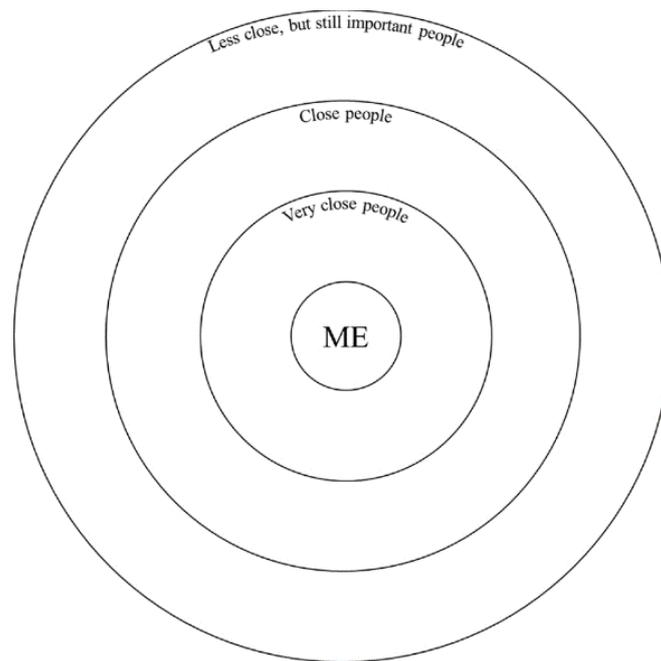
The six headteachers that agreed to participate in Phase 1 of data collection were sent further information, consent forms and interview questions. An interview schedule was developed and sent to participants together with an information letter and consent forms (Appendices 1-7). The headteachers were given the option to either conduct the interview at their school or at an external location with five choosing the former option. Logistics meant that the interview with Head-Qualitative was conducted by telephone and as we had met previously, the absence of face-to-face interaction did not seem to effect quality of data being generated (Bourque and Fielder 1995). At the start of every interview, an explanation of the study and a re-reading of the consent form was provided and included in the audio recording. It was evident in all cases that participants had previously read the interview questions and prepared answers in their mind, and knowing which questions were to come enabled some of them to answer several questions together in a natural conversational style (Brinkmann and Kvale 2018). Head-SPT, Head-Points, Head-Qualitative, and Head-PAT followed the interview schedule, and Head-Passports and Head-Continuum referred to lead the order of what was explored. Head-Continuum brought multiple sheets and books to illustrate how their system worked, and Head-PAT also explained by using an example sheet from Pupil Asset Tracker and invited their deputy to participate in some of the interview at points where they felt that this could be useful to the research. During the interview, participants expressed what they experienced, perceived and did, often with little interpretation or explanation by the researcher (Brinkmann and Kvale 2018). The exploration of a new policy such as AwL through research interviews provided participants an opportunity to reflect upon their enactment and discover new relationships between their own ideas and experiences as they heard themselves speak. Follow-up questions and probes were created through spontaneous analysis of participant responses throughout the interview by the researcher (Brinkmann and Kvale 2018). Often the researcher's interpretation of the participant responses was articulated and co-analysed with the participant (Brinkmann and Kvale 2018). Prior, during and after interviews, the researcher reflected upon Brinkmann and Kvale's (2015) criteria for a quality interview, with the aim of developing the researcher's technique with each interview.

According to Brinkmann and Kvale (2015:181), a quality interview is one where the interviewee provides answers that are 'rich, specific, and relevant'. The interviewer should strive to ask concise questions, which in turn allows the interviewee to give extended and detailed responses. Throughout the interview, it is crucial for the interviewer to actively follow up on and clarify the meanings of the interviewee's answers, ensuring a deep understanding of the relevant aspects. Additionally, the interviewer should 'continuously verify their interpretations' of the interviewee's responses to maintain accuracy and clarity over the

course of the interview (Brinkmann & Kvale, 2015:181). Interviews were transcribed within one day of the interview being conducted and as the recordings of each interview were being heard, the length of researcher's questions and quality of answers were reflected upon and adjustments in the researcher's approach were made for future interviews. Viewing data generation as an incremental process, new insights from interviews were explored with subsequent interviewees (Brinkmann and Kvale 2015), and for one point, Head-SPT was emailed a question that had not been explored with them during the study's first interview.

The main theoretical framework used in this thesis is from Maguire et al. (2015) who observed that policies in schools were not implemented all at once but enacted over time, and that they were interpreted and reinterpreted by several different stakeholders. To explore this, question 2 of the interview schedule used an interactive visual method designed to access headteacher narratives on the stakeholders involved during the initial response to the new policy during 2014 to 2016, and also once the system had been embedded in 2017 to 2018 (at the time of the interviews). The visual method aimed to explore headteacher narratives of their policy and how it was shaped by stakeholders, or as Kim (2016:3) puts it, it was an 'attempt to create a fit between a situation and a story schema about some experience or event that consists of who, what, how, and why'. The added value of a visual method is that it enables the 'research participants themselves (to) access and articulate their inner thoughts and feelings' while simultaneously providing a method for the researcher to see and access those thoughts (Hogan and Pink 2012:7). As participants' perceptions are not accessible through an objective process (Irving 2011), the researcher has to find ways to access memories and experiences and enable the participant to express these in a manner that connects with the purposes of the research (Hogan and Pink 2012). The 'hierarchical mapping technique' developed by Antonucci (1986) is a prominent affective name generator that is widely used in social psychology and social network research (Figure 9), but to the best of my knowledge this is the first time that it has been applied to schools' policy research. The instrument uses a standardised model of three concentric circles depicting three different degrees of closeness. As Glegg (2019) explains, in this model, each circle is related to certain types of social support characterised by closeness, composition, and duration. Nadoh et al. (2004) show that participants use three criteria to place ties into the three circles: type of tie (i.e., role relation), quality of tie (i.e., degree of closeness), and formal properties (e.g., regular contact).

Figure 9: A Standardised Network Map based on Antonucci (1986)



In this thesis, rather than the participant being the ego, the policy of AwL was the central part of the map and headteachers were asked to judge the influence upon interpretation and enactment rather than closeness of relationship. Headteachers stuck post-it notes<sup>6</sup> on concentric circles to represent the level of influence at the early stage of interpretation and sensemaking of AwL, and also at a later stage when their AwL was being used, with a more centred position indicating greater influence. The exercise acted as a memory recall aid (Barlow et al. 2011) for the headteacher and proved a stimulus and guide for the conversation, and so these maps are not only data in-themselves but also generated verbal data for the interview transcript. This visual method creates a new relationship between the data that a traditional interview conveys and the additional symbolic knowledge that a participant created map provides (Hogan and Pink 2012).

### 3.5 Approach to Phase 2 Interviews and Group Discussions

Both Gibson (2012) and Irwin and Johnson (2005) suggested that researchers cannot expect to establish a strong rapport and trust with children upon their first meeting, which prompted the researcher to design a period of being in the classroom as a helper prior to

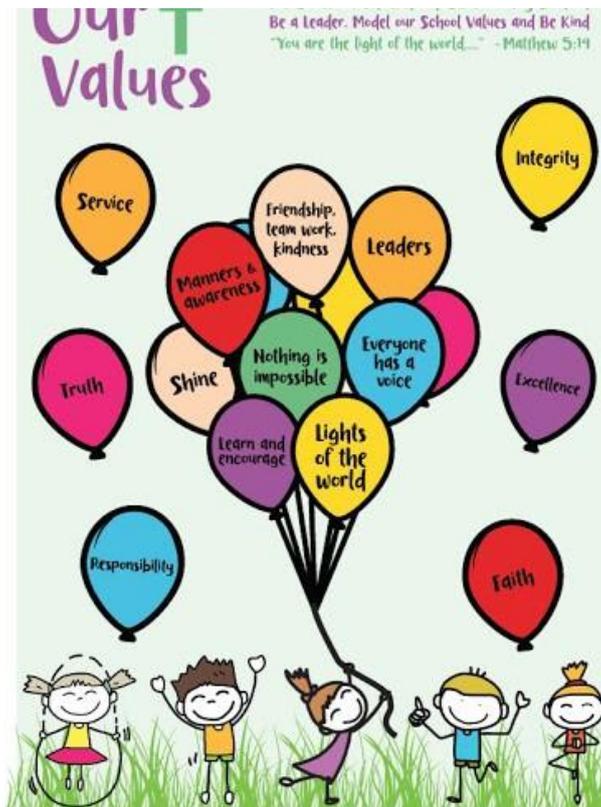
---

<sup>6</sup> The researcher completed this task for Head-Qualitative according to their telephone instructions and then sent them the map for approval.

conducting the interviews. Spending two days within both schools enabled an understanding of the language and assessment norms that teachers and pupils were using in relation to AwL. Hogan and Pink (2012) discuss the idea of becoming emplaced in ways similar to the research participants through participating in similar activities, and this time within the school shaped what was explored in interviews and group discussions together with the language used in the questions. Poet et al. (2018) had reported that teachers perceived that feedback to pupils had improved under AwL and this emplacement within the classroom enabled the researcher to see what that meant. For example, in the SchoolQualitative Year 3 and Year 6 classes, the researcher observed how one pupil's mistaken answer to a teacher's verbal question to the whole class was explored as a learning exemplar on the white board. Although not observed, Head-Qualitative and several teachers informed the researcher that feedback groups were commonly used in which the teacher summons four students to the front of the room and within earshot of the class they discuss the progress and quality of work of each of them rather than written marking of books. The researcher also observed pupils supporting or challenging answers given by their classmates. Research on AfL has shown that verbal feedback from teachers and peers often has more effect on learning than written feedback does, or test results provide (Baird et al. 2017).

The National Curriculum 2016 was designed to encourage mastery of fewer concepts in greater depth and to empower pupils to learn independently of teacher support. The teachers in SchoolQualitative frequently referred to their adoption of a system called Purple Learning that aims to encourage pupils to embrace challenge by taking risks and giving answers even when they are not sure and trying unconventional approaches in their workbooks. In the Year 4 class, the researcher noted the absence of ability streams during mathematics and instead slower learners of a particular topic were invited to sit closer to the teacher so that they can receive additional support while quick learners work more independently undertaking extension activities. There were also a number of after class interventions for slower learners, sometimes in corridors with support teachers, and sometimes at break times or during assembly with the class teacher in the classroom. Head-Qualitative, class-teachers and pupils all referred to the school's values which were displayed on some walls (Figure 10). This observation data was then used to inform the interviews and group discussions as these aspects such as purple learning, values, challenge, feedback groups and ability streams were explored.

Figure 10: Wall Display on SchoolQualitative's Values outside the Year 6 classroom



During the two days spent in SchoolPAT, much of the contextual narrative conveyed to the researcher by the headteacher, teachers and support staff surrounded the social deprivation of the families that the school served which resulted in many pupils displaying a lower level of aspiration. To address this, staff emphasised values such as excellence and determination and embodied the attitude of wanting to achieve in the motto (Figures 11 & 12). Teachers and support staff conveyed pride in the progress that they were observing the pupils were making, and the validation of their approach provided by an 'Outstanding' Ofsted judgement (Figure 13). In Phase 1, HPAT had explained that they had just added this school to their portfolio as it had been judged by inspectors to be failing under the previous leadership. Head-PAT conveyed that staff were trying to encourage pupils to take interest in each other's responses and to challenge or support or ask for explanation. Pupils seemed to generally be less confident than those in SchoolQualitative although they seemed no less happy or enthusiastic about their learning. In contrast to SchoolQualitative, written marking was an essential part of AwL, and so while helping Year 6 pupils with their weekly diary, some of them showed me how they were responding to written feedback from their teacher

and the researcher observed pupils asking for verbal clarification on the comments. The researcher noted the diagnostic questioning that teachers engaged pupils in to discern understanding of concepts, together their usage of visual explanations on the board or PowerPoint to clarify. During Phase 2 interviews, HPAT had described this questioning by teachers and responses by pupils as 'verbal Ping-Pong', and this approach has been found in the research literature to be a crucial element of AfL (Black and Wiliam 1998). The researcher also supervised pupils during music and religious studies classes, and it was evident that the school were trying to expose pupils to a wide curriculum.

Figure 11: A Wall Display in the Assembly Hall of some of SchoolPAT's Values



Figure 12: A Wall Display of SchoolPAT's Motto

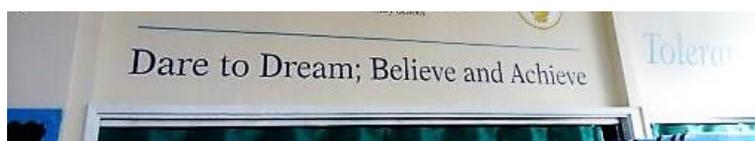


Figure 13: A Banner Advertising SchoolPAT's Ofsted Judgement



Physical space is at a premium in primary schools and so the researcher was flexible as to where a staff told the researcher that they should interview. In SchoolQualitative, only the headteacher had an office and classrooms and halls are often used by support staff even when the teacher is not teaching. This resulted in a variety of data generation sites including corridors and the staff room (Table 6). Flexibility by the researcher resulted in cooperation from the participants. For example, when a teacher or headteacher began saying something relevant during a natural occurrence, the researcher would ask if it could be recorded and it was always permitted; and when SQ-Y6T informed the researcher that they would have no time to for an interview, the suggestion of conducting it while they were on playground duty during lunch was accepted. This flexible approach resulted in SQ-Y6T bringing their lunch into the staff room while the researcher was chatting to Head-Qualitative, and they were happy to commence the interview indoors and then continue it on the playground. The excessive noise of children playing created an ‘eye of the storm’ privacy in-which only the researcher and SQ-Y6T could hear the conversation. Reading the transcript, the quality of the data is not lower than in interviews conducted privately indoors, and as SQ-Y6T shared the most candid reflections upon the gaming (Baird and Elliott 2018) that occurred under levels, perhaps being outdoors amidst the noise allowed them to relax. In both schools, the researcher’s presence during informal moments in the staffroom led to conversations about this research project that resulted in the subsequent conducting of interviews with SQ-Y5T and SPAT-Y3T who were outside of the original sample.

Table 6: SchoolQualitative Data Generation Sites

<b>Data Generation sites</b>	<b>Participants</b>
During a tour of school	Head-Qualitative
Within the classroom during class	SQ-Y4T, SQ-Y6T
On the playground during lunch	SQ-Y6T
Within the staffroom during a break	SQ-Y4T, SQ-Y6T
Seated separately in the corridor	SQ-Y4T, All Pupil Group discussions
Privately in the staff room	SQ-Y5T, Governors Focus group, Parent Focus Group
Privately in the headteacher’s office	Head-Qualitative

The pupil numbers were smaller in SchoolPAT and so some classes included students from two years. Portakabins were being utilised for some teaching, and one of these had an office where the interview with SPAT-Y4T and the Year 4 Pupil Focus group could be conducted privately. Some data generation with SPAT-Y4T and SPAT-Y6T also occurred in the

classroom both during a class and also afterwards. The interviews with Head-PAT were in their office, the first in the presence of their Special Needs Coordinator who Head-PAT invited to contribute to the interview, and the second impromptu one being conducted privately.

Table 7: SchoolPAT Data Generation Sites

<b>Data Generation sites</b>	<b>Participants</b>
Within the classroom during class	SPAT-Y4T, SPAT-Y6T
Within the staffroom	Year 6 Pupil Focus Group
Privately in a classroom	SPAT-Y3T
Privately in the staff room	SPAT-Y6T, Parent Focus Group
Privately in an office	Head-Qualitative, SPAT-Y4T, Year 4 Focus Group

Using the themes inducted from the Phase 1 headteacher interviews (Table 8) and researcher observations from being emplaced with the school, it was decided to explore the research questions with the headteacher and class teachers through both sit-down interviews and during natural occurrences so that the strengths of a well-prepared and knowledgeable researcher and the strengths of a conversational style could be combined (Zhang and Wildemuth 2009). The conversational approach allowed flexibility to record data during natural occurrences within the school day such as the classroom or playground, together with more formal one-to-one interview settings. This resulted in some transcripts including multiple small pieces of data generated over 2-3 days in both schools.

Conversations explored the thesis research questions with broad and open questions, such as, 'can you show me how you record progress?', and 'can you show me how you report progress to parents?' which prompted the interviewee to share their experiences with the researcher listening and reflecting and adjusting the interview direction smoothly (Zhang and Wildemuth 2009). To build a rapport, the researcher presented themselves as 'a learner, a friend, and a member of the interviewee's group, who has sympathetic interest in the interviewee's life and is willing to understand it' (Zhang and Wildemuth 2009:3). This is an approach that Denzin (1989) termed sympathetic identification where the researcher demonstrates empathy without giving advice or passing judgement.

Table 8: Themes from literature and Phase 1 data Informing Phase 2 Interviews

How the AwL system is used by the teacher – mechanics of the system
Teacher clarity of pupil understanding - planning, adjusting teaching
Differentiating pupils – Mastery, ability streams, coasting students etc
Age-related Expectations – teaching, judging and reporting
Ensuring integrity in judgements and data inputs
Moderation experiences
Teacher workload – improved under AwL?
Fitness for purposes?

As D’Souza (2013:413) asserts, ‘parents offer a unique perspective on what works and what doesn’t work in schools’, and so the group discussions with parents explored whether AwL had provided greater clarity about the progress, attainment and next steps of learning for their child (Popat 2015). Technical language was removed, and a flexible questioning approach was adopted (De Decker et al. 2012) with parents tending to respond by sharing stories about their children’s experiences with the AwL system and their experiences with interpreting school feedback about their child’s progress. The schools were in different geographical and socio-economic areas, but both struggled to recruit parents even though they were held near to the end of the school day or when a parent had indicated it would be convenient<sup>7</sup>, but no financial or material incentives were offered (De Decker et al. 2012). In SchoolPAT, the parent focus group was held in the staff room, but we were uninterrupted. Two mothers attended from the start with another arriving 15 minutes later. Another mother had provided some written comments conveyed through the school office. The two mothers that arrived together seemed to be friends and had similar positive opinions about the school and both were very adamant that information that the school provided about their children was clear. The mother that arrived late expressed that teachers did not convey next steps of progress for her son her clearly which unknown to her contradicted the other two parents. After this third parent had heard the complimentary narrative that the other two parents were conveying, she did not contradict them again which suggests that she wanted to conform to the group expectation (Acocella 2012). In SchoolQualitative, the parent focus group consisted of two mothers and so speaking sequentially without interrupting each other was unproblematic. One mother had children with special educational needs and the other had a child that was neither high achieving nor struggling, and so these testimonies offered an

<sup>7</sup> The SchoolQualitative parent group discussion could only be arranged on the morning of day three.

interesting contrast of experiences with AwL. It was also held in the staff room and although one or two teachers did enter to collect items, this did not seem to disturb the parents. Encrypted transcripts were sent to participants for verification.

The school governor focus group in SchoolQualitative was conducted privately in the staff room and comprised of two participants selected by the headteacher, with one being the priest representing the diocese (SQ-GA) and the other a parent who had been serving on the Assessment Committee during the past 18 months (SQ-GB). The researcher adopted a slightly more challenging approach (Brinkmann and Kvale 2018) so that the disagreements that governors had had with Head-Qualitative about the removal of testing to track pupil progress was not avoided. Participants were friendly and informative and spoke in-turn. Both governors shared sensitive data which prompted SQ-SGA to seek reassurance that the data could be approved by the governing board. An encrypted transcript was sent to both governors for approval and SQ-GB confirmed the accuracy on behalf of the board. At the end of the focus group, SQ-SGA suggested to Head-Qualitative that the governor with the most significant concerns about the school's AwL approach be interviewed by the researcher, but Head-Qualitative never arranged this.

A further CUREC process was undertaken prior to Phase 2<sup>8</sup> with special consideration being given to conducting group discussions with primary school-aged children. Gibson (2012) advises increased preparation and a different approach to adult group discussions so that children have a positive experience accessing their thoughts and experiences. As Gibson (2012) puts it, 'The more competent the interviewer is at tailoring the interview to the participants, the less competent the children must be for interviews to be successful (Gibson 2012:153). As individual interviews with an adult can be intimidating for a child, group discussions were chosen (Acocella 2012). Adler et al. (2019) point-out that pupils are experienced in talking to children of the same age that they are familiar with; and as they also have homogeneity of experience of AwL, pupils of the same class were interviewed together. The researcher presented themselves as 'a friendly adult... who wants to learn from children' (Gibson 2012:154). Pupils were viewed as 'competent creators, interpreters, and reporters of their experiences' and therefore, were 'partners in the research process albeit with differing cognitive, linguistic, social, and psychological needs to adults' (Gibson 2012:150). Having been emplaced in the participants' classroom, the researcher used language and content that was part of the assessment experiences of the pupil (such as feedback groups) rather than the language of the policy (Gibson 2012). Pupils seemed to

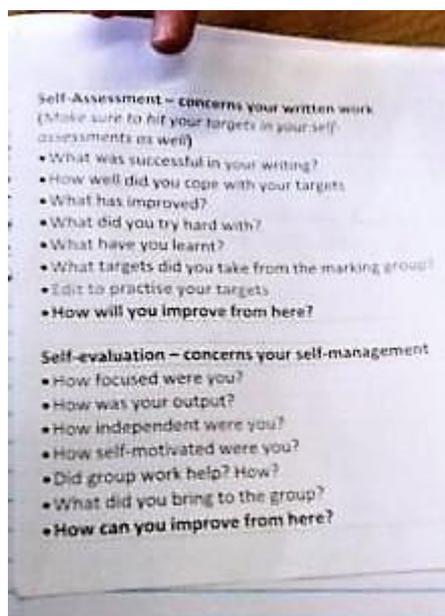
---

<sup>8</sup> CUREC Reference ED-CIA-19-113

often understand questions based on peer answers which they then added to (Adler et al. 2019). Patiently asking follow-up questions helped pupils expand their narrative which they sometimes related non-sequentially, and pupils helped each other to construct the picture and corrected each other in a positive and supportive way. The researcher used non-verbal gestures of interest and encouragement such as eye contact and head nodding. During the interview, conscious attention was given to asking questions and listening to responses in a manner that encouraged pupils to reflect. The researcher kept an awareness of whether pressure to conform to each other or a desire to please the researcher was influencing the responses but in the main they seemed authentic. The researcher tended to ask pupils to *show how* they were assessed in their books which gave them a prop to support their experiences (Figure 14). For example, Year 6 pupils in SchoolQualitative showed the researcher how they take responsibility for their improvement by using editing books in English and also talked the researcher through the self-evaluation checklist for both the quality of written work and also their management of their approach to their work.

In SchoolQualitative, the four Year 4 pupils chosen by SQ-Y4T to participate in the group discussions had differing timetables which meant that they could not meet the researcher together but needed to be interviewed in two pairs (two girls in one, and one boy and one girl in the other) for approximately 15 minutes each in the corridor outside of their classroom. This format of interviewing pairs of pupils from the same class allowed for a thoughtful discussion in-which participants were actively engaged throughout. Only having the researcher and one other pupil's words to focus on seemed to encourage a high level of discussion among participants, and also answers that followed-on from their peer's answer. SQ-Y6T had not selected participants and left it to the researcher to choose whoever they wanted to speak with during the French class conducted by another teacher. As all pupils wanted to participate, and as the researcher wanted to gain the best data from pupils with a range of attainment, together with the desire to practise interview method with children, five group discussions of six to eight pupils were conducted for 10-12 minutes in the corridor outside their classroom. Subsequently, parents were sent a letter through the school office inviting them to opt-out of the data being used and an encrypted transcripts were sent to all parents for their child to read and verify.

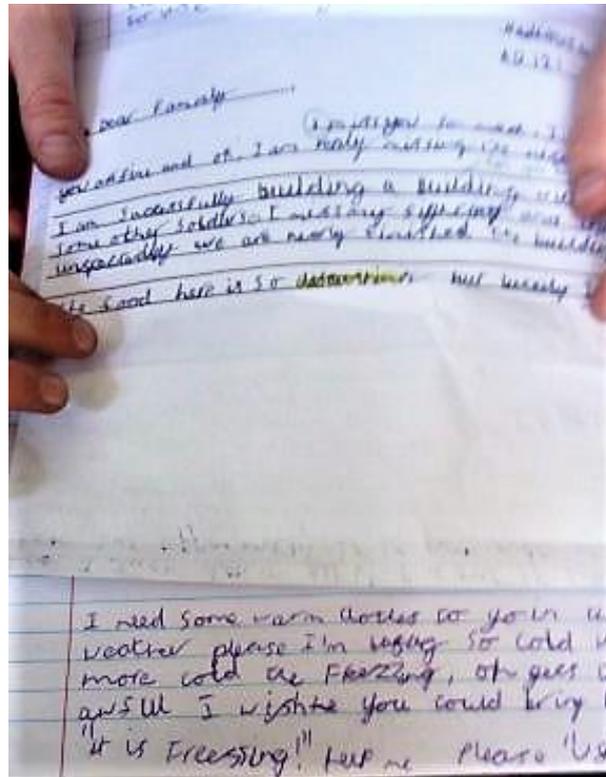
Figure 14: Pupils in SchoolQualitative Year 6 Group Discussions Showing the Researcher how they self-assess and self-evaluate their own work



In SchoolPAT, while SPAT-Y4T was teaching the rest of the class, they invited the pupils that they had selected to be part of the focus group to be interviewed in an office adjacent to the main classroom. Pupils seemed to feel flattered and excited about participating although one Year 4 pupil in SchoolPAT did cry at the start of the day and told SPAT-Y4T that they did not want to participate anymore and received reassurance that they did not have to and should do it only if they felt happy to do so. That pupil did choose to participate and did so actively. The discussions were kept short at 12-15 minutes to avoid saturation and boredom save the SPAT Year 4 discussion which had 6 students together who all had plenty to share. All six pupils participated positively and keenly and in a disciplined manner, raising their hand and waiting to be invited to speak. Much of the data became repetitive after 15 minutes but the respectful thing was to ensure that all participants said what they wanted to share, however, it did result in some fidgeting as pupils had to wait for each other. The SchoolPAT Year 6 group discussion was conducted in the staff room and there were two pupils who were more confident and vocal than others, however, all pupils were very respectful of each other and spoke in turn and listened to each other although discussion between participants was rare and only to remind each other. Pupils were able to articulate and evaluate the way in which they receive feedback clearly and provided some valuable insights. Although several support staff entered and stayed in the room for most of the interview, their presence

did not seem to affect the pupils presumably as they are used to their presence in classrooms.

Figure 15: Pupils in SchoolPAT Showing their Corrections in Response to Feedback



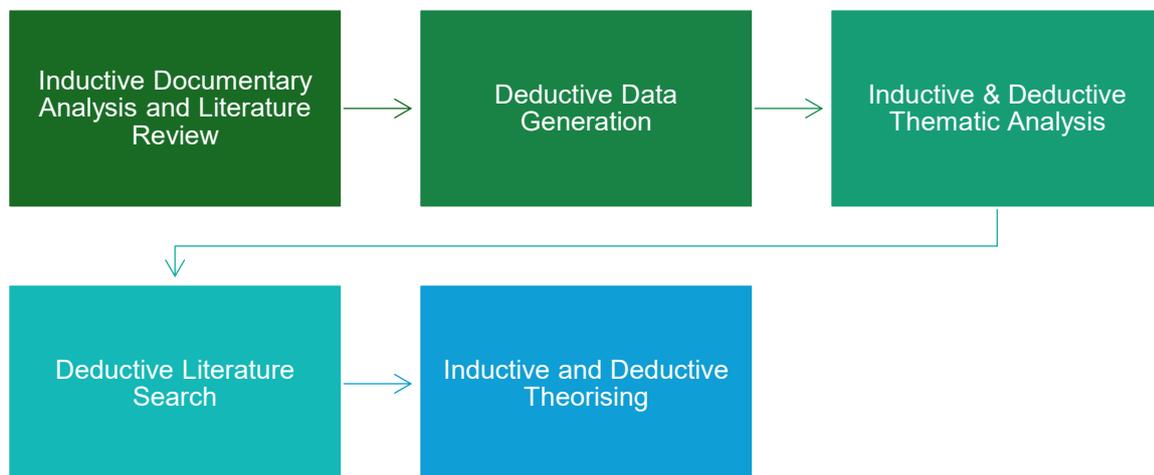
Although not predominantly an ethnographic study, by conducting in situ observations, the researcher has been able to deeply engage with the schools social environment has added some layers of understanding of assessment approaches that would not have been visible through interviews alone or may initially have appeared unrelated (Reeves et al. 2013).

### 3.6 Approaches to Analysis

A continuous reflexive deductive and inductive process has been adopted throughout this DPhil study (Berkovitz 1997) where understandings were gleaned inductively from documents and literature and then used deductively to create questions to generate primary data (Figure 16). The primary data has then been analysed inductively, and these themes were used to deductively search for literature on policy enactment to theorise. When the researcher commenced the exploration of AwL during their Master's degree, there was no academic literature on AwL and so the research involved an inductive analysis of policy documents to glean the rationale behind AwL which was then analysed deductively with

participants (Popat 2015). The interview and discussion data generated was then explored inductively and themes that were identified were related to the academic literature. The inductive analysis of AwL policy documents has been combined with an inductive analysis of academic literature and then the emergent themes have been explored deductively with headteachers in Phase 1. The data from Phase 1 was analysed inductively and emergent themes were explored with participants in Phase 2 deductively. Phase 1 interviews and group discussions were transcribed verbatim with the exception of fillers by the researcher from the commencement of the exploration of the interview questions. Observed or heard reactions which were important for the meaning of the sentence were included such as ‘raised eyebrow’ or ‘chuckles’ (Widodo 2014). All transcripts were sent to the participants for review. On two occasions, clarification was sought from participants on aspects in the transcript that were unclear to the researcher or on a significant aspect that emerged in a later interview. Transcripts were read dozens of times and using the research questions, themes from the literature and theoretical framework, relevant text was highlighted for further analysis and ideas that were repeated were grouped together (Auerbach 2003), and emergent themes such as mastery and feedback were recorded.

Figure 16: The Continuous Reflexive Process of Induction and Deduction in this DPhil Study



Phase 1 themes were then explored in Phase 2 interviews with several stakeholders within a school. For example, enhancing teacher understanding of pupil knowledge and skills was a stated aim of AwL (Popat 2015) and in the Phase 1 interviews, headteachers were asked how their AwL system affected teaching and learning? They answered it in a variety of ways that reflected their priorities. Head-Passports said that the biggest impact that they observed

was that teachers were more engaged in the analysis of the progress tracking system rather than just inputting data for the headteacher to analyse; and Head-Qualitative said that the removal of regular testing meant that teachers were gleaning a pupil's understanding through classroom questioning and marking their work. These comments were coded under teacher meaning-making and informed the unstructured exploration during Phase 2 interviews where class-teachers were asked whether they were clearer about what their pupils had understood and could do (Table 9). After all Phase 2 interviews and group discussions had been transcribed and verified by the participants, they were read dozens more times and then uploaded into NVivo and read again. Relevant text was analytically coded (Miles and Huberman 1994) into nodes that were created based on the deductive and emergent themes from Phase 1. Each transcript was read line-by-line and relevant text was highlighted and copied into those nodes with a lot of text being coded in the same nodes, and extra nodes were also added for emergent themes. Relevant quotes were then extracted with reference to the research questions. A further literature review was conducted, and these texts were integrated with the previous two literature reviews and the researcher's familiarisation with the data to develop themes and create a codebook (Terry and Hayfield 2020). Domain summaries of what participants said in relation to a particular topic or research question were added (Braun and Clarke 2019). Byrne (2022) explains that a reflexive approach to thematic analysis involves a reflection on: (1) the dataset; (2) the theoretical assumptions of the analysis, and (3) the analytical skills/resources of the researcher (Byrne 2022). This approach acknowledges that another researcher may develop different codes and themes using the same data and literature. In Braun and Clarke's (2019) words, reflexive thematic analysis entails 'the researcher's reflective and thoughtful engagement with their data and their reflexive and thoughtful engagement with the analytic process' (Braun and Clarke 2019:594). This iterative approach involves the researcher engaging with the data and literature and asking themselves what the data is conveying and how that connects with what the researcher wanted to know (Srivastava and Hopwood 2009). This process stimulated meaning-making (Trent and Cho 2020) regarding the third research question that sought to consider what the case of the enactment of AwL informs us about policy enactment in primary schools more generally.

Table 9: Examples of Thematic Codes

Code	Example Quote	Related Research Question	Literature Categories for theorising.	Domain Narrative
Curricula	Head-Passports: So, we found ourselves in a position where we have fundamentally more tricky curriculum, knowledge and skills to impart that were fundamentally different; and we had no way of measuring that other than outcomes led objectives. So, that was the starting point for us.	Why are they enacting AwL that way?	School context at the time of policy response	The curriculum has become a fundamental focus and tool of assessment.
Ethos	HC: Our ethos has always been to try and do some things differently and we felt that this was a big opportunity	Why are they enacting AwL that way?	Values. Sensegiving	Some headteachers relished the opportunity to design an AwL system that enacted their values.

Code	Example Quote	Related Research Question	Literature Categories for theorising.	Domain Narrative
Feedback	Head-Points: it's also to do with feedback, "so this is what you need to do". In writing, children have a tick-list so the success criteria will often be linked to the statements that you're trying to assess	How are primary schools enacting AwL?	Vertical Alignment. Clarity of next step of learning.	Assessment is an ongoing process rather than just a summative process.

Code	Example Quote	Related Research Question	Literature Categories for theorising.	Domain Narrative
Resistance	<p>HC: And beyond. I may be pre-empting one of your later questions. What we also didn't want to do was to marry it too tightly into the SATs system of testing which actually makes it very difficult because the only thing that people want to know is how you are going to do on the SATs?</p>	<p>Why are they enacting AwL that way?</p>	<p>Accountability vs Autonomy. Headteacher Values.</p>	<p>There are tensions between the need to hold schools accountable and the desire to give headteachers autonomy.</p>

<b>Code</b>	<b>Example Quote</b>	<b>Related Research Question</b>	<b>Literature Categories for theorising.</b>	<b>Domain Narrative</b>
Teacher Workload	Head- Passports: The work-life balance of my teachers was clear. And I decided not to break down those objectives into tiny-tiny objectives that they would create something like the old APP. I felt that they were onerous	Why are they enacting AwL that way?	Headteacher values. Teacher workload.	Staff wellbeing and retention are big issues across the system and AwL is considered part of the solution.
Tests	Head- Qualitative: What we now do is we have a system where the teachers don't use tests.	How are primary schools enacting AwL?	Headteacher values. Resistance.	Some headteachers emboldened by the communities of practice are willing to experiment with substantial changes if it matches their values.

The transcripts from Phase 2 data generation were read multiple times and checked against the themes from Phase 1 which were found to be pertinent. This prompted a progression to theory triangulation (Patton 1999) which involved a reflection upon the data while revisiting the academic literature and using multiple theoretical perspectives to interpret and analyse the data. This approach enhanced the depth and breadth of research findings by incorporating diverse viewpoints, thereby providing a more comprehensive understanding of the phenomenon under study (Flick 2018). One of the primary advantages of theory triangulation is its ability to enrich the analytical process. The use of theory triangulation is

particularly valuable in complex research contexts such as the enactment of AwL where single-theory approaches may be insufficient to capture the multifaceted nature of the subject matter. For example, Noble and Heale (2019) found that combining theories of organisational change, policy implementation, and educational psychology can offer a more holistic view of how educational reforms impact schools and teachers. They found that a multidimensional analysis helped to identify the interplay between different factors and provides a more nuanced understanding of the research problem.

Theory triangulation also contributes to the validation of research findings because when multiple theories converge on similar conclusions, the credibility of the results is strengthened as it suggests that the findings are robust and not merely artifacts of a particular theoretical lens (Patton 1999). Conversely, if different theories lead to divergent interpretations, this can highlight areas where further investigation is needed, prompting researchers to refine their hypotheses and explore additional dimensions of the phenomenon (Thurmond 2001). A wide range of literature was employed that both challenged and integrated existing theories, and reflection on those and the data lead to the development of new theoretical insights. By juxtaposing different theoretical perspectives, gaps, contradictions, and complementarities were identified which stimulated theoretical innovation, synthesize elements from various theories to construct more comprehensive and integrative models (Denzin 1978). These theories were synthesised and how they intersect with the data and how it contributes to the overall analysis has been articulated (Patton 1999). For example, the tensions between the school's leadership team and Governing Board on the appropriateness of the AwL system were first conveyed by Head-Qualitative during the stakeholder analysis facilitated by the hierarchical mapping exercise during Phase 1 of data collection and it was clear from the diagram and commentary that governors were excluded from the interpretation process. The topic was then explored with two School Governors during Phase 2 data collection and it became clear that the Governing Body had struggled to fulfil its legal obligations to hold the leadership team accountable without pupil test data and that some had resented not having these matters addressed by the headteacher and had chosen to leave their roles. Reflecting on this and the other hierarchical maps, it became clear that the headteacher had the power to invite and exclude stakeholders during policy interpretation and/or during implementation, not just in School-Qualitative but in all six schools. This prompted a revisit of theories by John Novak on invitational leadership (Novak 2009) which argue that educational leaders should actively invite less-powerful members of the community to participate in the institution's offering – the leader has the power to invite. When applied inductively to the stakeholder involvement data in this DPhil study, this power to invite involvement seemed to lie with the primary school

headteacher. This prompted a read of Ball et al.'s (2011) article on policy actors and policy work in schools, and reflecting on those theories it seemed that the centrality of the headteacher's role was underplayed as this power to invite involvement from, and indeed uninvite from, stakeholders forms part of their role as a co-ordinator of stakeholders which is a new contribution to our understanding of policy enactment in a primary school.

To aid interpretation, it was decided to integrate analysis and discussion with findings in every chapter. Integration is discussed in mixed-methods literature when handling both qualitative and quantitative data (Åkerblad et al. 2021). In this thesis, integration has been used for combining the presentation of analysis and data which allows for immediate contextualisation, enabling readers to grasp the significance of the data without having to refer to separate sections. When readers encounter raw data, they may struggle to understand its implications without accompanying interpretation. By integrating analysis, researchers can provide explanations, highlight trends, and draw attention to key findings as they present the data. This immediate contextualisation helps readers to understand the implications of the results more easily, enhancing comprehension and retention. Further, the ability to see the data and its interpretation in tandem allows readers to form a more coherent understanding of the research findings, minimising the cognitive load and facilitating a more integrated and holistic processing of information. Combining results and analysis in a single chapter also improves the overall flow and cohesion of the document. Separating these sections can lead to disjointed reading experiences, where readers must constantly switch between chapters to connect data with its interpretation. A unified chapter ensures a seamless narrative, guiding readers through the research findings and their implications in a logical and coherent manner. This structure not only makes the document more readable but also maintains the reader's engagement and interest. The continuous narrative helps to build a compelling story around the research, making it easier for readers to follow the progression of ideas and understand the significance of the findings. This approach allows for a more dynamic presentation of the research, where the interplay between data and analysis can be more effectively showcased.

### 3.7 Ethics and Positionality

CUREC processes were undertaken prior to both phases of data generation. Informed consent was vital and so details of the study were written in user-friendly language for each category (teachers, parents, pupils) and sent to participants together with consent forms. These were then read at the start of each interview to ensure understanding about the study and the right for participants to withdraw, and verbal consent was recorded even if written

consent was provided. Confidentiality in terms of content of conversations was preserved for adult participants. Primary school-aged participants cannot sign consent forms themselves to participate in data generation or understand how their data may be used and so consent forms were completed by parents and encrypted transcripts were also sent to parents. This was particularly important with the Year 6 pupils in SchoolQualitative as the teacher had not selected pupils to participate or obtained written consent from parents and because all pupils wanted to participate multiple group discussions were conducted and then transcript verification was used as a method of consent. Confidentiality in-terms of no-one knowing which actors were participating is not possible in small primary schools during school hours. There is so little physical space that either people would see who was being interviewed or if a private room was made available, then others would know that it was being used for. The Year 6 pupil discussion in SchoolPAT and the parent discussion in SchoolQualitative were held in the staff room and other people did come into the room. In the pupil discussion, the support staff that entered stayed for the whole discussion and listened in but as it did not seem to be distracting the pupils, and it was not a private space, the researcher did not ask them to leave. Confidentiality was particularly important for teachers and governors as they were willing to share honest views on and experiences with the AwL system but did not want their headteacher to feel that they were criticising them. How quotations can be presented with a locatable voice without compromising identities has been carefully considered.

Data integrity was ensured through verification and triangulation. As suggested by Brinkmann and Kvale (2018), during data generation, the researcher continually reflected upon participant responses and verified their understanding with participants through follow-up questions which also provided an opportunity for further explanation and clarification. Verification was then sought for the data in a written format by sending participants encrypted transcripts which enabled participants to see what they had said and how it would look if quoted in my thesis or publications. In Phase 1, the six headteachers triangulated each-other as the same questions were explored, and the Phase 1 interview data with Heads Qual and PAT was triangulated with their Phase 2 interview conducted over one year later. Their comments about benefits to teaching and learning were triangulated with class-teacher data who also triangulated each-other. Pupils naturally triangulated their peers, by reminding each other of details, and complimented each-other by sharing experiences that illustrated points made. Single case studies located in just two schools are always going to have limitations in terms of demonstrating that findings are applicable outside of their immediate context and so triangulation with the literature at multiple points of the research process has been essential with two research projects being particularly pertinent. As detailed in the literature review, Poet et al. (2018) conducted interview research on the

enactment of AwL with the largest sample of primary and secondary school leaders and teachers and so their findings have been used to reflect against emergent data in this DPhil thesis. Also detailed in the literature review is the book authored by Thrupp et al. (2018) which documents responses to a new assessment policy in six primary schools in New Zealand. This research provides an ethnographic insight into policy enactment and has informed some of the data generation methods and presentation of findings in this thesis. Together with these two studies, reference back to AwL policy documents and the public management literature on policy enactment enabled researcher observations and interpretations on the data to be verified. Although data is unique to the context, common themes were extracted and analysed with reference to the academic literature. Had a multiple case design been developed exploring several policies then it would have been possible to understand the differences and the similarities of enactment between the cases which some researchers claim increases the credibility of the theorising (Baxter and Jack 2008; Stake 1995), increased representativeness (Gustafsson 2017) and others allow wider exploring of research questions and theoretical evolution (Eisenhardt and Graebner 2007). This is disputed by researchers like Dyer & Wilkins (1991) who argue that single case studies are better than multiple cases for developing high quality theory because a single case study provides a deeper understanding.

The data generation in this thesis was relational and social and so being mindful of one's changing positionality throughout the data generation process has been vital. As Hopkins (2007) comments, 'The ways in which researchers are read and interpreted by research participants are also an ethical consideration that requires reflection throughout the research process' (Hopkins 2007:387). For example, when interviewing teachers, the dynamic was of a professional discussion among peers, but in relation to headteachers, the researcher was always the junior partner as the headteacher was the gatekeeper for all data generation. They did not interfere with participant inclusion criteria and welcomed interviews with teachers, pupils and parents but they did control whether conversations with school governors were had and if they were which governors were spoken to. Considerations of positionality in relation to pupil participants were central to the researcher's approach. After seeing the researcher helping within the classroom, pupils seemed to become comfortable, perceiving the researcher as another adult that was part of the teaching and pastoral team. Pupils were told at the start of the interview that they were the teachers during the discussion as they were explaining to the researcher how assessment works in their school. This positioning of the interested outsider seemed to give pupils confidence and the responses of other students helped students understand what was being asked, the type of information that should be shared and acted as a memory-recall. Managing bias within a

small study is challenging as purposive selection and self-selection coincide with limited choices due to low response rates. Inclusion criteria were used by the researcher to purposively select data generation sites and categories of participants with, and emails were sent to all 44 schools categorised as primary on the county council website. The participant headteachers were self-selecting and all were successful in terms of SATs results and Ofsted judgements and presented as being confident in their AwL systems with no lower-attaining schools responding. All parents were self-selecting and as the response rate was so low it was appreciated that someone attended. Therefore, data cannot be used for reaching conclusions but can provide some ideas for further research. Pupil participants were selected by class teachers who were asked to provide a variety of attainment levels. It is possible that teachers also thought about who they thought would be articulate and interesting or with whom they had a better relationship or whose parents they thought may permit the participation. This was considered to be a necessary compromise to ensure that children who may not be emotionally ready to participate were not selected, and parents that may be hostile towards research participation were not provoked. As Bryan and Burstow (2018) explain, school-based research involves competing demands.

## CHAPTER 4: The Types of AwL systems being used

### 4.1 Chapter Introduction

This chapter presents data from six headteacher interviews conducted in Phase 1 (October 2017 - January 2018) and Phase 2 (March 2019) to show what Assessment with Levels (AwL) systems looked like in six primary schools. The CAWL report (2015) emphasised headteacher autonomy while providing reflection questions. Headteachers were asked to consider the insights that assessments should provide about students' knowledge, understanding, and independence, including inclusivity for pupils with Special Educational Needs and Disabilities (SEND). They were also asked to reflect on how to communicate assessment purposes and data to pupils, adjust teaching to improve understanding, support progression for quicker learners without deviating from the topic, and record and depict progress. The six participant schools created, purchased or were given, different AwL systems. Two schools used the National College of Teaching and Leadership (2014) report and Clive Davies' approaches as a basis to create Passports and Points systems; two schools used the National College of Teaching and Leadership report (2014) and Wroxham Alliance approaches as a basis to create Qualitative and Continuum systems; and one school purchased an electronic systems called School Pupil Tracker and another school was given an electronic system called Pupil Asset Tracker by their Multi-Academy Trust (MAT). The Passports system, developed by the school in 2014, focuses on core subjects like Reading, Writing, and Maths, using Performance Indicators and KPIs for each year group. Teachers compile evidence from pupil work to support assessment judgements, integrating assessment into their teaching while ensuring a manageable workload. The system includes Progress in Reading Assessment (PIRA) and Progress in Understanding Mathematics (PUMA) tests. The Points system, based on Clive Davies' materials, uses spreadsheets with curriculum statements for each year group. Teachers identify KPIs necessary for a pupil to be ready for the next year-group's curriculum. The system includes a nine-point progress scale and emphasises mastery statements for quick learners. Moderation meetings with partner schools ensure consistent expectations. The Continuum system, developed since 2014, measures progress and attainment based on an assessment continuum. It focuses on a growth mindset approach, allowing pupils to progress at their own pace. The system breaks down reading into strands with curriculum statements for assessment, considering handwriting, sentence construction, composition, and spelling. Teachers track progress termly, aiming for significant improvements in pupil performance. The Qualitative system adopts a qualitative approach involving in-class conversations with pupils and written explanations to parents. Teachers use daily classroom interactions to assess pupils,

focusing on progress rather than tracking. Pupil attainment is tracked through curriculum statements, with progress categorised as 'Good,' 'Slow,' or 'Excellent.' Moderation ensures consistent understanding and accurate assessment. School Pupil Tracker (SPT) is an electronic tool designed to help schools develop their own assessment systems. It supports planning, recording assessments, tracking, analysis, and report writing, aligning with National Curriculum objectives. SPT allows schools to customise Key Performance Indicators (KPIs) and employs both Detailed and Broad Tracking methods. Detailed Tracking evaluates a child's attainment using National Curriculum records, while Broad Tracking relies on teachers' professional judgement or test outcomes. The Short-Term and Long-Term Grids provide a comprehensive view of student performance, with personalised elements like diaries, targets, and progress cards. Another electronic system is Pupil Asset Tracker monitors student progress and attainment across various frameworks. It uses National Curriculum Age Bands and Point-in-Time Assessment (PITA) to provide clear metrics. The system's flexibility allows customisation of assessment criteria and the weighting of statuses such as 'Beginning,' 'Developing,' and 'Embedded.' The Tracker's colour-coding feature helps identify students not making expected progress, facilitating accountability and targeted discussions with teachers.

Subsequently, a discussion offers additions to the literature on mastery teaching and also a discussion on the similarities and difference between mastery teaching the other approaches encouraged by the Expert Panel for National Curriculum Review (2011) such as Assessment for Learning (AfL). Both the Expert Panel for National Curriculum Review (2011) and the Commission for Assessment without Levels (2015) stated that AwL was part of a wider reform of the curriculum that would shift towards mastery teaching. Bloom (1971, 1976, 1984) observed that teachers' traditional practice was to organize curriculum content into units and then check on students' progress at the end of each unit. He reasoned, that the feedback would be much more valuable if they were used as part of the teaching and learning process to provide feedback on students' individual learning difficulties and then to prescribe specific remediation activities and he termed this as mastery learning (Bloom 1971). In using this strategy, teachers organise the important concepts and skills they want students to acquire into learning units, each requiring about a week or two of instructional time. Following high-quality initial instruction, teachers administer a formative assessment (Bloom, Hastings, and Madaus, 1971) that identifies precisely what students have learned well and where they still need additional work. The formative assessment includes explicit advice about what students must do to correct their learning difficulties and to master the desired learning outcomes. When students complete their corrective activities (after a class period or two), they take a second, parallel formative assessment that addresses the same

learning goals of the unit but includes somewhat different problems, questions, or prompts. The second formative assessment verifies whether the correctives were successful in helping students remedy their individual learning difficulties. It also serves as a powerful motivational tool by offering students a second chance to succeed (Guskey 2010). Along with the corrective activities, Bloom recommended that teachers plan enrichment or extension activities for students who demonstrate their proficiency on the first formative assessment. Enrichment activities give these students exciting opportunities to broaden and expand their learning. Bloom believed that nearly all students, when provided with the more favourable learning conditions of mastery learning, could truly master academic content (Guskey 2010). A large body of research has borne him out: When compared with students in traditionally taught classes, students in well-implemented mastery learning classes consistently reach higher levels of achievement and develop greater confidence in their ability to learn and in themselves as learners (van Haren and Harroun 2019, Cundiff et al. 2020, Conley 2020, Kruckenberg 2024). Researchers have found that certain techniques are particularly useful such as, diagnostic pre-assessment with pre-teaching (Guskey 2010); engaging all students in high quality, developmentally appropriate, research-based instruction in the general education classroom (Wiggins and McTighe 2005); progress Monitoring through regular formative assessments (Hattie & Timperley 2007); high-quality corrective instruction following formative assessments (Sternberg 1994); parallel formative assessments (Guskey 2010); effective enrichment activities that provide valuable, challenging, and rewarding learning experiences for learners who have mastered the material and do not need corrective instruction (Tomlinson et al. 2006).

#### 4.2 The Passports System

When levels were removed in 2014, Head-Passports was Assessment Lead in their school and commenced designing what a new system could look like by reviewing the research material while still expecting that the DfE may publish a centralised system,

'We found ourselves in a position where we have fundamentally more tricky curriculum, knowledge and skills to impart that were fundamentally different; and we had no way of measuring that other than outcomes led objectives. So, that was the starting point for us' (Head-Passports).

Head-Passports said that they knew that there would be lots of groups and companies looking to make money from this process and decided to reject those products due to their

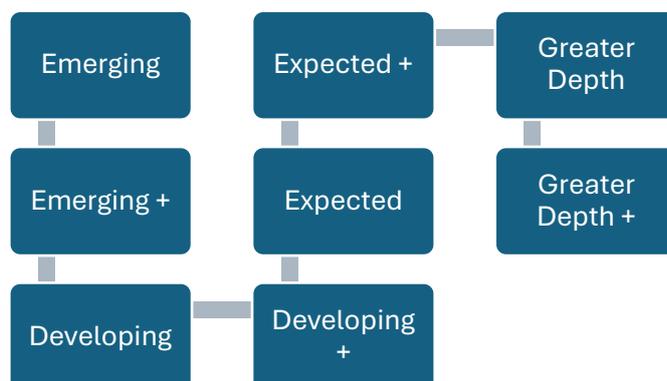
profit motive and viewed it as an opportunity to develop their own approach. Initially, Head-Passports looked at a company that they had used previously which was 'Not As We Know It' by Clive Davies<sup>9</sup> who produced objectives that were broken-down out of the National Curriculum. They then looked at what they perceived would be tested, and what the core subjects were, and initially 'mirrored what had gone before' (Head-Passports). At that point, objectives were tracked and crossed-off in a paper exercise and then transferred from Year to Year around subjects other than Reading, Writing and Maths. Teachers in SchoolPassports took those objectives and considered what they looked like for each Year group and decided to develop 'Passports to Success' for Reading, Writing and Maths.

To discern the attainment of a pupil, teachers create a body of evidence that supports the teacher assessment. These Passports are used in a formative way throughout the Year as teachers decide whether a child is on track. It was felt that assessment should feed into what the teacher does in the instruction process within the classroom, so teachers also looked carefully at how they mark and support pupils daily in the classroom. Within each Passports, they outlined Performance Indicators, and within each of those there were fundamental non-negotiables (KPIs) which formed the expectations for the Year group. The work-life balance of teachers was important to Head-Passports so they decided not to break down objectives into further detailed objectives as this may be too onerous and could possibly detract from the cycles of teaching and learning – 'which is fundamentally the most important thing that is happening' (Head-Passports). Whether a pupil had met the age-related expectations for the Year was determined by the percentage of those objectives that a pupil needed to attain. Head-Passports explained that teachers felt that 100% was not going to be realistic, but equally 60% was going to be too low because that would mean that 40% of the objectives had not been met which probably would include some of the non-negotiables, therefore, they decided that if a pupil attained 85% of the Year group objectives they would be considered to have met age-related expectations.

---

<sup>9</sup> <https://www.notasweknowit.co.uk/>

Figure 17: A Pupil's progress through a Year Group in the Passports system



Head-Passports commented that getting information across to parents so that they know and understand, 'has been an enormous challenge' (Head-Passports). So, teachers held more consultations and moved to producing three report cards during the year and Head-Passports runs assessment briefings for parents as there have been several changes to the assessment approach that they may be unfamiliar with. Head-Passports perceived that this has enabled parents to understand the terminology and content of the reports although primarily parents seem to want to know whether their child is meeting expectations or not. Parents are also welcome to bring children to parents evenings. Head-Passports insisted that teachers are careful as to what they share with pupils about their attainment. Key Stage 2 pupils take their Passports and know what is expected of them and are encouraged to be actively engaged with the objectives, but Key Stage 1 pupils are considered to be too young to be able to engage. Teachers are expected at the start of a lesson and the end of a lesson to mention learning objectives and success criteria. Extra feedback and separate trackers are provided for children that are in receipt of Pupil Premium or have SEND because they may not have quite as much progress in the same way or at the same pace as other pupils. The language of classification is not used in classroom conversations but is told to parents in the aforementioned report cards.

#### 4.3 The Points System

Head-Points had received negative feedback from local Headteachers who had purchased online packages and so they decided to refer to Clive Davies as they had purchased his materials previously 'Not As We Know It'<sup>10</sup> document as their starting point which uses

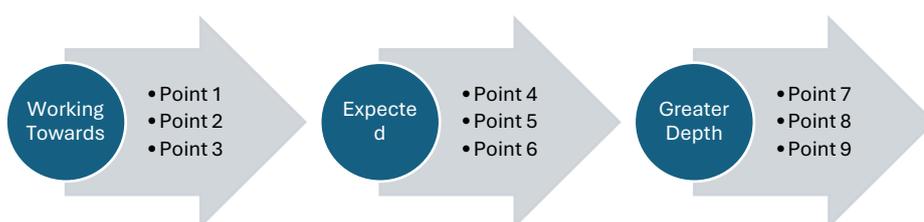
<sup>10</sup> <https://www.notasweknowit.co.uk/>

spreadsheets that have curriculum statements in each row and the school developed their own system based on this for each Year group connected to End-of-Year Expectations. Attainment is measured by the percentage of statements that children meet within their Year group. So, as they progress through their work, more statements are highlighted on the system and as that percentage increases the school decides 'on a grade', with pupils attaining 80% of those statements being the grade of a pupil meeting the expected attainment (Head-Points). Teachers then reflected on the statements, and considered which ones students had to have, and those that were less important, and set the former as KPIs as these were key statements that children needed to know to access the next Year's curriculum. It was decided that to attain the 80%, all of the KPIs must be attained, that is, a pupil that was not secure in their understanding of or competent in their application of a KPI could not be considered to be working at the Year group's expectations.

The teacher's judgement determined whether a statement is highlighted. Head-Points decided to start off by mapping the percentages throughout the Year but found that this was inaccurate because some of the statements were so broad that a teacher couldn't complete some statements until the end of the Year. For example, the statement on mathematical measures related to the three aspects of capacity, weight and measure which were being taught in a different order in different classes. Thus, Head-Points worked with the maths and literacy coordinators to split the statement further so that teachers were able to show that pupils had learnt something and give credit for that. They then looked at progress and put in a nine-point system which covered three levels: Working Towards, Expected and Greater Depth. Points 1 to 3 means they're Working Towards, 4 to 6 means Expected, and 7 to 9 is Greater Depth. Head-Points explained that this meant that 'parents could then understand how close they (their child) are to being a Working Towards' (Head-Points).

---

Figure 18: SchoolPoints AwL System Categories



The system continued to develop as teachers try different methods and discuss what works better for them, for example, the school moved to Maths Problems as their main teaching scheme and so were using more statements from that rather than Clive Davies documents. As Head-Points commented, 'I think we've refined it is what I would say. It's been a gradual... It hasn't stood still since we started in 2014. Each year we've responded to what questions do we need to answer in order to work out where children are' (Head-Points). With the Clive Davies documents, there are Exceeding statements so when a pupil has reached Working as Expected, they can move on to the Exceeding statements which are aimed at deepening their level of understanding, that is, mastery. These mastery statements do not link with the next Year's curriculum and therefore, Head-Points suspects that all schools may be struggling with how to track progress from one Year group to the next due to each pupil starting again at those statements of 'Working Towards' all the time. To address this, Head-Points and her colleagues are considering,

'using the idea that if a child is 'Working At' the end of Year 3 in order to make expected progress they need to be Working At the end of Year 4, and so on. But if we are going to accelerate progress and the children are doing better than that, then they need to be upping that level on the nine-point system which we've now got in place - we can say that they've moved so far. So, we can say that they are 4.4 which would be at 'Expected' in Year 4, but accelerated progress could be 5.6 which means they had a higher level of 'Expected' by the time they leave Year 5' (Head-Points).

Year 2 and Year 6 pupils are on a different set of statements because they are working on the DfE Interim Framework and so they would have had to complete 100% of them in order to be 'Working At. There are fewer statements than there are in Years 1, 3, 4 and 5, but

Head-Points explained that teachers still need to ensure that pupils are making progress in Year 1 in order to set them up to succeed in Year 2. For Year 2 and 6 teachers, The County Council runs SATs training to help understanding of the expectations. Head-Points also explained that the County Council employ consultants to run moderation on a sample of 25% of schools particularly examining that their SATs writing levels are accurate. SchoolPoints were moderated by them in 2016 for both Year 2 and Year 6 work.

The school employed a non-teaching Deputy Head, who manages all data, and they have developed a filter so that data can be viewed for different categories of pupils such as boys, girls, SEND, English as an Additional Language (EAL), or summer born, so that comparisons can be made between different significant groups. The Deputy Head takes the lead in terms of developing a system and also generating an analysis which they provide to the subject coordinators who consider how it relates to their subject and how best to add to it. It is then also passed to teachers who are also trying to complete it and suggesting possible problems. For example, analysis of all the tests and the analysis of all the classroom data showed that punctuation and grammar was correct in the tests but then don't use it in their writing, and therefore teachers were asked to focus upon this. Thus, the assessment system allows Head-Points and colleagues to analyse the strengths and weaknesses in the teaching and the outcomes which then influences teaching.

Sharing attainment with pupils depended on their age. For example, in Year 1, pupils have a rocket ship with their target on the card and they work their way through it and when they achieve it they get stars and then they meet the new target. In Year 2 pupils have a Charlie Checker who has a couple of targets. In Key Stage 2 pupils are more likely to have a selection of targets that they select from and a tick list where pupils tick when they think that they have achieved something. In Governors meetings there have been focused sessions on data so that governors can see what it means and what statements look like. Teachers have also run sessions for parents about how to understand a report and how to understand what we're talking about in terms of the whole language of assessment and the Deputy Head also used the school's YouTube channel. The Ofsted report in 2015 had informed Head-Points that parents wanted more information about how their children were progressing, so Teachers now report to parents four times per year which entails two face-to-face meetings with parents, and two written Reports. The first Report goes out in January and informs parents where their child is now and the likelihood of them being at the Expected level at the end of the Year; and the second one in July uses the nine-point system to explain where their child has reached in the Year. Consultations with parents revealed that they wanted tick

boxes stating 'they are expected to be at this, this is their strength, this is their weakness, these are there areas of development' (Head-Points).

#### 4.4 The Continuum System

Since 2014 Head-Continuum and colleagues have developed their own system for measuring progress and attainment which is based on an assessment continuum. Head-Continuum and one of their Assistant Heads relished the opportunity to create their own system, 'our ethos has always been to try and do some things differently and we felt that this was a big opportunity... we took over one year, and we've only just finished our final tweak really of taking out some of the things that we didn't need' (Head-Continuum). As a starting point, the Head-Continuum and the Assistant Head commenced conversations with each other in 2014 after reading the Expert Panel Report (2011) and considering the principles behind their ideal assessment system and the evidence that they may look for. They didn't want to purchase testing systems such as PUMA or PIRA as they felt that their pupils had particular challenges that needed a tailored approach,

'We wanted to do something that was right for the children with a very mixed demographic with 50% of children with EAL, 35% children with pupil premium, 35% of children on the SEND register, and so it was a very different picture to other schools who could just buy something in and probably could make it fit but we found that very difficult' (Head-Continuum).

Initially, the diocese and local authority provided no support or guidance, 'yesterday you used to use levels, tomorrow you won't - nobody said anything about it for a year. I didn't have any resources or training' (Head-Continuum). Over 18 months, Head-Continuum and the Assistant Head met with Dame Alison Peacock several times, and also with other school leaders but perceived that those leaders were developing systems that were similar to the previous levels system. Head-Continuum explained that they didn't want to 'marry it too tightly into the SATs system of testing which actually makes it very difficult because the only thing that people want to know is how you are going to do on the SATs? And when I say anyone, I mean externally, and also parents' (Head-Continuum). The concept of a continuum is based upon Carol Dweck's work on a growth mindset mentioned above, whereby a pupil is allowed to develop at a speed that suits them rather than in relation to Year-group expectations, and therefore the school had initially abandoned using such expectations. However, governors and visitors from the diocese and local authority insisted that they create and use them. As Head-Continuum explains,

'it would work without this expectation of Year groups, and children would move throughout the continuum and we would, through close monitoring of this, make sure that we picked-up children that sort of hadn't made as much progress as other children; or that teachers were teaching appropriately to ensure that children were able to meet the expectations of the skills; but in a way how we have finally done it is a numerical system which we had tried to avoid' (Head-Continuum)!

Figure 19: Assessment Principles behind the Continuum

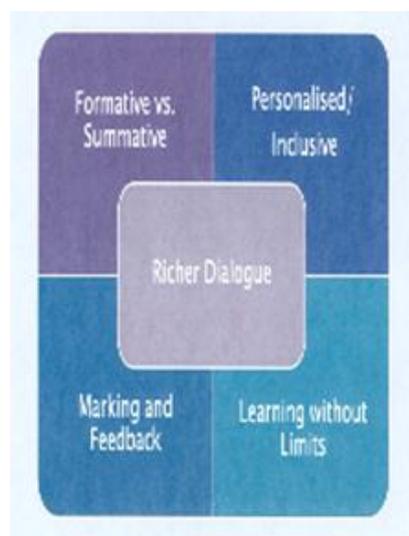
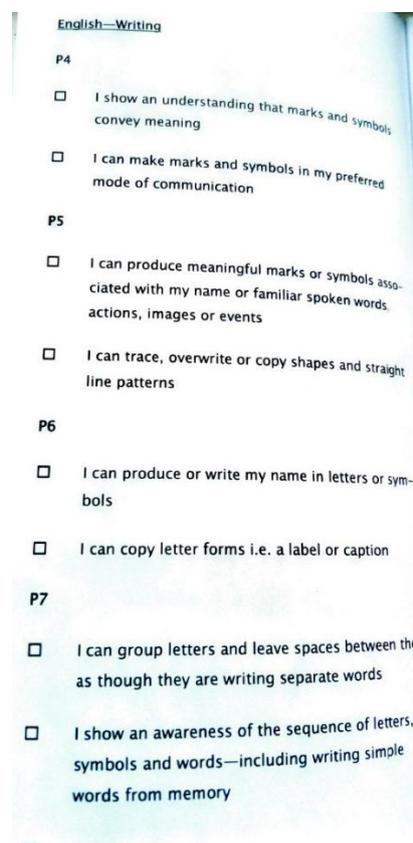


Figure 20: The Process used in the Continuum



In the Continuum, reading has been broken down into different strands and then a curriculum statement has been put in each of the strands. The assessment takes into account handwriting, sentence construction composition and spelling and to attain each point a pupil must produce one written piece of dialogue. The children and the teacher use these statements to assess what the pupil can do and then the teacher can plan the next lesson with that in mind. Thus, these statements are intended to work both as an assessment tool and a planning tool. The system does not attempt to assess the whole curriculum some fundamental aspects that the teachers think that the pupil needs, ‘these are the significant points that we measure’ (Head-Continuum). A curriculum strand will have 19 points that pupils work through as they progress through the school, with teachers being able to see points that were not attained (in the current or previous Year-groups) so that they can focus teaching on those aspects. The statements are not linear but are intended to be somewhat sequential from point 1 to point 19, however, Head-Continuum says that teachers will teach according to the needs of the pupils rather than following the sequence, and then the Deputy Head will split the data into Year-groups.

Figure 21: Example of Continuum ‘I can’ statements within English Writing Points 4, 5, 6 and 7



Conversations between the SLT and teachers centre around whether a pupil has achieved 2 points of progress in each curriculum strand. These need not be sequential and there are no KPIs within the points. A pupil is expected to make 11 points of progress during the Year which should be spread through different strands. Teachers will examine the areas that progress is being made termly and then focus on areas that seem to have less progress. The Assistant Head monitors the data on the spreadsheets that they have developed and they can also break it down to groups such as children with EAL, on pupil premium and these are born at the start and at the end of the year. Teachers provide the previous point the pupil had attained in a domain and the number that they have now attained without referring to expected progress. The teacher will then also discuss the next steps that the pupil needs to take,

'if they're on 14, it automatically generates what 15 is to let the parents know that they need to do. So, we've tried to have a learning emphasis rather than a number, and we group the numbers to say... we do say actually 'Expected', but we group in quite a big band, and "we expect your child to be somewhere about there", because we know the children may progress at different rates but get there in the end' (Head-Continuum).

For Year 2 and Year 6 teachers, the school gathers teachers from eight other primary schools to moderate using the SATs Interim Standards which Head-Continuum thinks works effectively, however, they do not think that this could occur for other Year-groups because the AwL systems are different. Head-Continuum feels that local authority advisors had wanted their AwL system should tally with the SATs system and so the school did try to do this, however, they found that the scoring system was too different and what was more important was to focus on the learning gap analysis in-which some test papers are used. The main cost that the school has incurred has been in terms of SLT and teacher time. Head-Continuum estimates that SLT time will have cost £15-20,000 and now transferring their paper system onto a web app will cost another £15,000. Teachers have received approximately 40 hours of ongoing training which Head-Continuum made a priority as they wanted the AwL system to be meaningful to the teacher and helpful for their lesson planning. Head-Continuum admits that they had not factored-in the turnover of teachers which would require continuous training in the AwL system. Head-Continuum recounts that SATs results were poor in the year prior to the AwL system being enacted with the school being in the bottom 3% of the lowest attaining in maths within the county. After two years of AwL, the school is now in the top 7% in maths, and in the top 3% in reading for disadvantaged pupils. Head-Continuum proposes that their AwL system may be more demanding than the SATs

curriculum and therefore, 'we automatically had higher expectations' from Year 1 which has meant that less support is being required by pupils in Year 6 (Head-Continuum). They also perceive that this has enabled the school to avoid narrowing the curriculum. Head-Continuum reflects that in their attempt to accurately capture the progress of each pupil they have created a cumbersome system that is unwieldy,

'The challenge with it is probably that it gives us too much information. What information do we actually need and who needs it, and who wants it and when?'

The system is very time-consuming on teachers because there are so many different strands... and we're trying to develop into something which is more practical' (Head-Continuum).

The Continuum system was really something enacted by Head-Continuum and their Assistant and it was heavily dependent on them manually entering and analysing data.

#### 4.5 The Qualitative System

For this school, the CAWL Report (2015) was considered an opportunity to create a system that worked for their pupils and to reflect upon what they considered to be the purpose of assessment. Head-Qualitative and their colleagues decided that it would be best for the teachers to develop a system that would work for them in their class to gauge whether the pupil could or could not achieve certain objectives. The English and Maths subject leaders invested many hours initially designing the system and looking at other systems and considering whether they could work for their school. Head-Qualitative said that they were not involved in the designing of the system but described themselves as 'the instigator of the change' who then delegated the task to classroom practitioners because they felt that 'it had to work in the classroom for it to be useful, and if it didn't work in the classroom then it was of no use you know in the system' (Head-Qualitative). Parents were not involved in the design, however, they are now involved in the discussions with their pupils about their progress and attainment on a termly basis. Key Stage 2 pupils were interviewed about assessment approaches that work best, and with the input of other teachers and Head-Qualitative, the school decided upon creating their own system which is continually being refined and has taken until 2018 to fully embed. As Head-Qualitative recollects,

'The first year of us implementing it - which would have been 2015 to 2016 - I had absolutely no data as a Head teacher which was a very scary year. There was the year that we were trialling and trying to implement a new system. And

then the year after that was very much about us trying to figure out what data we would be giving to governors and how that was going to look. And really, it's only now that I feel confident that we have got ourselves into a rhythm and routine with the assessments' (Head-Qualitative).

Whereas previously teachers relied on termly tests to provide the assessment which would provide a level to the pupil teachers now have to be sure that they understand their pupils, and that they are confident of being able to give a judgement on those pupils. Head-Qualitative was also very keen to restore professional trust back in teaching by allowing teachers to assess in their own way in the classroom using AfL and using those approaches to discern exactly what the pupil can and cannot do and how to fill the gaps. Thus, the AfL system is not separate to teaching but teachers are using questioning in a classroom on a daily basis, and then speaking to Head-Qualitative about what pupils can and can't do, and exactly what gaps they need fill and exactly what their next steps are and how they going to fill those gaps. Within classrooms, verbal feedback groups are used so the pupils have quality feedback which is thinking about what they need to do to improve and how to make their pieces of work better. Interviews with pupils were conducted to discern the types of feedback that works best, and verbal feedback was considered preferable, which has been confirmed by improvement in the pupils' work that teachers have noted. Head-Qualitative comments that, 'those are the kinds of conversations that we're now having, which were not being had before' (Head-Qualitative). The conversations between teachers and pupils do not examine whether the pupils are on track, but how can prove their progress. For example, every time they start a piece of writing, pupils may have,

'to think about "how is this piece significantly different to the previous page?", "how was this piece significant progress from the previous one?", "what changes have I implemented?", "how have I made it better, what have I made better?". So, everyone, parents and pupils, are constantly thinking about "what am I doing to improve my work?". And those are the types of conversations we're having now. So, rather than saying you're on-track, it's more a conversation of "this is where your next step is, and this is what we need to do to get you there". And it's constantly getting the pupils to think about what the next steps are but more crucially how they can get to those steps how they achieve those steps' (Head-Qualitative).

Although Head-Qualitative has kept teacher well-being and workload central to their motivation behind removing testing and written marking, she admits that 'Teachers have

found that extremely frightening and that has been one of the biggest challenges of implementing it'. Head-Qualitative recounted that they would find classroom set-up for testing simply because teachers did not know how else to assess and some have found the changes too much to continue with, 'I've lost a couple of teachers over the fact that we've done this because of the change. And it has been really difficult. It absolutely has' (Head-Qualitative).

Pupil attainment is tracked through an in-class tracker of curriculum statements. Progress is not measured numerically but by the progress that teachers are seeing in pupil books and through interviews with pupils and through pupil progress meetings. Pupils are then graded as making 'Good Progress', 'Slow Progress', or 'Excellent Progress' and so a pupil could have very low attainment but making excellent progress. Where a pupil is evidencing that they can achieve particular curriculum objectives independently then teachers are able to assess whether or not they are confident in certain key areas and this is how the school tracks attainment towards the end of year expectations. These expectations have been taken from curriculum objectives and teachers set what they expect pupils to have attained at the end of each Year group. Teachers do not deem that a pupil has to have 100% of the expectations but use their judgement to decide whether a pupil is approximately at the expectation even if they do not attain all statements. The exception are Year 2 and Year 6 where pupil attainment is based on the KPI's set by the National Curriculum Interim frameworks. Pupils are provided challenge and depth through Assessment for Learning (AfL) techniques employed by the teachers, however, interestingly (and in-contrast to the other participants in this study) the school has not followed DfE advice which states that pupils should not learn curriculum material from future Year groups to challenge some high-attaining pupils. This is also the case for SEND and other pupils who may be Working Below the Year group standard as a teacher may identify a gap in understanding which will mean revision of a concept covered in a previous Year group.

#### 4.6 School Pupil Tracker System

School Pupil Tracker (SPT) is a package that has been available for schools to purchase since levels were removed in 2014 and includes a step-by-step development plan for schools to use to develop their own AwL system with SPT. It attempts to aid planning, assessment recording, tracking, analysis and annual report writing. The School Performance Tracker (SPT) system emphasizes adherence to the National Curriculum objectives, offering a structured approach to pupil assessment. The system is designed to record detailed judgements, aligning with what is anticipated to be tested by the DfE and encapsulated

within the National Curriculum standards. SPT restricts the modification of these core objectives, although it allows schools to tailor Key Performance Indicators (KPIs) to their specific needs. The two primary tracking options provided by SPT are:

**Detailed Tracking:** This method aligns with Department of Education's guidance for the new National Curriculum, placing a significant emphasis on formative assessment. It utilises the National Curriculum records integrated within SPT to automatically determine a child's attainment level, assigning an Estimated Band, such as 'Y2 High'. This band is then used to auto-populate the Short-Term grid, which represents the current school year. The system also automatically calculates expected progress and Age-related Expectations (ARE) comparison percentages. The feature's ability to provide data on year-on-year attainment was a key factor in its adoption, as articulated by a Headteacher who believed it to be a requirement for Ofsted evaluations.

**Broad Tracking:** This option allows for professional judgement by teachers regarding a child's attainment or the use of test outcomes to benchmark against the National Curriculum standards. Teachers can directly use the National Curriculum documents to assess a child's current attainment level and input tracking scheme codes, like 'Y2 High', into the Short-Term grid. Alternatively, the use of standardized or scaled test scores provides a benchmark for the required standards. Both tracking options aim to facilitate a comprehensive and adaptable assessment framework, enabling schools to monitor pupil progress effectively while adhering to national educational standards.

The implementation of a Short-Term grid to track a child's attainment and progress represents a strategic approach in educational assessment. This system allows for the automatic calculation of expected progress and Attainment and Progress in Education (ARE) comparison percentages, providing a streamlined method for educators to monitor student development. The Broad Tracking option, as referred to by the School Performance Tracker (SPT), emphasizes the importance of broad judgements over meticulous record-keeping for every objective. Such an approach aligns with the principles of formative assessment, where the focus is on the overall progress and attainment of a child rather than an exhaustive analysis of each learning objective. This method can potentially reduce the administrative burden on teachers, allowing them to dedicate more time to instructional activities and personalized student support. Moreover, it facilitates a more holistic view of a child's educational journey, accommodating the natural variability in learning pace and style among students. By focusing on broader educational outcomes, educators can tailor their teaching strategies to better meet the diverse needs of their students, fostering an environment that

encourages all-round development. The efficacy of such tracking systems is contingent upon their ability to adapt to the unique context of each educational setting, ensuring that they serve as effective tools for enhancing learning outcomes and supporting informed decision-making by educators. Together with that the integration of Short-Term and Long-Term Grids allows teachers to discern gaps in understanding on the current topic and over a longer period of time. The Short-Term Grid captures immediate academic activities and attendance, providing a snapshot of a student's current engagement and achievements. This data then flows into the Long-Term Grid, which offers a comprehensive view of a student's progress over time. It includes a historical compilation of assessments, attendance records, and foundational stage data, which are crucial for identifying trends and patterns in learning. Additionally, the system allows for the inclusion of personalised elements such as a diary for anecdotal records, pupil targets to set and review goals, and a progress card that offers a detailed report of a child's development. The inclusion of historic annual reports and an evidence portfolio, which can contain uploaded photographs and documents provides teachers and parents with a multifaceted view of the child's educational journey. Tracking Points offers a numerical method to assess academic development. The TP scale is designed to increment by three points annually, establishing clear benchmarks for student achievement at various stages of their education. For instance, a TP of 15 is expected by the end of Year 2, progressing to a TP of 27 by the conclusion of Year 6, 'The tracking system really comes into play when we are identifying termly which children have not made progress and then that can feed into teaching and learning because that gives you the information to have that conversation with the teacher' (Head-SPT).

Golden Codes by SPT serve as benchmarks for student achievement within the curriculum framework, providing clear indicators of a student's progress and attainment. For instance, the 'Y6 High' Golden Code signifies that a student has met the comprehensive curriculum requirements for Year 6. This system also allows for the differentiation of learning, where students who demonstrate proficiency can be further challenged by deepening their comprehension and application of the material. The Golden Codes are designed to be utilized in a similar fashion to the previous levels, offering educators a method to gauge a student's position in their academic journey. This shift reflects a broader trend in education towards more nuanced and flexible assessment models, which aim to capture a fuller picture of student learning and development. One of the criticisms of levels was that pupils labelled themselves according to their attainment, and the headteacher explained that they do not use these codes to feedback to pupils, 'We share objective tracking with pupils that doesn't have any names, we don't give them codes we don't say whether that's good or bad. We just

say you can do these things, you can't do these things yet. That's quite literally it for pupils' (Head-SPT).

The SPT system offers the capability to automatically integrate different coding systems such as levels, P-Scales, Golden Codes, and EYFS Bands and generating a comparison and calculation of progress across these diverse data types. For a fee it can also be adapted to the needs of a local authority or MAT. The Annual Report writer feature use Golden Codes to generate a report for parents,

'Parents reaction has been quite good because what it will do is it tracks progress over time - years in fact - and they get a graph. So, where you know that children don't make progress for a term and then catch up the following term, parents don't get quite as panicky about that as they used to, because they can see that over time children are still moving up... They don't want codes. I spent a very long time because this system produces golden codes - which are levels (whispered) - and I spent a very long time trying to get parents to understand what these golden codes were and we realised after a year or so that we were wasting our time because that's exactly what parents want to know: they want to know, "is he about the same level as every other child in Britain; is he doing better or not? Is he speeding ahead or is he slowing down?" That's all they want to know' (Head-SPT).

As the schools in the Multi Academy Trust (MAT) use different assessment systems, Head-SPT described moderation as challenging but have started with using comparative judgement of pupil work in Writing, 'we have resorted across the partnership of schools, and across the Multi-Academy Trust (MAT), that this is the expectation for a child of this age, this is a better example, and this is a less capable example' (Head-SPT). The MAT was newly formed and felt that the use of standardised tests would help consistency, 'we have used national tests across schools. That's certainly something the MAT that we are a part of will do so next year with every Year 3 sitting the same test at the same time to give them some data' (Head-SPT). These quotes convey that in the early days of AwL, there was considerable concern about how to ensure educational standards within schools and between schools following the removal of levels and so they were exploring some standardisation in tests across schools and a common understanding of what expectations there should be.

#### 4.7 Pupil Asset Tracker System

The Academy Trust that SchoolPAT belongs to contracted the company Pupil Asset to create a bespoke system based upon the feedback from a meeting of assessment coordinators within the 30 schools that form the Trust. In SchoolPAT, pupils are expected to make five steps of progress in an academic Year so that they finish the Year at an 'Embedded' stage. For example, if they have left Year 3 in an Embedded stage, they will leave Year 4 in a 'Year 4 Embedded'. As SchoolPAT had been placed in special measures by Ofsted for historically poor performance, Head-PAT has set a target of six or seven steps of progress within a school Year, however, parents and pupils are informed of Age-Related Expectations (ARE) only in Years 5 and 6 when they are being prepared for National Standard Attainment Tests (SATs). If a Year 3 child is at Embedded they can attempt deeper Year 3 work, whereas prior to AwL, they would have been covering the Year 4 curriculum as the aim was to move through the levels. Within the first two weeks of a Year, teachers establish a baseline and then create a target for the end of the year as a percentage for the class and then each pupil is tracked individually. After teachers have set targets for the end of the year, Pupil Asset tracks how many steps of progress each pupil are going to need to make to attain that. So, for example if a pupil left Year 3 as Developing+ in a particular aspect (therefore didn't achieve Year 3 Embedded) they have an extra step of progress to make to get to Year 4 Embedded by the end of the Year. The Headteacher explained that the Year group targets are based loosely around national Key Stage 2 results, so, for example, if National results stated that 75% achieved the expected standard for maths, then the school would aim for a minimum of 75% of the class achieving ARE although that would be adjusted according to the prior attainment of the Year group. Although PAT has the function, the school did not use Key Performance Indicators (KPIs)<sup>11</sup> and throughout the Year there is ongoing formative assessment in pupil books and those notes enable the pupils and teachers to know what the next step in their learning is. However, the teacher would not relate their knowledge directly to the tracking system language as that may lack significance to the pupil. To parents, teachers present the pupil's attainment and progress at the end of the Year using terms such as, Better than Expected Progress or Expected Progress, 'they seem happy with knowing whether their child is meeting the expected standard or not' (Head-PAT).

Pupil Asset has provided the school with a designated person whom headteachers and assessment coordinators can contact for support or who can adjust the system according to

---

<sup>11</sup> They did introduce KPIs by Phase 2 data collection 16 months later.

the individual school's needs. Head-PAT and Deputy Head-PAT conveyed that the Pupil Asset Tracker helps them to hold teachers accountable because the colour coding enables them to instantly identify where pupils are not making the progress that they should be which enables them to have conversations with staff about why and what is happening for individual children. They felt that this is a significant improvement from the system with levels,

'We had some levels which were a bit of nonsense. If you had a 4b or 4c, there's no way to actually track whether a child is all there, but this is a very clear tracking system so you can literally see them moving through them... now you're not talking about levels, you're talking about, "for this year are these Year 1 children doing what they should be doing are these Year 2 children doing what they should be doing?". Whereas in the past, sort of the progress level 1 and level 2 was always a bit of an issue - it always appeared wishy-washy'... now every Year-group is expected to make the same progress, and we are tracking it the same way and so it's much clearer' (DeputyHead-PAT).

The school are providing enrichment activities for pupils who have completed Year-group objectives, however, the majority of pupils come from deprived communities with limited educational support at home and so teachers focus on ensuring that they are secure in their understanding of core concepts,

'we will look at how to challenge and stretch them, so for example, if it's maths and they've covered all the objectives for maths, we'd be looking at how actually we can enrich that so that people can develop their understanding, but we leave the Year 4 curriculum to Year 4 teaching... in our school we are riddled with... because really it's about embedding and making sure the students really are very secure and understand to the very best for that year group before they move onto the next, at the right time; and at the right time they should be doing that' (Head-PAT).

The absence of the common assessment language that levels provided has meant that the Headteacher and Deputy considered moderation to be more important than it was prior to AwL. They conducted this both within their Trust and also with three other local schools where examples of pupil work was discussed by teachers and agreement was sought as to the features of an Emerging, Developing and Embedded piece of work. This moderation is conducted with all teachers in Year 6 and in Year 2 in reading and writing in English, science

and maths, with teachers completing forms that transcribe those conversations and whether there is agreement in the judgements. And if all the other teachers agree with a teacher's judgement, the school leaders understand that to mean that the teacher's judgement is secure. Head-PAT felt that the SATs system was mis-matched with Pupil Asset Tracker (they also felt that the system with levels was mismatched). They felt that this was due to the inadequacy of the SATs system rather than the AwL system and have not adjusted their AwL system to align with the SATs. As the Headteacher remarked,

'I've got no problems with Assessing without Levels, it's given us more responsibility for how we say we are tracking and how we are monitoring, and what we call it and what we do about it, but it doesn't tally at all with SATs and SATs are archaic. They're trying to really advance the way we are assessing on an ongoing basis and giving professionals responsibility but at the end of the day they haven't given us any more responsibility at all because it's still all measured in a punitive way in that test on the day so the two systems aren't working together and they never will unless they take SATs away; or they think of a different way of doing it' (Head-PAT).

Formative assessments, tailored teaching methods, and communicating students' progress and future learning goals to both students and parents are fundamental aspects of education. Consequently, headteachers were never given the freedom to eliminate these practices, as emphasized by the CAWL Report (2015). AwL provided headteachers with the flexibility to carry out these educational functions in line with their values and the specific needs of their school community, ensuring both headteacher satisfaction and a high level of standardization across schools. National accountability tests in Years 2 and 6 in primary schools are designed to ensure a consistent evaluation of student performance nationwide (CAWL 2015). These tests are crucial for all school stakeholders, as they influence league tables and Ofsted inspections. As a result, the six schools in this study used the Year 2 and Year 6 National Test Standards to benchmark their AwL system and compare with partner schools. This practice is consistent with Braun and Maguire's (2018) findings under the levels system, where despite widespread disapproval of extensive testing, active resistance was rare due to the pragmatic compliance of school leaders and teachers (Keddie 2018). Ball (2003) observed that these accountability mechanisms compel leaders and teachers to organize their efforts around inspections and test results.

Fifteen months later during Phase 2 data collection, Head-PAT explained that the AwL PAT system had been adapted by teachers according to their needs. Despite being given an electronic AwL system by their MAT, Head-PAT explained that that the enactment of AwL was an ongoing process (Maguire et al. 2015) and they had simplified their progress reporting system as they could better discern what was needed as they were using it,

‘the system has been adjusted and changed, which is right, I think, for system leaders to continue to look at the system as being fit for purpose. So, we do not have the Emerging, Emerging + etc etc. I think there were 6 altogether and the Expected was 5’ (Head-PAT).

SPAT-Y4T explained that they found that the original AwL system was more complicated for them to interpret but easier to show movement as it involved all pupils starting at the bottom of the ladder of 6 levels and they then had to climb that during the Year,

‘I found last Year’s terminology of the levels... I was very confused. So, it was completely new to me and it was a lot of gradations – +, + developing, and all that kind of thing, so it took a long time to work-out exactly how that translated to in the class’ (SPAT-Y4T).

SchoolPAT’s updated system was simplified and depicted each pupil in green if they were At the Expected level, in amber if they are less secure, and in red if they are below the Expected level in a concept or skill based on a teacher ticking can do statements on the pupil’s electronic dashboard making judgements more binary,

‘While the new system is simplifying, so it is far easier for me to describe who is at where they should be, who is Above, who is Below, but I think it is harder to pull out from that progress because you might have a child who starts the Year at the Expected level and finished the Year at the Expected level. So, I guess the interpretation is, yes they made progress through the curriculum although it is harder to show it’ (SPAT-Y4T).

While showing the tracking system on the laptop, SPAT-Y4T explained the difficulty that they face for showing gradations of attainments within a classification especially for higher and lower attaining pupils as they may be classified as at Expected progress throughout the Year rather than moving through levels,

'it is quite possible that at the end of the Year she would still be At Expected, and then it does not demonstrate the progress so much. I mean if you look at this boy, so he has moved from working just Below to becoming Expected, so you can clearly see that he has crossed into the other group' (SPAT-Y4T).

They proceeded to explain that under the original Pupil Asset Tracker approach that the school was using at the time of Phase 1 data collection, monitoring each curriculum statement had become unwieldy and so a few Key Performance Indicators (KPIs) were chosen which are now monitored through pen and paper rather than electronically. Meeting the KPIs is now the definition of a pupil being at the Expected level and usually a photo of the pupil's work demonstrating that they can do it is also uploaded onto the electronic system at the end of every short-term. It also keeps a record of all this attainment for each pupil throughout their primary school career.

#### 4.8 Chapter 4 Discussion: Embedding Standards and Standardisation through AwL

This thesis uses the enactment of the policy of AwL as a case study to explore what was happening in the wider paradigm of primary schooling between 2011 and 2019. As stated in the Research Design chapter, an aim of this thesis is to document difference and synthesising similarities and what is evident from the data presented above is that the six participant schools have enacted different AwL systems while attempting to ensure that national standards are complied with. Lewis et al. (2020) contend that discussions about 'raising standards' have essentially become a meta-policy for education systems worldwide. The term standard denotes an authoritative benchmark and level of achievement that guides and limits individuals, with standardisation originating in scientific activities and manufacturing (Timmermans and Epstein 2010). Landri (2022) contends that standards and standardisation are interdependent - standards rely on the process of standardisation, which in turn facilitates the dissemination of standards across various activities. Helpfully, they outline four types of standards which were visible in the data presented above:

- 1) Design Standards - specifications concerning tools and products;
- 2) Terminological Standards - a common classification;
- 3) Performance Standards - the threshold of outcomes and;
- 4) Procedural Standards which are the protocols to follow.

Fenwick (2010) noted that these standards frequently emerge from the interplay between normative guidelines, rules, and evolving configurations of knowledge and expertise within political contexts. According to Tsatsaroni and Koutsiouri (2022), standards function as instruments to simplify complexity and enable specific governance forms, and they are often presented as voluntary norms which can mitigate the perception of imposition and enable standards to move from being a norm to normal. Landri (2022:28) remarks that once standards become 'invisible,' they can perpetuate the inertia of a social environment, making them challenging to alter or even recognize. This perception of standards being normal and voluntary is compatible with the trend towards autonomy in education. The Organisation for Economic Co-operation and Development (OECD) has emphasised to its member states that there is a connection between school autonomy and student performance, 'political faith in autonomous schooling as key to a world-class education...' (Keddie 2019:6). The OECD proceeded to connect this autonomy to the devolution of power from central administration directly to schools combined with strong accountability mechanisms that evaluate certain standards. To determine if standards lead to standardisation, and whether standardisation leads to uniformity, it is important to explore how standards are adapted in different contexts, considering their varied outcomes and influential power (Riese et al. 2022). Table 10 summarises some of the main features of the six AwL systems and what is striking is how similar they are despite each headteacher having their own views on what good assessment looks like and each school having a different context. These differences influenced early decisions on AwL as schools SPT and PAT that were part of Multi-Academy Trusts (MATs) chose or were given electronic AwL systems so that there was standardisation across their MAT, whereas the other four schools developed their own AwL systems. However, the components of the AwL system are broadly similar.

Table 10: Similarities and Differences between AwL Systems

	Pupil Asset	Passports	Qualitative Approach	Points	School Pupil Tracker	Continuum
Developed by school	X	✓	✓	✓	X	✓
Given by Trust	✓	X	X	X	X	X
Bought by School	X	X	X	X	✓	X
Numerical System	✓	✓	X	✓	✓	✓
Progress judged against Curriculum Statements	✓	✓	✓	✓	✓	✓
To challenge quicker learners, work to deepen understanding of current topic is given	✓	✓	✓	✓	✓	✓
Uses Tests to moderate progress judgements	✓	✓	X	✓	✓	✓
Year group expectations	✓	✓	✓	✓	✓	✓
Progress based on number of Year Group KPIs attained	✓	✓	✓	✓	✓	✓
Uses Interim Standards in Y2&6	✓	✓	✓	✓	✓	✓
Inter-school moderation used to secure teacher judgements	✓	✓	✓	✓	✓	✓

The headteachers in all six schools, and teachers in SchoolPAT and SchoolQualitative conveyed that the basis of planning and assessment under AwL are curriculum statements. These statements can be long and contain several constructs and so there would be some variation in which constructs are being taught, however, all six schools identified curriculum statements that they considered to be essential for the pupil to master and labelled these statements as Key Performance Indicators (KPIs). Although not the subject of this thesis, a closer examination of these KPIs is likely to reveal that they are similar because all six schools utilised nationally standardised KPIs for Year 2 and Year 6 that are published by the Department of Education to benchmark their own KPIs against for the preceding Year groups. These KPIs for each Year group become the basis for end of Year Expectations that pupils are trying to meet so that they are prepared for the subsequent Year's curriculum. Figure 22 is an example of the Year 6 National KPIs for Reading.

Figure 22: The 2018 Year 6 KPIs for Reading published by the National Association of Head Teachers (NAHT)



**Year 6 reading**

Key performance indicator	Performance standard
<p>Applies a growing knowledge of root words, prefixes and suffixes (morphology and etymology) - as listed in English appendix 1 of the national curriculum document - both to read aloud and to understand the meaning of new words that are met</p> <p>Increases familiarity with a wide range of books, including myths, legends and traditional stories, modern fiction, fiction from our literary heritage, and books from other cultures and traditions</p> <p>Checks that the book makes sense to the reader, discussing the individual's understanding and exploring the meaning of words in context</p> <p>Summarises the main ideas drawn from more than one paragraph, identifying key details that support the main ideas</p> <p>Retrieves, records and presents information from non-fiction</p> <p>Participates in discussions about books that are read to the individual and those that can be read independently</p> <p>Provides reasoned justifications for their views about a book</p>	<p><b>With reference to the KPIs</b></p> <p>By the end of Y6, a child's reading should be fluent and effortless across all subjects, not just in English</p> <p>A child can:</p> <ul style="list-style-type: none"> <li>• discuss the purpose(s) of the language that is read and understand why sentences are constructed as they are;</li> <li>• focus on all the letters in a word so they do not, for example, read 'invitation' for 'imitation' simply because they may be more familiar with the first word;</li> <li>• accurately read individual words, which might be key to the meaning of a sentence or paragraph, to improve age appropriate comprehension;</li> <li>• read independently, including books they would not choose to read;</li> <li>• compare characters, consider different accounts of the same event and discuss viewpoints (both of authors and of fictional characters), within a text and across more than one text; and</li> <li>• reflect on feedback regarding the quality of their explanations and contributions to discussions.</li> </ul> <p>A child understands the majority of terms needed for discussing what they hear and read such as metaphor, simile, analogy, imagery, style and effect.</p> <p>A child applies the skills of information retrieval eg in reading history, geography and science textbooks, and in contexts where they are genuinely motivated to find out information, such as reading information leaflets before a gallery or museum visit or reading a theatre programme or review</p>

These Year 2 and Year 6 KPIs are also used in inter-school moderation meetings for those Year groups in all six schools and so they act as a benchmark for standardisation between schools too and Headteachers mentioned that they were aware the pupils leaving Year 1 and Year 5 needed to be prepared for the Interim Frameworks and thus the SATs seem to have a washback effect on AwL as it has been found to do with other aspects of schooling (Brown et al. 1996; Baird et al. 2017). This dominance of SATs within primary school life seems to also have a standardising effect on the AwL systems as pupil attainment on SATs is an unspoken validator of whether the AwL system was successful at tracking progress and attainment. Indeed all six participant Headteachers said that judgements made using their AwL system did tally with pupil attainment in SATs especially in the second year after some adjustments had been made to the AwL system following the previous year's results.

Five of the six schools use tests to moderate teacher judgements on pupil progress – a couple choose to purchase tests like PUMA or PIRA whereas others created their own. The change in the use of tests under AwL is that they are not viewed as the arbiter of pupil progress as teacher judgement and evidence from work that pupils have produced now have at-least an equal status when making pupil progress judgements, and nor is test data sufficient for this purpose. Teacher judgement has been enhanced by the focus on curriculum statements when planning lessons as the AwL system then tracks attainment according those. One of the reasons why teachers may have been less secure in their own judgements on pupil progress and consequently more dependent upon test data under levels, was because the criteria under levels were different to the curriculum statements being taught in the classroom. AwL seems to have achieved increased security in a teacher's own judgement while standardising through the focus upon teaching and assessing according to the curriculum. This focus upon the curriculum is also connected to the desire to embed mastery learning which all schools have committed to by pre-teaching and supplementary teaching topics to pupils who have been identified as needing extra support rather than allowing them to fall behind or streaming pupils into ability groups. Quicker learners in all schools are also not moving to the next topic under AwL but are being given activities to deepen their learning of the current topic. This focus on curriculum development across all subjects was increasing during Phase 2 of data collection as Ofsted's inspection framework of 2019 had an increased emphasis on this.

The exploration of AwL when it was first introduced aimed to discover how the new systems would be different and how these differences would change a schools system that had been standardised by levels – it seemed to be a radical move (Popat 2015). AwL seemed to be promoted by government due to its compatibility with a desire to give headteachers increased autonomy, however, the data presented above suggests an increase in standardisation that AwL has achieved. Previous policies such as the introduction of the National Curriculum in 1988 standardised the curriculum offer within schools, and National Literacy and Numeracy Hours in 1999 standardised some classroom practice overtly to the consternation of the teaching profession and scholars (Hargreaves 1994, Wang and Day 2002, Webb 2004, Day 2005). However, AwL seems to have achieved standardisation surreptitiously by focusing the profession upon the autonomy promoted by the policy. The results of Federici's (2013) survey study revealed a positive relationship between school headteacher self-efficacy and perceived job autonomy; and headteacher self-efficacy and perceived job autonomy were positively related to job satisfaction and negatively related to contextual constraints; and contextual constraints to autonomy were negatively related to job satisfaction (Federici 2013). The CAWL Report (2015) emphasised such headteacher autonomy while providing some scaffolding in the form of reflection questions. The first

question asked headteachers to consider what insights the assessment should provide to discern the security of the students' knowledge and understanding of the subject matter, concept, or skill and whether they can apply that with a growing independence. To ensure inclusivity for pupils with Special Educational Needs and Disabilities (SEND), thought should be given on how they can demonstrate their learning. Secondly, headteachers were asked to reflect upon how they would communicate the purpose of an assessment and also the assessment data with pupils so that they know their next steps in learning. Thirdly, headteachers were asked to encourage teachers to reflect on how they could adjust teaching to improve pupil understanding and to consider whether some pupils need extra teaching or practice. Fourthly, headteachers were asked to consider how to support progression where learning is secure in quicker learners as they should not be taught outside of the topic being taught to the rest of the class. This might involve assessing whether pupils can apply the concept in different contexts or to explore a concept in greater depth before moving on to new material. Fifth, headteachers were asked to consider how to record and depict progress. None of the headteachers in this study have resisted implementing AwL according to these principles even if they resisted numerical reporting and classroom testing as in the case of Qualitative or an attempt to resist Year-group expectations as in the case of Continuum.

Practices such as formative assessment, adapting teaching approaches, and conveying progress and next steps of learning to pupils and parents, are accepted parts of schooling and so headteachers autonomy to remove them was never an option and it can be viewed that the reflection questions provided by the CAWL Report (2015) emphasise that. What AwL provided was latitude to enact these teaching functions according to a headteachers values and to meet the needs of the stakeholders in that school context which seems to have worked by providing sufficient autonomy to enable headteacher satisfaction while maintaining a large degree of standardisation between schools. National accountability testing in Years 2 and 6 in primary schools are implemented to ensure uniformity in evaluating student performance across the nation (CAWL 2015) carry high stakes for all stakeholders in a school as they are used to compile league tables and inform inspections by Ofsted. This has prompted all six schools in this study to use Year 2 and Year 6 National Test Standards to benchmark the standards in their AwL system and also to benchmark with partner schools. This approach is in-line with observations by Braun and Maguire (2018) under the levels system that despite disapproval of these extensive testing regimes, active opposition was unusual as school leaders and teachers embody pragmatic compliance (Keddie 2018). Ball (2003) noted that these accountability mechanisms prompt leaders and teachers to organise themselves in response to inspections and test performance. Bevan

and Hood (2006) observed that across public services in England, accountability process combined targets with 'an element of terror' (Bevan and Hood 2006:517); and Braun and Maguire (2018) explained that teachers had a defence mechanism against a fear of failure by attempting to anticipate what inspectors may want to see.

## Chapter 5: AwL Internal Moderation as the enactment of an Audit Culture

### 5.1 Chapter 5 Introduction

The CAWL (2015) Report acknowledged that eliminating levels would place greater emphasis on moderation. Chapter 4 explores moderation between schools, while this chapter examines moderation within a school under AwL. CAWL (2015) suggested that schools consider how to ensure the quality of assessment data, the time teachers need to record it, and the appropriate frequency for collecting and reporting it. They emphasized the goal of obtaining precise data on pupil understanding and ensuring unbiased teacher interpretations. Although detailed guidance was not provided, the National College for Teaching and Leadership (2014) conducted research with 34 Teaching Schools to explore alternative assessment methods beyond levels. They highlighted an example where a school developed software for lesson planning and identifying next learning steps. Teachers used this software to capture evidence of student achievement against national curriculum objectives, creating electronic profiles that showed the child's achievement standards relative to school and national expectations. Teachers could comment on the evidence, allowing stakeholders to see not only the child's achievement level but also the process, skills acquired, and content covered. These e-Profiles were connected to the main school server, enabling both internal and external moderation through adjustments, edits, sharing, and publishing. In this chapter, the process for achieving this in the participant schools is described using class teacher and headteacher interview data. What transpires is that this process has raised the standard required to justify teacher judgements prior to data being entered into the school's system.

### 5.2 Moderated Input in SchoolPAT

Head-PAT explained that the system tracks according to a Year-group expectation with progress being measured each half-term and they triangulate judgements using different types of data,

'we will do a summative test, we'll do standardised tests, we'll do teacher knowledge, we'll do looking at books and weighing-up the evidence that we see in books with national curriculum objectives at the end of the Year - so it is a combination of different factors' (Head-PAT).

SchoolPAT use summative assessments three times per Year through a program called Pixel that provides a gap analysis in pupil understanding and then the teachers use the analysis tool there to then plan the next sequence of lessons, so they can target these gaps. standardised tests at the end of each school Year with a focus pupils going into Year 2 and Year 6 when they will sit national tests, and also for pupils with special needs at several points during the Year especially when they are applying for extra funding and support for them. Performance on these tests is then used to identify gaps in understanding and targets are decided in nationally test subjects, numeracy, oracy, reading and writing, and an action plan for learning for students that may be missing previous targets is created,

'The action plan changes every term because children are moving and learning - it does not happen in a predictable pattern... if that child is not meeting their personalised end of Year target, then we have got to look at that child whether they are targeted for Greater Depth or even if they are targeted to not quite get to Expected, either way, if you know that they are not going to meet that target then we look at that' (Head-PAT).

A first draft of that target is created at the end of the School Year during a meeting when pupil books are handed over to next Year's teacher and that is reviewed by that teacher and the Senior Leadership Team (SLT) in September/October time after the teacher has had some time to get to know the pupils. SPAT-Y4T explained that they meet regularly with the Senior Leadership Team (SLT) to discuss the progress of every pupil in their class and in those meetings teachers are themselves challenged to adjust their teaching to help the pupil,

'There is a lot of rigour in those meetings and you're challenged, that middle group of children of yours are they challenged, which of these can you get to Greater Depth. Which of these are at risk of not achieving the end of Year expectations? The group Below, which of these children can you get to the middle rank. So I think you are less likely to not get them on and see people coasting, because there is always a push to see children converting from At (Expectations) to Above, so it does not seem quite as much of a risk for me as it did in the past' (SPAT-Y4T).

This internal moderation process is designed to increase the security of a teacher's judgement and establish trust in `those judgements among their colleagues and it forces teachers to accumulate evidence from a pupil's work throughout the Year, 'we hold quite a high evidence-base in their books so the work that they are doing every single day is how I

collect my evidence to show that they are making progress' (SPAT-Y4T). This results in teachers collating a portfolio of evidence of progress for a pupil which in-turn results in teachers knowing pupils very well,

'We know our children really well, so for example, there are 2 children in my class that are targeted to get age-related Expectations by the end of the Year and they're yellow at the moment. And I have done case studies of those to be able to explain to Head-PAT as to why they're yellow at the moment but I have also done QLAs (Question Level Analysis) of their work and their assessments to know which bits I need to plan to be able to shift them' (SPAT-Y4T).

The headteacher and class teachers explained that under AwL, judgements on pupil progress are made based on the triangulation of data from multiple sources of pupil workbooks and test performances to ensure that teachers are secure in their judgements and that all colleagues have trust in that as under levels some teachers reported that judgements did not reflect the reality of a pupil's performance,

'we did-away with teachers just inputting data on their own. It's definitely joined with leaders and with Senco, saying, okay, you are saying that this little girl is At Expected for Writing, can you show me her book, get her book out and let us have a look and see' (Head-PAT).

SPAT-Y4T shared that these moderated input meetings usually take 3 hours with pupil progress in each subject being analysed systematically with particular focus on 'those that are on the cusps - what has been done to support them... it is committed to support these children from crossing the boundary into Expected or crossing from that boundary into Greater Depth' (SPAT-Y4T). They stated that they found that the atmosphere was very supportive to help them to help the pupil but there was an onus on justifying judgements with data so they will take examples of work from those that are Above Expectations, those that are At Expectation, and those Below Expectations, 'they will often ask us to take, not necessarily all of the work, but maybe a cross-section, so two tops, two middles, two bottoms' (SPAT-Y4T). They explained that the context of the school has also influenced this approach of moderated input. As explained in Chapter 4 Section A, SchoolPAT was placed in special measures due to poor performance and Head-PAT chose to take-on the challenge of leading it as her second school in 2016. They had not felt able to rely on data provided by the previous leadership team and so used testing of pupils extensively in the early years to aid with benchmarking,

'I've heard along the grape vine that previously the tests were really important, because two or three years ago, I think there was lots of issues obviously before Head-PAT came. So, I think that the senior leadership weren't very familiar at the start and they needed quite clear data to support at the start. I think now it is carried on and there are less tests this Year that we have been asked to do. So, we were asked to do one at the beginning of the first term and we did some just after Christmas as well and we will do some at the end of the Year as well, just to keep an eye on progress from that point of view in terms of the test scores' (SPAT-Y4T).

During Phase 2 data collection in 2019, Head-PAT explained they were in the early stages of encouraging teachers to create their own tests initially in mathematics so that there is better alignment between what has been taught and what is being assessed which they felt would add security to the teacher's judgement,

'we are moving away from buying-in tests and the expectation... we're starting with maths this term, so if you're teaching fractions, next term, you need to write your own assessment on what you're teaching, because teachers should deliberately know, I am planning to... and I am going to know that they have done it, is when they have answered the questions that I have set them... So, once we have got a whole suite of assessments we'll know that those assessments are planned deliberately to match what we think needs to be taught' (Head-PAT).

They explained that all teachers have been provided resources to help them write these tests both for expected learning and to push pupils to greater depth in a topic. These teacher-developed tests will then be treated as another piece of data for moderated input,

'Then, that is extra evidence to say that the child can do it, and all of those summative, quick snapshot, summative tests will be QA'd by the maths lead who is also a national mastery lead' (Head-PAT).

Much has been written on the validity of test data in determining what a pupil knows and does not know. In SchoolQualitative, it became an issue of conflict between staff and School Governors and more is detailed on this in Chapter 5. However, in SchoolPAT, the SLT has taken the view that no type of data should be relied upon completely and each must be

triangulated including numerical test results. Head-PAT explained that they took a pragmatic approach because much of the school system required numerical data for accountability purposes,

'schools are numbers-driven and you will never get away from that, and I think any school that says that they are not numbers driven is fooling themselves, because at the end of the day, there is a national standard, there is a floor standard, and if you fall below the floor standard, a flag is raised and you are going to have pressure on you anyway. And anyone who is coming in to your school or whatever, externally, they are all going to want numbers. So, you can never ever escape from the numbers game; it is always going to be there' (Head-PAT).

Rather than reject numbers completely or regard them to an oppressive tool, SchoolPAT have tried to ensure that the numbers that they report are accurate, 'So, I think you have got to find an authentic way that you can maintain integrity as a school, and still know that I am providing numbers, but the numbers that I am providing you are accurate 100%' (Head-PAT). This commitment to integrity is a central value of AwL and this is discussed in more detail in Chapter 5.

### 5.3 Moderated Input in SchoolQualitative

Head-Qualitative describes moderation as 'a key thing' as they have in-school moderation of every single writing book before a pupil moves to the next year group prior to making a final judgement of pupil's attainment and progress. These meetings involve the pupils, their teacher one member of the senior leadership team, and the next Year's class teachers to ensure that there is a continuity of understanding. They also undertake inter-school moderation with schools in their partnership, as-well-as organising an annual moderation with schools from several different partnerships where teachers are split into different Year groups and the approach of comparative judgement was employed to try and ensure consistency of understanding. The moderation process is different for Year 2 and Year 6 pupils because these years have the Interim Framework provided by the DfE and in the past two years these Years were moderated externally by the local authority and Head-Qualitative was pleased that the moderator wrote to her governing body to say how impressed they were with their assessment system. In the two summers that that school had sat SATs, the attainment scores assessed by teachers tallied with SATs results especially in reading and grammar. In writing, an external moderator had felt that their judgements were

too harsh and so they have now adjusted them to be more generous and they now also tally with SATs results.

This collective problem solving extends to moderation within the school as teachers discuss the work of a pupil with a colleague or as a group of colleagues, 'We come together really... So, I would take a writing book to another teacher - to the Year 5 or Year 6 teacher - and we would look together at that piece of writing to check that we all were happy with where it was and the progress that they have made' (SQ-Y4T). As pupils are not tested and progress is not measured numerically, targets for attainment are not set but instead an individualised vision of progress is created collectively for each pupil.

'we are looking at the progress they're making between pieces of work but within that we have an understanding about age-related - what our expectations would be... We're getting a better shared understanding of what they should be looking like in our experience, but we're still questioning, are they moving enough, so even if someone is looking about right between bits of work, if it's not looking like they're moving, why are they not moving? If you just look for numbers you could easily say, they're fine - that's two completely different conversations' (SQ-Y6T).

SQ-Y5T said that they felt, 'Much more secure in my judgement' (SQ-Y5T) through this system and felt that their judgements were challenged by SLT just as much as under the system with levels when they were reporting numerically. The difference under this AWL system is that rather than examining numerical data on a spreadsheet, teachers gather and discuss examples of work produced by the pupil and as the context for that is known best by the class teacher, SQ-Y5T felt that their views had more importance to SLT,

'I feel now that my judgement now is more valued, and the discussions about them with senior leadership - they take it on board. I mean, they look at the book as a snapshot, whereas I have worked with that child for a long time' (SQ-Y5T).

When I asked SQ-Y5T whether verifying their judgement with test data would make it even more secure they did not feel so but reflected that it could possibly help the pupil develop test experience prior to entering secondary school.

## 5.4 Moderated Input in SchoolPassports

Head-Passports explained that due to the new National Curriculum being more challenging, teachers had to improve their subject knowledge and also their understanding of what the new expectations were. Several staff meetings focused upon looking at what was to be classified as Emerging and Expected, through looking at exemplar materials, pupils' work, and moderated work samples both within the school and also with neighbouring schools. As the school is not part of a Multi-Academy Trust (MAT), they have formed a partnership with local schools and have created a development plan in-which Head-Passports leads on the organisation of termly moderation meetings. In 2017, Year 2 pupils work that was moderated was found to have 100% alignment with the judgements of teachers from a neighbouring school,

'we are a one form entry school so we also make a point of linking with our other partner schools to conduct moderation to see if we were linking - to see if we were levelling. So is (our) year-related expectations in Year 1 looking like (another) school's age-related expectations in Year 1?; so we did a little bit of cross matching. If you've got single year groups you need to do that otherwise you could be sailing alone completely off on a tangent. So that was important' (Head-Passports).

However, as teachers were still unsure about their judgements following the change of system, Head-Passports purchased PIRA and PUMA tests from Hodder Education<sup>12</sup> for reading and mathematics which pupils took termly to provide an objective measure of pupil attainment that could support teacher judgement and stimulate gap analysis,

'teachers professional judgement was the fundamental thing that I wanted to get right... What we were saying was "we have a score here of 112, we have a workbook here with this amount of evidence and we have a Passports with this number of texts where the child is making progress... the emphasis in my school is on the teacher's professional knowledge and understanding of that child, based upon the body of evidence and some moderation' (Head-Passports).

Pupils and teachers have progress meetings in which teachers will classify pupils as Emerging or Emerging+; Developing or Developing+; Expected or Expected+; Greater Depth

---

<sup>12</sup> <https://www.hoddereducation.com/assessment>

or Greater Depth+. Head-Passports was keen to encourage teachers to query the quality of their data rather than just use it as a grade,

‘who isn’t Developing, why haven’t they made that step that they should have?; who is flying and what are we doing for them? And if this child is struggling “do we know why that’s happened?; what’s changed?; what interventions and measures were there?; what are we doing different to help their child develop?’ (Head-Passports).

Head-Passports held pupil progress meetings each term with teachers and the pupil, and then phase leaders or key stage leaders have conversations about the data in their phase and they talk about the outcomes and why that outcome was reached. Teachers have to justify their judgements and look at work and decide whether they are being accurate. Pupils who are underachieving are discussed and plans for them are created. Thus, the data is there to ask questions. The school also uses Oxfordshire Pupil Tracker<sup>13</sup> which enables headteachers to generate whole school data for their own analysis and to present to governors and Ofsted.

## 5.5 Moderated Input in SchoolPoints

Head-Points was concerned about the absence of a common national assessment language under AwL and the misperceptions that that could cause about a pupil’s progress and attainment,

‘Some people talk about ‘Working At’ and perhaps some talk about ‘Working Above’, some talk about ‘Mastery’; with some of the tracking systems, when you go into a school in September they will tell you well 80% of the children are ‘Working At Expected’ because they’re ‘Working At Working Towards’ which is where they’re expected to be at that time of year; whereas we would be talking about ‘Working Towards’ until they meet the ‘Expected’ level’ (Head-Points).

To improve certainty, Head-Points believed that schools like their one are using standardised testing more than they used to with products such as PIRA and PUMA by

---

13

<https://static1.squarespace.com/static/5da58c1361113f4a0a4fb658/t/5df8a989ec20c045b7c6b976/1576577417818/Pupil-tracking-procedure-2017-21-1.pdf>

Hodder Education being a big influence in schools under AwL so that teachers have more security in their judgements of pupil progress. These uncertainties have increased the importance of moderation meetings with schools in the partnership to ensure consistency of expectations of pupils. This is a local authority school and that is part of a partnership of eight schools that has existed historically based around a secondary school. Some of those schools are part of academy trust but as this partnership existed prior to AwL and mass academisation, the headteachers have decided to keep it and they meet every term. Teachers in Year 2 and Year 6 meet twice per year and bring different pieces of writing and maths so that they can compare, and check that they're meeting the same criteria. As this is a three-form entry school, such moderation meetings are also held within the school b some of the statements are subjective. As Head-Points explained,

'you had one class that got only 36% had made 'Expected' and another class had got 76% - so what is it that's making a difference because the children aren't that vastly different. And it's the interpretation of those statements and so we are now having to spend a lot of time going through the statements asking "what does that mean?; what does it look like in a child's work?; how many times do they have to show it?" ... and then there's the amount of input that pupils have in an unaided piece of work - what does that mean? Does it mean that they're completely unaided which is what Secure Fit was meant to be. Or, actually have they had quite a lot of input through dictionaries, thesaurus... to have everyone working at a level playing field is a can of worms' (Head-Points).

This approach exemplifies the recommendation in the National College of Teaching and Learning (2014) Report that evidence from pupil work should be given the teacher's professional interpretation focusing on both the progress made and the process undertaken, including the skills acquired and the content covered.

## 5.6 Moderated Input in SchoolContinuum

HeadContinuum explained that the Senior Leadership Team (SLT) can use the AwL system to have conversations with teachers about the progress of individual pupils or the class,

'what happens is if the teachers haven't taught or if they don't assess it, so sometimes they'll have poor progress measures because they haven't looked at two whole strands because they haven't taught it. So, we can say "right what you need to do in the next two weeks is to teach shape and measure; you need to

look where children are at and you need to plan effectively to make progress on that next time and also look at statistics”, or something like that’ (Head-Continuum).

Therefore, although the desire was to measure pupils across a Key Stage rather than against Year-group expectations, teachers would still be held accountable every half term as the senior teachers would be measuring pupil progress against the curriculum statements that were expected to be taught in a particular term.

## 5.7 Chapter 5 Discussion: AwL Moderation as the Enactment of an Audit Culture

The CAWL Report (2015) did not provide guidance on moderation other than for headteachers to consider when it might be necessary to help ensure consistency and accuracy of assessment judgements. Each school in this study created a moderation strategy tailored to their specific needs, guided and supported by the headteacher's vision. Teachers acknowledged the importance of leadership in moderation but also felt a shared responsibility for it. Both the teaching staff and the headteacher believed that moderation had enhanced teamwork within the school, with headteachers noting it had promoted a common understanding among everyone. The schools in this study have adopted an approach known as social moderation which seeks to ensure consistent teacher judgement through collaborative efforts (Adams and Anderson 2019). It has been found that using moderation as a sociocultural learning experience improves teacher judgement (Klenowski and Wyatt-Smith 2013). A social-cultural approach aims to make moderation a form of a community of practice whereby teachers learn from each-other and by participating in the process to develop their assessment judgement so that integrity improves (Wenger 1998) and Smaill (2013) went further by advocated for professional learning to be the primary aim of social moderation. Scholars such as Maxwell (2002) argue that schools often face a dichotomy between using moderation for accountability purposes or improvement purposes and can struggle to achieve both, but as AwL judgements are not used for accountability purposes, this can be achieved more easily. Under AwL, moderation has three purposes: to ensure valid, reliable, and accurate teacher judgements; serve as professional learning to improve teaching, learning, and assessment practices; and to provide 'quality assurance' or 'quality control' of learning, teaching, and assessment by aligning the judgements of different teachers within and between schools (Harlen 2007:20–21). This latter purpose is being achieved by using National SATs Frameworks in Years 2 and 6 combined with discussion on examples of work as each school has different AwL systems unless they are in the same

Multi-Academy Trust (MAT). Effective relationships between the school and the MAT, the school and other schools, and among teachers within the school, were all crucial in developing an understanding of moderation and implementing it as a process (Adams and Anderson 2019).

Under AwL, inter-school moderation is still based on externally provided criteria and judgements are centred around whether a piece of work has met that criteria in the opinion of teachers from more than one school. Usually, teachers from different schools will take turns to lead those moderation activities. AwL moderation practices seek to address the other two purposes by combining a focus on National Curriculum standards together with socially created meanings. Engaging in moderation against externally provided standards aids the clarity of comparisons but it often overlooks the social constructions of standards (Adams and Anderson 2019). AwL moderation within one school was centred around standards in the form of learning objectives and success criteria set for the Year group. Together with these broader objectives, the class teacher, headteacher, and other senior teachers set objectives unique to an individual pupil and measure their progress from a starting point and what the student was capable of in the previous meeting compared to where they are now. AwL treats interpretations to determine attainment and interpretations to determine progress differently – the former uses externally mandated standards for judgement whereas the latter defines progress as subjective and personalised to a pupil. In the context of Scottish primary schools using teacher judgements for summative judgements and accountability purposes, Adams and Anderson (2019) argued that ‘the absolutist version is concerned with proving, whilst the progression-based version is concerned with improving’ (Adams and Anderson 2019:11). By absolutist they mean making judgements on pupil attainment for accountability purposes and so these judgements are high stakes for multiple stakeholders and so proving them is needed to generate trust. In the context of AwL what this thesis reveals is that primary schools require progress judgements for in-school tracking to also be proven. Teachers in Schools Qualitative and PAT explained that they held regular meetings between the class teacher and members of the Senior Leadership Team that required them to take examples of pupils work to evidence the judgement on progress that they had made. A large part of these discussions focused on individual children's progress relative to their prior achievements based on the experiences of teachers. Teachers presented narrative accounts of selected pupils learning journeys, with questions from the SLT eliciting more information. SATs criteria were only referred to for Year 2 and Year 6 pupils otherwise these moderation meetings were framed by individual learner objectives and success criteria. Such narratives of learner progression require planning with the headteacher asking for data and work examples of certain specific pupils while the class

teacher also needs to be prepared to show to the SLT that they know the progress and needs of every pupil in their class and that they have a plan to support each one in the way that they need. Klenowski and Wyatt-Smith (2013) note that moderation can change classroom practices and AwL moderation is seen as part of the planning, teaching and assessment cycle.

Allal (2013) posits that achieving transparency and reliability in teacher evaluations, alongside the dependability and validity of external assessments, are complementary objectives within an assessment framework. Tolo et al. (2020) explain that school leaders must strike a balance between trust and control, as they are responsible for ensuring that tasks are performed according to expectations. These debates around trust and accountability during the enactment of AwL were framed by arguments initiated in O'Neill's 2002 Reith Lectures and clarified in her 2013 article on the concept of intelligent accountability. O'Neill (2013) argued that to achieve 'intelligent accountability' in our assessment system, we need to ensure that trust can be placed or withdrawn based on clear and comprehensible information. She explained that this involved enabling stakeholders to independently evaluate the performance of pupils, schools, and the education system and she felt that current complex metrics used to compare student performance did not present the necessary information for making informed trust decisions (O'Neill 2013). What O'Neill (2013) overlooked was that teachers themselves did not fully comprehend what pupils could do either. Braun and Maguire (2018) reported that teachers in two primary schools in London had a sense of deep unease with assessment tasks and were engaging in them without believing in them. Indeed, the Expert Panel on Curriculum Review Report (2011) stated that the main rationale behind the removal of levels was that they obscured the strengths and weakness in pupils understanding as they were best fit criteria and how teachers had lost trust in the levels being reported by the previous Year-group teacher due to a mismatch between that and the results of a baseline test. Headteachers in this study have described levels as being 'made up' and a teacher has detailed the pressure that a headteacher had put them under to allocate levels to pupils that they did not deserve. Therefore, in-school moderation under AwL firstly seeks to ensure trust between teachers. and between teachers and the school's AwL system that they are using, followed by encouraging trust between all stakeholders. During the five years prior to AwL, the government in Scotland had introduced summative assessments based on teacher judgements (Adams and Anderson 2019) and required schools to sample and review evidence of learners' work; agree on strengths and next steps; and provide feedback to improve practices (Scottish Government 2010: 3). Interestingly, Adams and Anderson (2019) reported that moderation in primary school was being conducted in multi-age and multi-stage groups where children were not judged against

a Year-group expectation much like that proposed in SchoolContinuum. Although reference to Scotland was not explicitly made by any participant in this DPhil study, that approach to summative judgements is very similar to those adopted by schools under AwL and this is indicative that teachers in both countries were contemplating notions of restoring trust in teachers by trusting teacher's assessment judgements while requiring a higher evidence threshold from these teachers during that period.

Tolo et al. (2020) introduced a concept they call internal institutionalized active inquiry, which they say involves making a teacher's work visible for documentation, scrutiny, evaluation, and enhancement. They argue that this approach serves as internal quality assurance, fostering questions and internal disagreements as crucial for quality improvement. These scholars did not expand on this concept further or illustrate it although they did give several examples of school stakeholders in Norway working actively to enable successful implementation of AfL. The data from this thesis is helpful in illustrating what this concept may mean. For example, Head-Passports purchased external tests for reading and mathematics which pupils took termly to provide an objective measure of pupil attainment to aid identification of any gaps in understanding. These test results were combined with the teacher's experience of the pupil and evidence from the pupil's workbook combined with the pupil's AwL Passport that had been tracking progress against curriculum statements. Pupils and teachers then held progress meetings in which teachers classified pupils according to the AwL measure such as Emerging or Emerging+ and pupils were made aware of their next steps of learning. The Expert Panel Report (2011), CAWL Report (2015) and headteachers in this DPhil study emphasised that an aim behind the enactment of AwL was to restore autonomy to schools and increase a teachers sense that they were a trusted professional. However, being a professional trust after the enactment of AwL refers to the triangulation of a teacher's judgement with evidence from tests, workbooks and senior teachers perspectives prior to inputting progress data into the school's AwL system. This reminds one of the debates in the scholarship during the 1990s and 2000s around the public service's adoption of New Public Management approaches that require management and measurement (Hood 1991). Apple (2004) argued that such approaches require the constant production of evidence that tasks are being performed both efficiently and correctly, a concept he referred to as an audit culture which seems to be the approach adopted in AwL moderation. Rather than trusting teachers, the question that arises is whether trust has now been fully replaced by control systems (Becker 1996); can a teacher have autonomy without being autonomous?

Groundwater-Smith and Sachs (2002) argued that an audit culture requires professionals to relegate their personal judgement and self-regulate by adhering to externally imposed directions, and Claus Offe (1999) argued that, 'I do not need to monitor those whom I can trust,, nor do I have to force them to do what I expect them to do or to call in third parties (such as courts) to enforce my claims... Trusting begins where the reach of these media of control ends, or where we decide to stop calculating, enforcing, and monitoring - if only on the basis of an assessment based upon prior monitoring and the knowledge derived from it. ' (Offe 1999:52–53). Thus, Offe raises the question of how many times a professional's judgements be assessed before it is unnecessary to control them? Bailey (1997) stressed the importance of professionals using their discretionary powers as autonomous practitioners, rather than simply following the orders of headteachers, inspectors, or governments. The latter process indicates a lack of trust in professionals' ability to perform their duties properly, as Frowe (2005) comments, 'one cannot claim to trust professionals while imposing extensive regulations to dictate their behaviour. This approach is self-defeating, as any notion of trust that requires regulations to enforce it inherently undermines the trust it aims to foster' (Frowe 2005:51). The problem for headteachers is that trust always involves some risk and educational attainment is high stakes for all school stakeholders (Tolo et al. 2020). This has prompted a redefinition of professional trust shaped by the audit culture.

## **Chapter 6: Findings and Discussion on the Enactment of Assessment for Learning through the Enactment of Assessment without Levels**

### 6.1 Chapter Introduction

This chapter presents data generated from interviews and group discussions conducted over a two-day period in SchoolPAT and during a two-day period in SchoolQualitative in March 2019, which was approximately 15 months after the Phase 1 data collection that was presented in the previous two chapters. This approach adopted in this thesis is that both enactment and context are fluid and continuously evolving, part of an ongoing process (Maguire et al. 2015), and by exploring the enactment of AwL at two separate points in time insights are provided into how schools were engaged in actualising the ambitions of AwL. Choosing the enactment of AwL as the case study allows for the complexity of this process to be elucidated (Pal 2005), and this chapter presents data from school stakeholders within their everyday settings, illustrating the intricate dynamics of how social interactions shape beliefs and decisions (Feagin et al. 1991). Data is presented from a combination of individual Multiple Perspective Interviews with the headteacher and class teachers within the same context (Vogl et al. 2019), together with group discussions among homogeneous groups of pupils, parents, and governors. While common themes and topics are used to synthesise the data, and attempt has been made to ensure the locatable voice of these different stakeholders is preserved and by doing that it is amplified. The result is a chapter that provides rich details and facilitates triangulation on the understandings among stakeholders on how the enactment Assessment for Learning (AfL) was enacted through the enactment of AwL and how teacher professional judgement has to satisfy a higher evidence threshold before being trusted.

As detailed in the Literature Review, research indicates that Assessment for Learning (AfL) should be aligned with curriculum standards and integrated into classroom teaching. Teachers must combine AfL with their subject knowledge to meet students' learning needs and provide effective feedback that enhances understanding and clears up common misconceptions. Teachers need strong pedagogical skills and the ability to lead discussions as eliciting students' thoughts and facilitating discussions are vital to help teachers adjust their instruction and provide immediate feedback. Effective feedback is central to AfL as it stimulates student engagement and self-regulated learning. While passive students often struggle to use feedback effectively, active participation and increased student autonomy in learning and assessment are associated with better outcomes, higher engagement, and fostering a sense of responsibility for their learning (Heintink et al. 2016). The data presented

below conveys how these indicators of successful implementation of AfL were being practised by teachers and pupils in both Schools PAT and Qualitative.

The second theme emerging from the data highlights the efforts schools undertook to ensure assessment judgments were consistent and accurate. Each school crafted a moderation strategy tailored to their unique needs, guided by the headteacher's vision. The schools implemented a social moderation approach, aiming for uniform teacher judgments through collaborative efforts. Across schools, the National SATs Frameworks for Years 2 and 6, along with discussions on work samples, were utilized, as each school had distinct AwL systems unless they were part of the same Multi-Academy Trust (MAT). Strong relationships between the school and the MAT, other schools, and among teachers within the school were essential for understanding and executing moderation. Within a school, leaders cross-check teacher judgments using various data sources, such as tests, teacher insights, and evidence from student workbooks resulting in a more robust evidence requirement prior to those judgements being trusted.

## 6.2 Phase 2 Data from SchoolPAT

### 6.2.1 Ongoing Classroom Dialogue

Continuous dialogue and verbal feedback was a feature of SchoolPAT's AwL process. Head-PAT explained that, 'We do a very much more fluid approach, where teachers are... we say it is like watching a game of ping-pong – it is very fast, it is very quick-paced and the teachers are continually showing something, and then the children have a go, and then they are back, and then they are having a go, and they are coming back' (Head-PAT). Teachers are expected to focus on one small objective every day with the whole class so that they have a very clear understanding of who the children that have fallen short of their objective are. Every pupil, even those with SEND, are taught the objective for that Year group, 'we do not teach-up and we do not teach-below. If a child is struggling, then we scaffold-up to reach that objective, but we do not teach out of the Year group objectives' (Head-PAT). The teacher is expected to use a combination of verbal and written feedback, 'sometimes I make a sort of a choice to do written or verbal feedback to different children perhaps depending on the type of child or depending on the lesson' (SPAT-Y6T). Some of the Year 6 pupils explained, 'Feedback in homework works like... so we'll go home and do our homework and then they will mark it and they will tick it in green pen and if we have got it wrong... they'll probably call us up and ask us if we could try it again' (SPAT-Y6pups). They explained that

their teacher usually consolidates the written feedback with verbal feedback if they have some wrong answers which they perceive has two benefits:

- 1) The verbal feedback clarifies the meaning behind the written comments.
- 2) The written feedback can be referred back to and act as a reminder when producing future work.

'sometimes you do not really understand it (the teacher's written feedback) and you then have to put your hand up and then they'll come over and read it out to you if you do not really understand... and they will probably work with you or help you' (SPAT-Y6pup).

'I would say [verbal feedback] with the green pen as well, because if, for example, they said, one time, SPAT-Y6T said, I cannot help you on this, because if they do not have green pen then we would not know directly what to do, we'd just repeat the same thing that we have done again. So, SPAT-Y6T writes in green pen and explains some of it, like what you have to do and it helps us remember it' (SPAT-Y6pup).

In SchoolPAT, this one-to-one work with the teacher or teaching assistant can be called a workshop and the written feedback guides the pupil as to how they are expected to engage in the classroom. The written feedback is also part of the dialogue and pupils are encouraged to write responses to teacher feedback.

'The feedback group is like, in English we have, if you get GWT it stands for guided work with the teacher. GWAT stands for guided work with learning assistants, and IW stands for independent work. And a tick and a T is, learning objectives met, and the challenge task set, and a T is learning objective not yet met, additional task set. VF is verbal feed-back comments. Yellow highlighter is check your spelling, green highlighter is you have done this part really well, and TF is you write your feedback... It helps quite a lot and I think it really is quite clear, and once you get used to it really is quite good... In English, when I have written some stories, I would get a yellow highlighter and then I would have to get a dictionary, and find out, and now I have done that, I actually know how to spell that word properly' (SPAT-Y6pups).

Head-PAT explained that they like to see teachers engage in written dialogue with pupils in their workbook as they feel that it is a gesture of respect for the pupil,

'for me personally, I like to see that there is some sort of indication of that in books with the teacher's response to the children marking, and if the child is reflecting and taking the time to reflect and think about it, then I think the least that we can do is respond to that thought. Because if you were having a verbal conversation with the child, and the child told you something, you would not just ignore that and walk away, you would respond. And so, for us, it is about showing that it is valued' (Head-PAT).

There are also opportunities for whole class problem solving and marking, 'Sometimes it can be like, SPAT-Y4T sets us a challenge on the board. Like in addition or a subtraction, like a decimal trick... and then we like, we have to write in our books, and then we transfer it. And then once we finish it, we check it and see if it is right' (SPAT-Y4pup). The dialogue also continues at home particularly where some pupils need some additional support,

'because my son has so many issues it is called a home/school diary, so if there is an issue at home, at school, we can communicate that way. Which is one way of putting it down because he will not always talk to somebody at school and so from that aspect it worked really well. So, it's a good two-way thing which you can write in as well. So, you have got your children writing their response to the teacher's feedback, but then you could also write something if you done some work. And then we can say how we're really interested in doing that and so we can make that comment in the book and then you are sort of really communicating at school and it helps them as much as it does us' (SPAT-Parent).

This continuous dialogue between teachers, pupils and parents seems to have created a community of learning where parents respect the efforts of teachers and also have self-respect in knowing that their efforts outside of school are valued by teachers.

### 6.2.2 Embedding Mastery

Much of this whole class teaching and feedback approach comes from the desire to embed a mastery approach into English primary schools and move-away from streaming pupils as

Head-PAT explains, Whereas, if you think, 5 or 10 Years ago, you can have 6 different groups and each different group have got a different set of activities that the teachers spent all night before preparing 6 different situations, 6 different ways. SPAT-Y4T commented that 'Our kind of mantra is that it is weighed upon quality first so what is best for an individual should be best for the whole class'. Head-PAT argued that that expectation was unreliable as it was not based upon evidence of the pupil's performance as some pupils would not attempt that task if they were in a lower stream, 'it is impossible for the teacher to know the 30 kids who manage to achieve that objective, because the teacher has pre-decided that you can achieve only this. SPAT-Y4T explained that once an expectation was set by the teacher then pupils tended to perform according to pre-determined expectations, 'Yeah and the children sort of fall into that because that is your fixed perspective - I think that is quite possible'. Contrastingly, under AwL, they felt that the dynamic had switched, 'So now you are in reaction to the children on the spot a lot more, I think, since the mastery approach'. They proceeded to explain that the switch was because teachers and teaching assistants would be reviewing all pupils work in the classroom and provide feedback to whoever needs it rather than targeting the same pupils in every topic, 'it is more sort of, on the spot differentiation rather than planning in-advance – it is more responsive this way because you could plan in-advance what you think the children might be doing but it does not normally work out that way'. Head-PAT explained that teachers will make judgements based on evidence from pupil work to ensure that their 'concrete and pictorial representations are supporting their representations of the concept' (Head-PAT). The tone of the feedback has also changed to pushing the pupil to think deeply and produce more rather than judging right and wrong, 'So that might be challenging some children and having a word with them, can you develop that, and what about this number, will this? With others we might be going through the process of, are they slowing down, and looking at the explanation' (SPAT-Y4T). There is scope in this mastery approach to differentiate for pupils who are much slower learners and their needs for greater scaffolding, and so they may be differentiated teaching approaches or extra sessions outside of the class and if that does not provide effective, a differentiated timetable is provided,

'So, I have got a class of 23 and there are two really who do not really access everything that we are doing, just two, and they are the ones that could be left behind, so they might have... different provision in-terms of what they are actually doing, but the vast majority of the rest of the class would be given everything. If I think back to even four or five years ago, that would not have been the case' (SPAT-Y4T).

Parents have noticed and appreciated these efforts to keep all pupils studying in the same class rather than having slower learners taught a separate curriculum segregated from their classmates,

'In my daughter's class there's two of them that are on the same level, they still keep them, they do not keep them on the side, they do give them their own work but they keep them in the same class, they do do certain things away from the other children, but they do give them separate work, but they do work in the classroom which is good, they do not single them out at all which is good' (SPAT-Parent).

One of the criticisms of levels was the incentive to move quickly through the curriculum rather than ensuring that pupils were secure in their understanding of a concept and give quick learners an opportunity to deepen that understanding. Head-PAT explained that teachers in their school are taught to teach a concept over a period of time and to vary one element of that each day rather than trying to complete topics quickly,

'it is variation really, in-effect, which is what the whole thing is about; it is about keeping as much the same, and changing 1 thing - reduce the cognitive load by keeping the numbers small, keeping the steps small, and then as the children become more familiar with the concept, very quickly, the teachers can give them more sophisticated questioning and sophisticated problems to solve, but all children are accessing those things at the same time' (Head-PAT).

SPAT-Y4T explained that there was less differentiation in topic, and less pre-planned targeting of support, and an increase in spontaneous targeting of support,

'Yeah so the difference in terms of providing those differentiated worksheets and activities is that it does use a great deal of time but you still have to make sure that you still differentiate in terms of how you approach the lessons so that everybody might be doing the same thing, but the support might be different and you might be floating around to certain children more, but it might be different children each lesson. You might be directing your teaching assistant to go to some children more than others. So that might be challenging some children and having a word with them, can you develop that, and what about this number, will this? With others we might be going through the process of, are they slowing down, and looking at the explanation. So, it is more sort of, on the spot

differentiation rather than planning in-advance – it is more responsive this way because you could plan in-advance what you think the children might be doing but it does not normally work out that way’ (SPAT-Y4T).

There is also an onus on teachers to require students to provide more robust evidence of their understanding of a concept,

‘the focus we have here is that it is about giving children conceptual understanding, you know, beginning making sure you that their concrete and pictorial representations are supporting their representations of the concept... and then as the children become more familiar with the concept, very quickly, the teachers can give them more sophisticated questioning and sophisticated problems to solve, but all children are accessing those things at the same time’ (Head-PAT).

Head-PAT explained that together with the teacher knowing how secure the pupil’s learning is, they are keen for pupils to feel that they are secure in their understanding and for teachers to glean the thinking process that the pupil underwent to get to the correct answer. For this, the pupils have to write reflections either throughout the lesson, at the end of the lesson, or it could be one at the end of 2 days, depending on what the teacher thinks is appropriate. In these reflections, the child self-reviews their learning process and what they found difficult which is the crucial thing for the teacher to know so that they can scaffold the pupil when a similar problem is tackled in another topic. This encouragement of independent reflection was mentioned during the pupil discussions. One of the Year 6 pupils showed me that their teacher had written the acronym IW in their homework book and a couple of pupils explained that this meant Independent Working and so I asked them to define that and they explained that:

‘Because if you are always working with someone you know you can’t learn as much, but if you are working on your own you’re learning a lot more’ (SPAT-Y6pup).

‘I would say it sort of means there is... you can do it if someone is helping you, for example, if Billy was helping me with English and I was helping her that could be independent working as well. I think it is just when you do something mostly on your own, like without a teacher’ (SPAT-Y6pup).

These comments show the awareness that pupils had of the importance of taking responsibility for their learning and being actively engaged. Passive students often find it challenging to utilize feedback effectively, whereas active participation enhances future learning. Greater student autonomy in learning and assessment leads to improved outcomes and higher engagement, promoting a sense of responsibility for their own learning.

### 6.2.3 Learning from Mistakes

In both Schools PAT and Qualitative it was evident during my time in the classrooms that pupils were not being taught to not equate mistakes with failure but they were being encouraged by their teachers to understand the process in their thinking that created the mistake and the steps that are needed to get the correct answer,

‘for us, the correcting of mistakes, the identifying what had happened the day before is a very, very important part of the learning journey, and so, we do “response to marking” – AfL marking is part of the next lesson it is not an add-on, it is not a time that we need to find another lesson for. Every day, every lesson, in maths and English, starts with 5 minutes purple pen marking. So, that means that... they have to respond to the teachers marking (in green pen)’ (Head-PAT).

Head-PAT explained that this purple pen marking and green pen response is part of the process of teachers accessing the meaning-making of pupils,

‘it means that the child has had to concentrate on what the teacher has said, they had to process the request and try and put it in place. Sometimes, we miss one, sometimes the child does not do it, and sometimes they do do it, and sometimes it isn’t right. And then the teacher will actually write on there, that we are going to workshop this’ (Head-PAT).

This approach helps all pupils to have a secure understanding of fundamental concepts and skills so that mistakes are repeated less frequently, helps the pupil keep-up with the rest of their classmates and allows the teacher to focus attention on pupils that need support for that topic. It also avoids the whole class from having to listen to a repetition of concepts that they may have already mastered as they would be completing an activity that consolidates their learning while others respond to the feedback. Head-PAT explained that this implementation of an Assessment for Learning (AfL) marking approach was made possible

because of the removal of levels as previously a successful pupil was often one that was learning curriculum content from the next Year-group's curriculum,

'it became content-driven, and it was not about depth, and sustained learning, and so, that is why I am in complete favour of the changes that have happened, but schools just need just time in unpicking the difference between... and I think that levels confuse the difference between tracking and assessment, because assessment is knowing where your child is truly, and the tracking tool can be anything. You know, we can stop using Pupil Asset and use something-else - it does not matter. The point is, that the teachers have a very deep knowledge of the assessment of where their child actually is according to the national curriculum, and that is what any school has to match up' (Head-PAT).

SPAT-Y6pups conveyed that they found this combination of written feedback and attention from the teacher the next day in the classroom was particularly helpful to their learning as written feedback was not always understood by some pupils, and another pupil explained that when several pupils struggled to understand a concept or task, the teacher will teach the whole class on the board to ensure that there was consistency of understanding. When asked if verbal feedback without the green pen marking would be sufficient, the Year 6 pupils said that they preferred having both as the written comments signpost the specific point that needs improving and also acts as a reference for learning of what needed altering when this challenge was encountered previously. A suggestion that one Year 6 pupil made which was supported by the rest of the participants was that it may be useful to have a separate book where they record teacher feedback in all subjects as they do not always have books with them when they are at home and so can reflect upon it. When I shared this with Deputy Head, SPAT-Y6T, they took it as a sign of learner development as the pupil evidently cared about their progress and wanted to take responsibility for improvement,

'that just makes me so proud of them. The shift in their attitude to learning in the time that I have been at the school is just... there are nowhere else... it is unbelievable, the fact that they want to take that feedback home. And actually, the children that you spoke to were, some very studious, some not academic, some not studious – you had a real range' (SPAT-Y6T).

One thing that emerged frequently in conversations with staff and parents was the context of the school when Head-PAT became the headteacher in 2016 compared to 2019 when Phase 2 data collection occurred. The pride that SPAT-Y6T refers to is for the care that

some pupils now have towards their learning as a result of the approaches that the teaching team have adopted.

#### 6.2.4 Focus on Learning Outcomes

SPAT-Y4T explained that every lesson has a learning objective and the green pen and purple pen approach is used to ensure that that objective is met, 'the focus of the lesson was relative clauses so my feedback would be on relative clauses. So, this sentence does not contain a relative clause, could you rewrite it inserting one in. So that might be my feedback to a child linking it to the learning objective' (SPAT-Y4T). They emphasised the need to prioritise that as soon as possible after the original lesson so that it remains fresh in the pupil's memory. With the Year 4 class they do use verbal feedback if they have previously given written feedback and another question to answer,

'There is sometimes the element of verbal feedback as well, so I might be, as I say, going home talking to them or I might have spent a long time with the child, the previous lesson and actually I have given them a lot of feedback so I might just write "VF", verbal feedback, because I do not think there is too much more value in me giving them another task because we have talked so much in the previous lesson together' (SPAT-Y4T).

Head-PAT explained that the homework feedback being linked with the learning objective is also connected to the culture of pupil self-reflection and that it determines some of the lesson plan for the next day,

'Yes, you need to mark in-depth at least 10 or so kids depending on how you think the lesson has gone, because it goes back to that diary entry that I was talking about earlier, and making sure that if that child has a reflection and is showing you that there is a problem, then now in the lesson the next day we want to address that problem, so you cannot stockpile your marking and mark it all at the weekend, because you have missed that fluid journey of learning that happened. So, teachers do take my work home' (Head-PAT).

Head-PAT also explained that this has altered the quantity of homework questions as teachers are wanting to glean the thinking process of the pupil,

'we have the philosophy that we would rather see 4 or 5 high quality, tough, challenging questions in a book answered, than a page of 20 questions that have all been answered and so it is a mechanical repetition rather than a structured and deliberate variation. So, we would rather there be 5 questions where the child has really been challenged, and on reflection, oh I found number 4 very difficult because I was not quite sure about... – that is much more useful to us than just a page of them going through mechanical questions' (Head-PAT).

This ethos on setting tasks that challenge a pupil appropriate to where they are in their development is central to AwL. One of the criticisms that the Expert Panel (2011) had about levels was that pupils could coast if they were making sub-levels of progress. Head-PAT explained that AwL had changed their approach from giving *some* pupils work from a future topic to challenging *all* pupils to do activities that provided learning on the same topic at greater depth,

'We have, this Year, adapted our approach slightly in that we teach the whole class to greater depth standard... we know that not every child is going to make it. We scaffold as much as is right so that everyone accesses Expected standard work, everyone has the opportunity to access greater depth, we just know that some of them will get there and some of them will not... The whole thing is about challenge, the whole thing is about shifting the large mass up. So, we have got the large mass to Expected, and so, now we have got to try and shift them up to the next one' (Head-PAT).

SPAT-Y4T explained that they analyse whether the pupil can explain why they answered a question the way that they did to determine the security of the pupil's understanding and that this can be easier to ascertain in mathematics,

'I think it depends with different subjects but if I were to think of maths, I think you know someone is Above in maths obviously if they are able to access the material very well and apply it, but also if they are able to justify their thought processes, their reasoning and I think it is sort of tapping into that element of their work that really would provide them with a challenge. So, challenging our higher ability maths children to solve reasoning problems, to explain methods behind their recent problems and I think to get them to talk more about their learning, is one of the ways in which we might challenge them' (SPAT-Y4T).

Initially, the school tried to measure multiple curriculum statements to determine progress but they then condensed that to seven Key Performance Indicators (KPIs) for each subject in each Year group and they record the date when these have been met and attach the evidence from the pupil's work.

'previously was all in the same system, Pupil Asset, so we would simply tick boxes about statements of curriculum areas, and it worked, so colour coded. I think it was yellow or orange and green or something like that, with obviously green being that they were completely secure in that particular statement. I think we moved-away, I am not entirely sure of the reason for that, I think it was issues for management as they literally had numerous statements. Having said that, KPIs are similar in many ways in that you have got a list of statements but we do it on paper now is not electronic but it is a list of statements, you might tick it, you might date it when you think you have seen these things. So that is more like breaking down the levels to enable us to say they can do all of these things so they must be at Expected' (SPAT-Y4T).

SPAT-Y4T explained that this is then used to aid their understanding in everyday teaching on which pupils are likely to need extra support to catch-up and also during a future topic that is similar or building-on a topic taught previously.

#### 6.2.5 Extra Teaching to Reinforce Learning

As discussed in the Literature Review chapter, the main criticism that the Expert Panel (2011) had of levels was the way in-which they obscured the gaps in understanding that pupils had which in-turn prevented teachers from knowing how best to give support to them. While I was in both schools PAT and Qualitative, I witnessed class teachers asking pupils to remain in the classroom during school assembly or during morning breaktime to have a concept explained again by the class teacher and teaching assistant. One of the findings of research of Assessment for Learning (AfL) was that the immediacy of providing feedback and reteaching usually improves student learning outcomes and so SchoolPAT had embedded this into their approach. I explored this in my conversations with class teachers and Head-PAT they explained that they have two types of planned interventions and one type of more spontaneous intervention that they call a workshop.

The first type of planned intervention is generated by the PIXL electronic system based on the data inputted by the class teacher as SPAT-Y4T explained,

'you enter all their data into the spreadsheet and you then upload that to the PIXL website and they kind of break it down and they might suggest; they call them therapies which are basically interventions (for example) for a group of children that are struggling on rounding tens, here is a suggested therapy that you could work them on. So, some of those interventions we use but not all of them - we pick and choose which we think might give them support' (SPAT-Y4T).

The second type of planned intervention occurs after pupil test data had been inputted by the class teacher. The Deputy Head then analyses and interprets it and shares it with the headteacher, 'We look at it and if we have got seven children who are Below in Reading and I know that inference is getting in the way' (Head-PAT). The headteacher and deputy then invite the adults working closest with a pupil such as a Teaching Assistant, Special Needs Coordinator (SENCO) or class teacher to attend a pupil progress action plan meeting where they review the action plan that the class teacher has to have prefilled with a plan for suitable interventions, 'we are identifying the children who cannot fall through the net and then that action plan meeting is about putting something in place' (Head-PAT).

A workshop is organised on the same day when a pupil's response to the purple pen marking in their workbook reveals that they have not fully understood the teacher's written explanation and proven that by not answering the question set by the teacher to a satisfactory level, then the teacher will write in their book that they should attend a workshop that day at break time or at another suitable time outside of a lesson,

'So, a workshop is an on the spot because we go for a philosophy of teach-up and not catch-up. So, as we see a little bit of a gap, that may be developing, opening; on the day, we try and stop that. It could be in the afternoon during assembly where the teacher will say, I cannot come to assembly today because I am going to keep these 3 kids, because we need to go over what we went through today. And then, that is recorded in their books as well - that they have had an extra session on that one particular thing' (Head-PAT).

Parents also mentioned that they were pleased to be informed by teachers when these interventions were taking place, 'they always keep us in the loop' (SPAT-Parent).

'They do, they did this with me and my youngest because they just dropped behind a little bit in the maths. And so, they ask, is it okay if we give them extra one-on-one so he's had three extra maths sessions one-on-one per week to bring him back up to

level by the end of the Year' (SPAT-Parent). This continual monitoring of progress together with immediate identification of additional support and communication with parents is particularly important for pupils with Special Educational Needs and Disabilities (SEND) as they often need more time to grasp concepts and skills or they may always be working below the standard expected for their age.

#### 6.2.6 Pupils with Special Needs and Disabilities

While integrity in pupil progress judgements seems to have become more robust under AwL, conversely it seems to have obscured them for pupils with Special Educational Needs and Disabilities (SEND) in both SchoolPAT and SchoolQualitative. As levels were a continuous system from Year 1, a pupil could work at a lower level even if most of the rest of their class had progressed faster as they got older. Situating a SEND pupil's progress on a level also allowed teachers and parents to identify approximately where their learning was comparative to the rest of their class. By classifying students as being Below, At, or Above the Expected standard and resetting this each Year, the only category that most SEND pupils would get would be 'Below Expected' at every assessment point. SchoolPAT's Special Needs Coordinator (Senco) explained that the version of Pupil Asset Tracker that the school used in 2017-2018 provided a clearer perspective but the adjusted version obscured this,

'if they are significantly Below, the old system used to show that as well – "that they are in Year 4, but they are working at Expected for Year 2", whereas now it just shows that they're significantly Below. So, we might previously have a child that was making small steps of progress while being in Year 4 but to the Year 2 level, if that makes sense? Whereas, now, it is hard using that online tool to show the small steps of progress for those who are working Below their Age-Expected... it is not quite as transparent' (SPAT-SENCO).

To counteract the deficiency in this classification, teachers admit that, 'we are much more reliant on standardised assessment to show progress' (SPAT-SENCO), because the child is consistently sitting Below and is 'not getting any closer' (Head-PAT). SPAT-Y4T explained that these tests guide which interventions to use and then a test is administered to determine the effectiveness of the intervention and to plan the next,

'We try and identify gaps, plug those gaps with the teaching assistants who will be taking children at various points in the day with an aim to try and plug those gaps. And that is across the board, so it could be a maths intervention, it could

be an English intervention... Just to try and make sure that they are close to the main bulk of the class' (SPAT-Y4T).

SPAT-Y4T explained that the information on the AwL system from the previous Year-group's teacher means that they know when a pupil comes in their class working Below where they should be for their age and they make a specific plan for them to make what they called 'accelerated progress' while not overwhelming the pupil. The teacher is also cognisant that the level and quantity of work that a SEND pupil produces may mean that more examples of work may be needed to evidence the attainment of a curriculum statement construct.

One parent of a pupil with SEND explained that they received several written reports during the school Year from their teacher who then took the time to also explain them verbally which seemed to help significantly, 'they explain really clearly so you understand what they have done, where they are, and where they are going to go next. And then you can incorporate that at home' (SPAT-Parent). SPAT-Y4T explained that they try to ensure that feedback to both pupils and parents was specific, 'I find that it is very helpful to give quite specific targets for parents to work on: they learn their three-times tables by a certain point, they need to be checking their work and they need to proof-read things or they need to be reading every night' (SPAT-Y4T).

#### 6.2.7 Reporting Pupil Progress and Attainment to Parents

Head-PAT explained that schools have a statutory obligation to report to parents whether their child is working At Expected or Below Expected levels for that age and so they have decided to use those classifications in all Year groups. SPAT-Y4T felt that this was far easier for the parents to understand although they commented that even under levels they would not tell parents about sub-levels and would be telling them whether their child was where they should be for their age. The Expert Panel (2011) had argued that reporting levels obscured the focus for parents to know what they could help their child learn at home but SPAT-Y4T felt that they were doing that under levels and was unconvinced that this had been changed under levels,

'In some ways the messages are very similar to what they were with levels than what they are after levels. Whether levels has had a huge impact, I am not sure in what we discuss with the parents. I do not know if it has because in some ways the gaps are still there, and you still address the parent in a similar way regardless of the level' (SPAT-Y4T).

SchoolPAT was situated in an economically deprived area where the educational levels of parents may be variable and so teachers had to discern what they would be able to help their children with but as a minimum a parent could listen to their child reading a book aloud,

‘some parents may not fully grasp what you are trying to say or they may not be particularly literate and there is a limited amount of support that they can give, but the parents might not be able to help their children with that or with their subordinate clauses, but you know that that parent listens to them reading and that is a real benefit’ (SPAT-Y4T).

The deputy headteacher and Year 6 class teacher, SPAT-Y6T also did not feel that AwL had improved parents engagement with their child’s learning,

‘I have not seen an impact. I think that some parents take an interest and support their children. I think that would have happened probably with or without levels, but for this school the majority of parents are happy to send them in and let us improve their work. They do not read with them, they do not do spellings with them, they do not work on their homework with them’ (SPAT-Y6T).

To improve the connection between the classroom and home, SchoolPAT had invested in an electronic application called ClassDojo<sup>14</sup> which allows teachers to upload photos of children’s work during the school day so that parents can see what their child is achieving and it also allows parents to upload photos of homework and extra-curricular activities. SPAT-Y6T believes that this system is a positive addition but that most of the interaction is from classroom to home,

‘They are beginning to interact a little bit with our Dojo links which is our reward system where I can send pictures of work home and get back, “oh well done” or “that’s really good”; what I do not see is parents sending me pictures of, “oh look at them reading their book”, or “look at them reading the newspaper at home”, or “we went to the museum”, or “we went for a walk and they pointed out this leaf” or something like that linked to what we were doing (in the classroom). We have not got that link at the moment’ (SPAT-Y4T).

---

<sup>14</sup> <https://www.classdojo.com/en-gb/>

SPAT-Y6T reflected upon whether some parents struggle to understand what their child's next steps of learning because the communication was not as regular as it needed to be as teachers relied to heavily on written purple pen marking in pupil's workbooks and on limiting communication to report cards and parent meetings at the end of the school Year,

'I do not think they always do (understand). I think that we do not communicate those next steps for learning every week or every term; we do that at parents evening. And obviously their homework has feedback and so that is clear with next steps... and so that is a good communication tool. But I do not think we communicate individual next steps regularly enough I suppose for them to be able to do anything about it if they wanted to, really which is food for thought' (SPAT-Y6T).

This issue of timing of feedback was also commented on by a parent that could not attend the focus group with me but sent a written comment.

'We love them (report card and parents evening). Very informative. I understand all the information given. If I had one critique it would be that I would like a little more time to digest the comments so that we could speak with the teacher concerned before the end of the school term' (SPAT-Parent).

I read this comment to SPAT-Y6T during my interview with them and they felt very pleased to know that that parent had had a good experience with teacher feedback and felt that the criticism about timing was reasonable, 'that is absolutely a fair point for those parents that do want to support the learning of their children' (SPAT-Y6T). After the interview SPAT-Y6T admitted that the late sharing of the report card was a deliberate ploy because teachers had been challenged more than they were comfortable with when providing the report card earlier in the term (June) two Years previously.

#### 6.2.8 AwL in the Wider Curriculum

At the time of data collection in 2019, Ofsted had indicated that their inspection framework would begin to focus more on core and foundation subjects and not just those tested nationally and thus compel schools to widen their focus in the curriculum,

'We're starting to look at assessment in the foundation subjects and assessment in science, not necessarily tracking as rigorously as we would for maths and

English, but certainly for teachers to have an understanding and have a knowledge of how the children are progressing, for design and technology, for art – but that is an approach rather than a system of tracking’ (Head-PAT).

Head-PAT explained that they were preparing this for a launch in the subsequent school year across all Year groups. SchoolPAT use a topic curriculum and teachers were in the process of selecting key statements from the National Curriculum for each topic in each subject and from that they will extract success criteria from an electronic package called Common Sense<sup>15</sup>, that both the teacher and pupil will be expected to reflect upon as they work,

‘the project outline sheet will have all of those objectives on there, and the success criteria for the child as well, and so, the child will be, the whole time, reflecting themselves and thinking, can I do that, and can I do that, and then there’s space for the teacher to assess that as well’ (Head-PAT).

The Year 4 pupils explained that the structure of science lessons compelled pupils to utilise higher-order skills such as evaluation, ‘In science, we write down in our book about our experiments ., We always use the sheet then we used to put down how well we thought the experiments were and we ... put down our conclusion and methods and stuff’; and also self-assessment, ‘we used to put down how well we thought the experiments were and we... put down our conclusion and methods and stuff’ (SPAT-Y4pups). The Year 4 pupils explained that while much of the feedback in PE was to improve behaviour, they do get feedback in subjects like Music, ‘we get feedback from the teacher and he says to move your hand over and make sure your fingers are in the right place’ (SPAT-Y4pups), but would appreciate more feedback across subjects so that they improve themselves and develop their repertoire in subjects like art, ‘I think yes because people will get used to them the more they get it right, the more stuff they could do it in art’ (SPAT-Y4pups). One pupil felt that the lack of feedback in all core and foundation subjects resulted in a reduced motivation for some of their classmates, ‘sometimes some people, they do not try, and they just like sit there, and do not do anything, and then it kind of gets to me, because it is a subject that we have to do; it is not like it is like it English and maths. but like... sometimes we might need that’ (SPAT-Y4pup). That pupil proceeded to explain that they were referring to transferrable skills giving the example of needing to use drawing skills from Art when studying Religious Education.

---

<sup>15</sup> <https://www.commonsense.org/education>

One pupil shared that the type of feedback that they got in geography tended to be on literacy skills rather than on geography, 'its mostly spelling, because we do writing in geography in our books' (SPAT-Y4pups). Deputy Head, SPAT-Y6T. explained that this focus on developing literacy and numeracy skills was a deliberate policy to reinforce core skills as that had been a concern of inspectors,

'although we are looking at specific subject skills, we are still having a huge portion on the basic literacy and maths skills. Partly that was a conscious decision of the leadership team because the levels were so poor in school and children had massive gaps and to move from special measures to good to outstanding, we had to make a conscious choice' (SPAT-Y6T).

These quotes reflect the findings of the House of Commons Education Committee Report on Primary School Assessment (2017), which concluded that the high-stakes nature of the assessment system negatively impacted teaching and learning. This pressure resulted in a narrowing of the curriculum and an emphasis on 'teaching to the test.' The Committee noted that the focus on achieving results in English and maths often came at the expense of other subjects, such as science and the arts.

### 6.3 Phase 2 Data from SchoolQualitative

#### 6.3.1 A Culture of Collective Problem-Solving

School Qualitative's Deputy Headteacher, SQ-Y6T explained that they were creating a culture where all teaching staff take an interest in a pupil's progress rather than just for the pupils in their current class 'we're trying to get everyone to see the kids going from a to b, and that you're either teaching them, have taught them or are going to' (SQ-Y6T). They also want to encourage peer dialogue and not just dialogue between a class teacher and SLT so that all teachers have a 'helicopter view' (SQ-Y6T) – a broader perspective on the attainment of pupils across the school, 'So SQ-Y4T has got a kid right now, a real high flier, so we can have a chat about him on Tuesday, and we're going to get SQ-Y5T down from Year 5, and I'll be there, and we are going to talk about him because we're part of the system' (SQ-Y6T). They refer to this culture of considering the continuity of a pupil's primary school learning journey as "opening-up classroom doors" and it encourages a deeper understanding in every teacher of a larger number of pupils, 'I could sit down now I think, and have a pretty good conversation certainly of about 20 of those kids in Year 5. I could give you fairly-good

breakdown of where they are, because we've been consistently talking about constructing a plan that then feeds into the classroom with that AfL' (SQ-Y6T). They hope that this sharing of experiences about a pupil may help a teacher to access an experience about the pupil or of a similar situation which could assist a colleague to successfully navigate an issue or optimise the learning experience and outcome of a pupil,

'the more we talk about the problems that the teacher has in finding the right levels for them, and how to move them, the more people are free to talk about how hard it is to teach, with no expectation of what they should be doing, but are they doing enough?; are they moving?; are they shifting?; are you happy that they are doing the best that they can do?' (SQ-Y6T).

Ultimately, the culture that SchoolQualitative are trying to embed is designed to reduce fear within teachers who may be reticent to focus on what pupils are struggling with as it may imply that they are not good at teaching them. This concern could also influence teacher judgements on pupil progress and so it is hoped that this culture of dialogue increases integrity in data entry,

'There's no pressure on moving from one number to another which some can massage. I think that's important... it's really important. The last few data drops that we had to send to the governors, a couple of classes said that, you know what, I'm not entirely sure, so we didn't present the data. So, rather than putting in false information that we're just going to be chasing, we actually identify putting in a plan' (SQ-Y6T).

This collective culture also extended to whole-school teaching programmes such as maths fluency and has altered the content and tone of staff meetings towards a culture that encourages experimentation within the classroom and sharing of what works,

'that has been a whole school push to look at fluency in staff meetings and coming back together and sharing the ideas as to what has been done and sharing things that have worked well, and sharing things that we are discovering, so there is a lot of shared discussion about what we are doing. Our staff meetings are really about everyone sharing because we are all working on the same thing' (SQ-Y4T).

### 6.3.2 Classroom Dialogue

AwL seeks to enact one of the principles of Assessment for Learning which is to create an ongoing dialogue between teachers and pupils so that daily instruction is adapted to the needs of pupils, 'it is not just, as it was with levels, where you are assessing at key points throughout the year and taking that level. But it is much more of an ongoing process and we're doing it daily, and it is really then about taking it almost lesson by lesson' (SQ-Y4T). SQ-Y4T explained that their teaching of a curriculum statement would be split into several concepts that they would then teach in a series of lessons. Initially each pupil would be given a lot of support and then that would be reduced as the pupil became more secure in their understanding. At the end of each concept, a worksheet is completed by each pupil that is then marked collectively in the class. The teacher has a list of key points that they wanted the pupil to know and they tick this as students convey answers as a means to triangulate their own perception. The gaps in understanding then inform the teacher as to which pupils will need additional support and scaffolding in future lessons. SQ-Y5T explained that they record all the skills that need to be re-taught or repeated and that they verbally inform SLT as part of ongoing conversations about all pupils. Together with this, they make a written judgement on whether they think that a pupil is Below, on-track, or Above where they should be. As part of this, teachers identify a specific area that a certain child is not on track and analyse why that progress might have stalled,

'and that conversation then I feel is more powerful than saying that this child needs so many points progress throughout the year; they need to hit this at the end. The discussion is no longer about, where they need to be, it is about the progress that they are making which I think is powerful and I think within that, you get where they need to be anyway' (SQ-Y5T).

This sharpening of focus on the underlying reasons why a pupil may not have made the expected progress is educational for staff as it makes them think on *how* children learn a particular topic or skill rather than just monitor *whether* they have learnt it. It is also understood by pupils as evidenced by this passage from a Year 6 pupil discussion,

'When others, like say the faster ones are helping the slower ones, it is never a way of just giving the answer to them. We never give the answer to that, but we always help them go through the method - but we never go straight to the answer.'

There's keys and there's locks, and there is always a door that you need to get past and there is this one thing that you can always see that unlocks it, and you have done it.

And then you see a point in the question where you can accelerate, and it suddenly clicks, and you are off.

Yeah. And the answer is never the priority, the answer is secondary, the method is priority' (SQ-Y6pups).

What is evident from the passage above is the depth of understanding that some of the pupils have on the importance of understanding the process of learning and making effort to get answers rather than just being correct.

### 6.3.3 Embracing Challenge

SchoolQualitative are trying to encourage students to take responsibility for improving their own learning and choosing company that challenges them. At the end of a topic, pupils engage in two types of self-assessment in discussion with a classmate that they sit next to: 1) their progress according to their learning target and 2) their personal study skills,

'did they use group work effectively?; were they focused?; did they manage to achieve independence?; are they happy with their outcome? - so talking like this and a pair they can pick up on these things to the evaluation. They know that they work really well together, but are they beginning to feel that is a bit stale and they need someone else to work with - someone else they can pick brains to sharpen them; just because they know that they work well together doesn't mean that you want to sit next to each other all the time. Trying to get *them* to make those decisions' (SQ-Y6T).

During group discussions, Year 6 pupils in SchoolQualitative showed me their self-assessment and self-evaluation reflection question sheet which they said they used mainly to reflect upon written work in all subjects (Figure 14). This desire to challenge themselves or to be challenged so that there is continuous improvement in their work is embedded in pupils thinking from Year 1 as Head-Qualitative and the Senior Leadership Team (SLT) have decided to adopt the approach advocated by Diane Pardoe (2009) who argues that pupils learn best when they feel comfortable to be

challenged by an activity and to challenge themselves which is known as the 'danger zone', and that they will only do this if teachers are willing to do this and if the feedback that they provide praises the effort in tackling the challenge rather than just for completing the work correctly.

'very early down in the school, it starts with the understanding that the children can actually judge how challenging a piece of work is and can start to account for their role in it - I think that's what it really comes down to... don't say that something is right or wrong, don't say something is hard or easy, but take control of that and say, "I am currently finding this challenging, it's in my danger zone, and therefore, I need to find a way of making it challenging rather than dangerous", and similarly someone sitting there saying this is easy - that isn't good enough' (SQ-Y6T).

This notion of challenge is also continued into classroom discussions as SchoolQualitative encourage pupils to not only say what is in their mind but to listen and respond to what their classmates have said by either "supporting", "expanding-upon" or "challenging" it as some of the Year 6 pupils explained,

'Yeah, we use that for everything. If we do something in class... the whole class – that can be different ideas. And we can challenge, support and expand it. If you support, that means you agree with that;

Challenge is like, no I think like this...

And then expand is where you can expand on it – I don't think you can challenge and expand.

You say agree and expand.

And, sometimes you get a chance to challenge, for example, I think there was one the other day where one pupil challenged another, and then, that pupil challenged the person challenging.

It is like a debate.

Yeah.

You are... it is not like a proper debate where we actually are doing a debate.  
We are doing English but we debate.

We debate with SQ-Y6T...

Yeah, you have to be original in your answer' (SQ-Y6pups).

The extract above from one of the group discussions with Year 6 pupils conveys the interactive dynamic that teachers have created within the classroom in all subjects. Pardoe (2009) posits that pupils will tend to want to stick to activities that are in their comfort zone if they are afraid of making mistakes and so the psychological connection between mistakes and failure must be ended to enact Dweck's theories of a Growth Mindset. It was evident in the group discussions with pupils that they understood the distinction, and in this section they explain that they do not fear mistakes as school values underpin this ethos,

'we know we can learn from that (mistakes) as a class, we can see what other methods other people have done and we can build on that.

It is a school value, sort of, learning from my mistakes, and the GRIT' (SQ-Y6pups).

Duckworth (2016) argued that learning required certain psychological traits which they termed GRIT which stands for Growth mindset, Resilience, Independence and Teamwork and this is what the pupils were explaining that had been taught to them. Other Year 6 pupils conveyed their analysis of different types of mistakes,

'There are two types of mistakes – there's silly mistakes and there's mathematical errors. Silly mistakes is where you ...

You are daydreaming and you are not really thinking about the questions.

You might see it as a division, and it is actually a multiplication.

And there are sometimes three, because you might find a mistake and then correct it, but it is usually a fact that it is a mathematical mistake and you have actually got the method wrong, and you would not get in a silly mistake

And SQ-Y6T likes mathematical errors because they can help with them, but if it is a silly mistake they can't.

So mathematical errors are where you have not really come across that method before ...

Or you have, but you might have forgotten it' (SQ-Y6T).

Other pupils shared that mistakes were used as a tool for SATs preparation as pupils reminded themselves of common errors that occur when answering a question by deliberately formulating a question with those errors for their classmates to answer with the expectation that pupils could identify the error and have the confidence to challenge and point it out,

'So, we tried to do mistakes that we might make in the test and so if you doing arithmetic you might write down the mistakes that you might make, so you write it out, and then you write out the wrong one in the answer book, and SQ-Y6T moves-on and then people will put a hand up, and then challenge that answer' (SQ-Y6pups).

The tone with which adults respond to mistakes was considered important by pupils as it impacted particularly on these psychological traits,

'If it's a mistake, the class is like... SQ-Y4T sometimes will tell you, just like, close try, good thing; but sometimes she will say, that is wrong, but she does not say, that's wrong, you should work more, why are you not doing enough?, because she's really nice on giving feedback' (SQ-Y4pup).

'And if we admit to SQ-Y6T, they do not get angry at us, but they, sort of, make a joke about it.

SQ-Y6T says, put it down as a next step and practice when you come first thing in the morning.

*Researcher: So, the teacher's attitude is also important?*

Yeah, that is really important' (SQ-Y6pups).

#### 6.3.4 Embedding Mastery

The 'Mastery' curriculum enacted in 2016 meant that ability-streaming was no longer used and so pupils labelling themselves according to ability was absent from conversations, 'there being no fixed-ability grouping, there is no, "I am in that group"' (SQ-Y4T). During group discussions, when Year 6 pupils were asked whether they thought it would be a good idea to have ability streams none felt that that would be desirable nor did they know whether they were working Below, At or Above Expectations, 'if you asked any of them, they would have a rough go where they fit ability-wise in the class, but that is not embedded - it is not the culture of the school, and so they do not have that as the final feeling. So, a lot of them still, hopefully, see each other as equal' (SQ-Y4T). These categories are communicated to parents during the Year if the pupil is below Expectations so that everyone can help them both inside and at home, and for all parents at the end of the School Year when report cards were issued but they are not a label. Rather than pupils staying in an ability group throughout their school career, all pupils in the class learn a topic together and extra support is given by the teacher on a topic-by-topic basis meaning that it will be normal for different pupils to be receiving help from adults and pupils are also encouraged to assist each-other, 'I am not presuming that any of the children are going to find it easy or find it difficult. We take it and do it on a need-by-need basis' (SQ-Y4T). When I was in the Year 4 classroom, SQ-Y4T explained to me that pupils working Above Expectations complete work for a topic quicker than most of the class and so they are given more challenging problems to complete on that same topic or they may be asked to create their own questions for others to answer. Slight variations to a topic are used to make it progressively more challenging,

'we have been looking at two-digit numbers, divided by a one-digit number. Some of them tomorrow will be looking at a three-digit number divided by a one-digit number. But not all of them. So, the ones that have been identified as still needing some support with this will stay on two-digits and the rest will then move on. So, it keeps them all within the same context and the same mathematical topic and pushes the ones on that are ready' (SQ-Y4T).

This messaging from the teacher is also needed from those that assist them both in terms of the feedback that they provide and where they focus their time, 'the way that you deploy the

TA's has to be different, because they're not now there to staple things to walls and make things look pretty' (SQ-Y6T). Teaching Assistants were now considered as professionals who meet with the teacher regularly to understand learning expectations and are expected to identify gaps in pupil understanding and inform the teacher. Rather than the Teaching Assistant working with the same child or group of pupils every week, they would share the teaching of those pupils with the teacher to ensure that those pupils are also purple learners, 'We had, maybe three years ago, still a point where we would talk to teachers about their class and we would regularly hear from certain places that, I don't know, I don't teach them, and I'll have to go and check with the TA, which for us, was criminal' (SQ-Y6T). If teachers were not teaching all pupils within their class then there would be a number of pupils who were excluded from the whole-school conversation on pupil progress that SchoolQualitative are trying to establish.

### 6.3.5 Extra Teaching to Reinforce Learning

While I was present in the Year 4 class, I witnessed a group of children being held in the classroom to receive some extra teaching from SQ-Y4T and one of their Teaching Assistants while their classmates attended assembly. SQ-Y4T explained that approximately half of the class required reinforcement of some fundamental mathematics concepts that were covered in a previous Year group which would then also make it easier for them to grasp the next topic enabling them to continue working with the rest of the class.

'I have got two set intervention groups at the moment. So, there is about 10, 8-10 children that have got gaps from their previous knowledge that we have identified, and so they are working outside of the maths lesson. So, they're now working at some of the concepts from lower-down the school. My groups are all looking now at fraction's work, because when I finish division, the whole class are going to be doing fractions and so it is almost a bit of pre-teach to go over the basics for them, to give them a bit of a head start when we start fractions in the class. And so, it is partly for that. But we also we go through some of their gaps' (SQ-Y4T).

This commitment to keeping all pupils in a class together when introducing a new topic is a significant change under AwL and the new mastery curriculum, and puts the emphasis on the teacher to provide additional teaching or a varied instruction rather than just classifying the pupil to be of lower ability.

### 6.3.6 Verbal Feedback Groups

A central part of SchoolQualitative's AwL pedagogy was the use of verbal feedback groups during classes rather than providing written feedback on pupil's homework. Head-Qualitative explained that, 'our Year 6 teacher has not taken a book home in three years to mark but when you talk to the children about the feedback that they get, they can talk to you about it and how they're going to improve their work... where they are not responding to the feedback we know that there is an issue and that there is a problem with the feedback that we are giving' (Head-Qualitative). SQ-Y5T described the process, I had 4 children in my group, and they were doing point, evidence, explain, and two missed the evidence quite clearly. One child had a very strong point, evidence, explain, and it was very useful for those two to see that one, and also to see that that one could also still be improved from the point that it was; and they will often work together and they will have their own individual targets, but it was really useful and powerful' (SQ-Y5T). Some of the Year 6 pupils explained that they have a level of autonomy when they enter the feedback groups as some have pieces of writing that are polished and others enter it knowing that they have many edits to do, 'Before we look into the actual work in the feedback groups, we do something called a professional check' (SQ-Y6T). The reason why the pupils have to judge their own readiness is because they know that their work will be critiqued by the whole feedback group, and teacher collectively, 'there's usually five or six people, and we go through our work on the board' (SQ-Y6pup). SQ-Y5T observed that pupils learnt from teacher feedback given to another pupil, 'you would share your feedback with them, and the other children would comment on it at the same time so that they are all learning from the child's piece of work and then applying it to their own' (SPAT-Y5T).

Learning by listening to teacher feedback on classmates work was also understood by younger pupils in Year 4, 'they say to you, but they are not really saying it just to you, they're saying it to the whole class. Because if you have not done it, and someone else has not done it, they're not just saying it to you, they're kind of saying to both of you' (SQ-Y4pup). They sometimes write that verbal feedback in their book so that they can refer to it later and another Year 4 pupil said that they sometimes write feedback given to a classmate, 'you can then write in yours about someone else's' (SQ-Y4pup). Contrastingly, one of the Year 4 pupils expressed that they would prefer private feedback just from the teacher, 'Yeah, I would prefer it, if it was just singular feedback, because it is only the teacher that is going to give you feedback'. However, it is this added element of dialogue on improvement that seems to distinguish verbal feedback groups from written feedback. One of the Year 4 pupils also said that they listen to the conversations within a group when they are not participating

and 'try and see if there is something really good, and I can put it into what I have done'. SQ-Y5T felt that written feedback was more useful for when a pupil was mastering a concept well but those who are struggling need to see why they not understanding, 'they're finding it difficult for a reason, and they need discussion' (SQ-Y5T). This dialogue also helped teachers and pupils be clearer on their next step of learning. SQ-Y5T kept a list of next steps for each pupil in their class so that they have an overview of how they have developed throughout the year and the pupils can see their progress too as part of the ongoing classroom dialogue. This was confirmed by one of the Year 4 pupils, 'SQ-Y4T has these sheets that she gets out of the folder; we do not know what they are – she never shows them to us. And I think she ticks them off, if we can do that. and she sees that we have done that, and if we can do that then she takes it off on the sheet' (SQ-Y4Pup).

Part of pupils ownership of their self-improvement is their process of practising their writing and mathematics skills. Year 4 pupils explained what their different books were for,

'And our red books are English skills books.

In your main book, you basically write your stories, or your diaries, or whatever the piece of writing you're doing.

And then, we have a skills book which is basically the background, like planning, and we can like, plan sentences or like, yeah. And then in maths, we also have a main book where we do everything. So, the main book is for doing like, the Deepen It sheets. And the skills book is the blue one, where we practice the methods, we do the calculations' (SQ-Y4pups).

From an early age, pupils learn to evaluate the difficulty of their tasks and take responsibility for their learning progress. They practice their writing and mathematics skills in separate books — one for main work and another for skills practice. This approach helps them plan and refine their work, promoting continuous improvement.

### 6.3.7 Pupils with Special Educational Needs and Disabilities (SEND)

During the group discussion with parents a personal perspective from a parent, with two children with Special Educational Needs and Disabilities (SEND) under SchoolQualitative's AwL system was very insightful. Their son was a Year 4 pupil in the school and their daughter had now entered secondary school,

'I do not have academically-driven children, so for me, the removal of levels was very difficult, because I had a child who was in Year 4, still on what would have been P-levels, and when you are looking around at other schools for those children, it is very, very difficult because they ask, what level is your child at? And, because there is no definite level that everyone understands, if you go to the school down-the-road or a school in another county, there is nothing to actually say where your child is, because the school is so different, and so, for me the removal of levels was really, really hard' (SQ-Parent).

A parent explained that their daughter had a dedicated Teaching Assistant who helped her to be able to enjoy learning and she was able to read, 'Although she was not learning in the same way or the same content as her peers, it did not matter, because at-least she was still learning. And, she was happy, which for me was the most important thing' (SQ-Parent). The challenge emerged when they tried to secure a place at secondary school as they required a level or SAT prediction which this pupil did not have, 'you were walking into a mainstream school, and I am showing them books – this is what she is reading, she is a P4 level, we do not know if we are ever going to get to a P6, and they just look at you like, why are you even here?!' (SQ-Parent). In contrast to the increased clarity documented by teachers about pupils next steps of learning under SchoolQualitative's AwL system, that parent found it more difficult to know what their children knew and whether they were making progress than under levels,

'we had the levels, we knew what levels were, and we knew where we were going next. We could say, we need to move on from that - to get to that, we need to do X, Y and Z. We need to do this next, and how is that? Do we need an intervention? Do we need something else brought in to actually work, so that we could actually move forward? Now, I find that when I go to the parents evening, and obviously it is with the things that he struggles on, and I have had so many conversations with the teacher, but I know that he really, really struggles with the – where is he? Is he two years behind? Is he one year behind? How do we improve this situation? And, I am kind of a person where I like to know where everything is' (SQ-Parent).

Under SchoolQualitative's AwL system their son would be classified as Working At the Expected level because it would be relative to what was expected for them rather than relative to the Expected Standard for the Year-group.

### 6.3.8 Reporting Pupil Progress and Attainment to Parents

SQ-Y5T explained that AwL had positively impacted on their experience of giving feedback to parents in-person at parents evening, 'I would say the conversation with parents is better because they are not just looking for levels and they are not just looking for a category anymore' (SQ-Y5T). They explained that the conversation is about their child's work over the Year and they may talk through examples of the improvement in their workbooks so that parents can see progress. The two main benefits that they provide are that both pupils and parents are very clear on their next steps of learning and they also felt that the pressure to perform had reduced on children,

'I feel that it does take a lot of the pressure off of the children then, because the children do feel the pressure. If they are not attaining and if they then look at their progress, they then feel better about themselves as well. I have had some really nice parents evenings where you can say, the children are making this much progress, and they can see the progress through the books because of the way we structure, and the way that the children are actioning the feedback, and the discussions are back-and-forth, and I know more, myself, about the child as well. And so, I feel it is more powerful' (SQ-Y5T).

At the end of a School Year, teachers give a written report card to parents and SQ-Y4T explained that the front page defines the child as a learner where the teacher gives them a star rating on their attitudes and characteristics relating to learning followed by a written comment on this. For English and maths, pupils also have a star rating but this relates to whether they are making excellent, good, or slow progress for each learning area in of English or maths, and in the comment is a statement is made about their attainment and whether they're confident, and working at end-of-Year expectations. And then for all the other learning areas like science, computing, geography, history, the teacher gives them a star rating for attitude to learning in those subjects. SQ-Y4T perceived that parents appreciated this format as it was more essenceful compared to a previous iteration and the comments were tailored to a specific child, 'I think a lot of parents felt that if you have a very long report, it is almost harder to be more specific to that child' (SQ-Y4T).

### 6.3.9 AwL in the Wider Curriculum

Head-Qualitative explained that they were changing from a topic based curriculum to a subject curriculum to make sure that children 'think like a geographer, think like a scientist'

(Head-Qualitative), because they were influenced by Myatt (2018) who argued that pupils struggle to connect their learning during primary school with the subjects that they are taught during secondary school. SQ-Y4T said that they felt that this made it 'easier to make sure that we are doing the skills that they need' but 'possibly there isn't currently a rigorous assessment in place... like we have for English and maths' (SQ-Y5T). They proceeded to explain that some subjects like science have a curriculum that is structured in a manner that make it easier,

'I find with science it is also quite easy to implement, because of the way that science is structured. They have got their investigations, you can pull the name and talk to them about their investigations, and it also shows you a lot about their knowledge and understanding of the subject and how to go about the scientific investigations' (SQ-Y5T).

SQ-Y5T explained that the school has embarked on a project for all teachers to see how a topic or concept is being taught from Reception class to Year 6 with a particular focus on how a concept or skill should be developed in Key Stage 2. Both SQ-Y4T and SQ-Y5T explained that the school has had a particular focus on improving written literacy and so the coherence in this throughout the Year groups was more developed, 'Maths is probably not as refined as literacy' (SQ-Y5T).

#### 6.4 Chapter 6 Discussion: Enacting AfL by reducing Cognitive Load

When Phase 2 data collection commenced, it quickly became apparent that an exploration of AwL was largely an exploration of how AfL has been enacted,

'So, what you've done is taken AfL and put it right at the centre of the pedagogy, planning... And, that's basically what AFL is. ('Absolutely'). Where do you see the difference between AfL approach and a AwL approach? (Researcher).

'There should be no difference - assessment is assessment' (Head-Qualitative).

'So, are you saying to me that actually you have done a big piece of work on AfL? It is Assessment for Learning – it is not really about a tracking system? It is an in-class type of assessment' (Researcher).

'Exactly. It is, and that is what is making the teachers able to make an informed judgement and we can trust teachers to bring the data to us six times a Year and

we can input the data, because we know the quality of what is going on in the lessons is very high, and that the teachers know the kids very well, and that is what assessment is – it is knowing your children well, knowing what they need next' (Head-PAT).

This certainty of understanding that the enactment of AwL was largely the enactment of AfL expressed by both headteachers surprised me because the Expert Panel Report (2011) did not mention AfL and the CAWL Report (2015) only mentions it once in relation to it being part of initial teacher training. However, the data presented above and my own observations shows that many of the practices associated with AfL were being embedded. Teachers seemed committed to the core principles of AfL to create meaningful reflection on learning as part of a process of creating self-regulated learners. It seemed that teachers felt accountable and were focused on students achieving their goals and the alignment of the AwL system with curriculum statements seemed to assist with that. In both schools, teachers were drawing out students' thoughts and leading discussions, and feedback was grounded in that dialogue between teachers and students or among students themselves. These conversations drew out students' pre-existing ideas to address prior knowledge and misconceptions, and then that was used as a starting point to build understanding of concepts and also next steps of improvement in skills. It is this integration of classroom activities, homework, conversations and feedback that is indicative of successful enactment of AfL which is what Head-PAT is conveying in the quotation above. Teachers conveyed how they would adjust teaching approaches and plan extra support to ensure that all pupils had grasped key concepts and skills both by interpreting pupil learning in real time and also based on marking homeworking. These interpretation and adjustment skills for all pupils in a class are also evidence of the enactment of the Expert Panel Report's (2011) mantra of 'high expectations for all' being combined with AfL principles of teacher responsiveness to pupil needs.

Students were provided with opportunities to discuss feedback with their teacher and then given time to act on feedback during lessons and as part of homework tasks. The pupils conveyed their excitement and engagement with their work and teachers feedback and a knowledge of their own personal learning goals and a clarity on how to achieve them and indeed evidence that they re-read teacher feedback. There was no sense of passivity with pupils showing a positive attitude and active role in their own learning demonstrating that teachers had fostered autonomy and responsibility. Pupils involvement seemed to be meaningful and useful for future learning and they were encouraged to self-assess which emphasised a shared responsibility for learning and assessment. It was noticeable that

pupils did not refer to test scores or end of year grades but only to qualitative feedback in both schools. The school leaders have played the pivotal role in enacting AfL by establishing a school-wide culture with a clear vision and expectation for AfL use. Teachers have been provided support in using AfL-related teaching strategies, principles of good feedback, and effective questioning. This support included instructional resources, materials, and examples together with time to prepare AfL activities. Practice-centred collaboration was also important, with teachers engaging in conversations about formative assessment and teaching, and collaborating on shared problems. This collaboration was essential and particularly emphasised in SchoolQualitative, with teachers working together and engaging in communities of practice for long-term learning of AfL techniques and to problem-solve to discern best solutions for all pupils. Underpinning this, is a positive school climate based on trust, mutual respect, and cooperation which gave teachers confidence to try new approaches. Teachers were trusted to create assessments that they felt would most authentically glean what had been taught in the class and they were supported in that process by senior teachers.

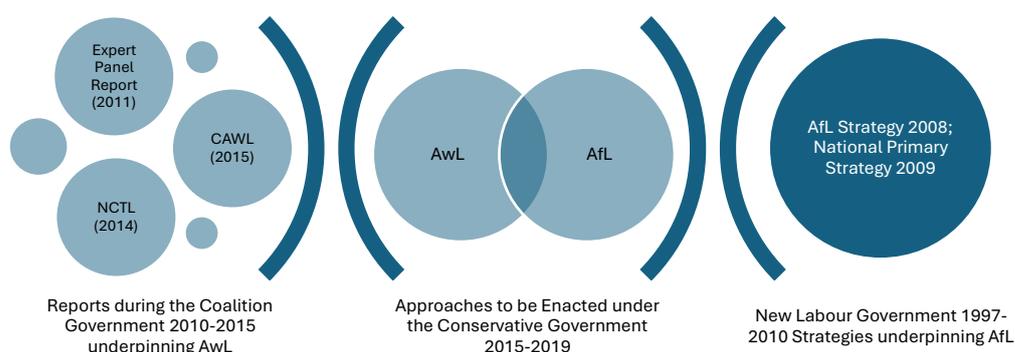
As it was not specifically an aim of AwL, it is necessary to consider that the enactment of AfL has been the product of policies and guidance that influenced teachers prior to AwL. The enactment of Assessment for Learning (AfL) was a priority for during the New Labour government (1997-2010) who launched an Assessment for Learning Strategy backed by £150m of government money over three years for the professional development of teachers in AfL. The idea behind the 'Making Good Progress' pilot was to introduce single level tests twice per year for teachers to use to check their own judgements and for financial rewards to schools that could show progress of two levels for pupils over a Key Stage on the National Standardised tests (SATs). The government argued that putting AfL together with pupil tracking would make formative and summative assessment complimentary,

'where the learners are in their learning, where they need to go and how best to get there had come to mean assessing pupils frequently according to national curriculum levels and sub-levels to track progress, setting new levels as targets and then working (somehow) to attain them' (James 2011:14).

Concerned about the variable enactment of AfL, in 2008, Ofsted conducted an evaluation of the Department of Children, Families and Services National Primary and Secondary Strategies' impact on the implementation of AfL, observing 232 lessons and finding only 10% to be outstanding, 40% were satisfactory and 11% were inadequate. In that report, Ofsted explained what good practice was and illustrated it through specific examples which

resemble many of the practices enacted in Schools Qualitative and Continuum. In these outstanding lessons, inspectors observed that teachers effectively communicated learning goals and used assessment data to set appropriate objectives, and pupils were well-informed about their learning targets and understood how to achieve them. Teachers frequently revisited these objectives, promoting self-assessment and aiding in the planning of future lessons. Skilled teachers posed well-crafted questions suited to pupils' understanding levels, allowing time for thoughtful responses and encouraging pupils to justify their answers. Ofsted (2008) gave the example of an infant school in a disadvantaged area where teachers used regular 'learning stops' for pupils to discuss their progress which resulted in significant improvement; and of a Year 4 class where the teacher managed pair and group work effectively to sequence photos and create captions, explaining their vocabulary choices, while the teacher facilitated discussions to improve vocabulary and reviewed learning multiple times in different ways. Ofsted (2008) also emphasised that good AfL included detailed marking of written work that informed pupils of their strengths and areas for improvement. They gave the example of a middle school that had a policy of regular written feedback consolidated by pupil reflection time on that feedback within class. Inspectors found that schools where classes were inadequate lacked leadership guidance and monitoring of lesson plans.

Figure 23: Policy Recommendations under the Coalition Government (2010-2015) have combined with Policies under the New Labour Government (1997-2010)



Ofsted's (2008) evaluation and the AfL Strategy informed the New Labour Government's Primary National Strategy (2009) which concentrated on the enactment of AfL in literacy and

mathematics. However, that Government was committed to the levels system and made substantial investments in pupil tracking and planning tools to assist educators in using national test results for monitoring and setting targets. They called this approach Assessing Pupil Progress<sup>16</sup> which provided multiple sub-levels for teachers to measure pupil attainment against prompting teachers to use frequent mini-summative assessments designed to enhance test performance. Although the intention was to assist the enactment of AfL, the focus on the levels system resulted in teachers conflating learning with performance, emphasizing frequent assessments to monitor progress and set new targets, rather than promoting genuine learning,

‘one can also argue that what was being promoted was no longer formative assessment, as part of pedagogy, for the purpose of enhancing real and lasting learning, but frequent mini-summative assessment to secure higher performance on tests to meet prescribed targets’ (James 2011:13).

These national strategies had the status of being guidance but compliance with them was judged by Ofsted inspectors. The aim was still to raise standards as measured by levels and assessment for learning continued to be seen as an instrument for this purpose. It is evident in the data presented in this thesis that the policies and training provided during the New Labour government informed and influenced the teaching practices of staff as they enacted AwL. Whereas at the start of this DPhil study, I had situated AwL as being sourced in Chapter 8 of the Expert Panel Report (2011), the Phase 2 data makes it clear that a large part of AwL involves the continued enactment of AfL which commenced approximately a decade previously. Before AwL several scholars argued that the accountability systems that governments in the UK and elsewhere used were preventing the successful consistent enactment of AfL. In their review, Hopfenbeck et al. (2013) observed that several scholars had concluded that high-stakes testing systems often conflict with the principles of Assessment for Learning (AfL), leading to a disconnect between AfL methods and the requirements of high-stakes testing and accountability systems (Kirton et al. 2007). This misalignment can cause tension with teachers’ values and result in feelings of deprofessionalization (Berryhill et al. 2009). Ball (2003) observed that the accountability mechanisms of a neoliberal system of public sector governance required ‘individual practitioners to organize themselves as a response to targets, indicators and evaluations’

---

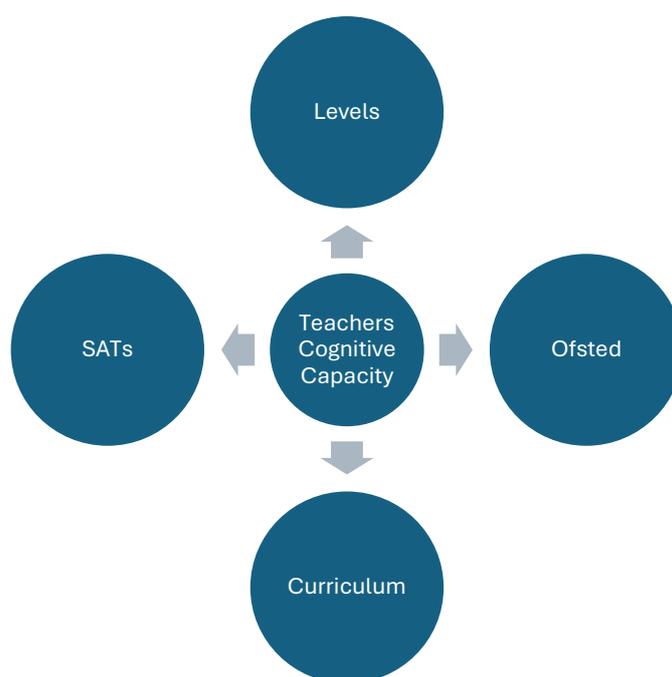
16

<https://webarchive.nationalarchives.gov.uk/ukgwa/20110203130903/http://nationalstrategies.standard.s.dcsf.gov.uk/node/20683>

and 'to set aside personal beliefs and commitments and live an existence of calculation' (Ball 2003:215). Wilson (2009) contended that the structure of state standards shaping the curriculum in the UK and USA has resulted in a dominant form of summative assessment. This dominance has overshadowed formative assessment, thereby becoming the primary influence on classroom instruction. Consequently, the combined roles of formative and summative assessments limit the curriculum and impede effective teaching. And Braun and Maguire (2018) titled their article about their interview study in two primary schools under the levels system as, 'Doing without believing – enacting policy in the English primary school'.

The data in this thesis suggests that it may not have been the national curriculum standards, accountability testing or inspecting that impeded the enactment of AfL but it may have been the requirement to calibrate with levels together with those accountability processes. Rather than conflating all of these, it is important to emphasise that in England each has its own criteria. The National Curriculum has statements which define what should be taught; Levels had separate criteria that teachers had to measure pupil progress on the curriculum against; SATs tested constructs not necessarily taught in the curriculum and so Year 2 and Year 6 teachers had to prepare pupils for that; and inspectors used Ofsted's guidance handbook and framework to make judgements. Whereas previous scholarship has tended to attribute the failure to enact AfL approaches to the pressure imposed by accountability process, this thesis posits that having four different criteria may have created a cognitive overload for teachers which made it difficult for teachers to then focus upon the enactment of AfL despite believing it to be pedagogically effective. This cognitive overload was caused by the need to interpret too many incongruent criteria which resulted in a lack of focus on the enactment of AfL – it was just one too many things to focus upon. Although many professionals may have advocated for the removal of SATs or Ofsted inspections, the Expert Panel Report's (2011) recommendation for the removal of levels seems to have been a savvy decision as it simplified the in-class tracking for teachers to the National curriculum statements that formed the basis of their lesson plan.

Figure 24: The Overload of Incongruent Criteria on Teachers Cognition



Braun and Maguire (2018) reported the stress that primary school teachers felt under levels due to multiple changes in assessment policies and strategies causing what they called pre-emptive enactment where teachers try to anticipate what Ofsted and the government may want them to do and they use their interpretation skills to enact an approach or policy. Cognitive overload occurs when the combined demands of external stimuli and internal thoughts surpass the available attentional resources (Sweller 1989). Therefore, cognitive load levels significantly impact both learning and performance in real-world settings (Goldinger et al. 2003). As schemas become more detailed and integrate more declarative knowledge into a single unit they provide clarity to the teacher. This was part of the rationale behind the Labour government's APP system that created many sub-levels, however, when it was combined with multiple schema that are incongruent it can become too complicated for teachers to be able to use continuously throughout the day. The demands of these multiple incongruent criteria seem to have diverted cognitive resources from one of the the main objectives of enacting AfL (Sweller 1989). This shift in mental resources creates "extraneous" cognitive load, reducing the working memory available for processing information to achieve desired outcomes (Chandler and Sweller 1991). Chandler and Sweller (1991) identified three types of cognitive load based on task characteristics: intrinsic load, extraneous load, and germane load. Intrinsic cognitive load refers to the working memory burden inherent in the lesson's content and pedagogy; extraneous load can include structural content like the criteria of levels, SATs and Ofsted criteria; and germane load

might include the awareness of students' prior knowledge, assessment of verbal and nonverbal comprehension, and monitoring students' attention levels. Any cognitive load not germane to the task reduces the cognitive resources available for target performance. It is these germane elements of teaching that AfL encourages and AwL has enabled the enactment of AfL by reducing the extraneous load on teachers by simplifying in-school pupil progress tracking.

## **Chapter 7: The Headteacher as Coordinator of Management Approaches towards Stakeholders**

### 7.1 Chapter Introduction

Chapter 7 presents an analysis of stakeholder influence in the initial interpretation of the AwL policy and also 2-3 years later after the chosen AwL system had been enacted which was co-produced with the school's headteacher. These stakeholders are not just leaders and teachers but in-line with Löfgren et al. (2018) and Wilkinson and Penney's (2021), pupils are also considered as policy actors in the processes of translating and enacting policy in school. The aim of Section A is to make the complexity of policy enactment process 'clearer' and capture the 'who' of stakeholder involvement in the process of policy enactment. Thus, the policy of AwL is used as a case to illustrate the process of (re-)interpretation, (re-)translation and (re-)enactment throughout the policy enactment process in a primary school. The data was generated through a hierarchical mapping technique (Antonucci 1986) used as part of the interviews with six headteachers between October 2017 and January 2018 to try and understand how headteachers created policy spaces and established a mobility for the policy of AwL from theorisation to enactment by engaging with school stakeholders. Rather than the participant being the ego, the policy of AwL was the central part of the map and headteachers were asked to place stakeholders according to their influence in the interpretation stage and also later after policy enactment. The hierarchical mapping technique was part of a semi-structured qualitative interview using a thinking-aloud approach that created visual and verbal data concurrently and both these types of data are presented. Analysis of this data reveals that the uncertainty that headteachers faced in the interpretation of what AwL should entail resulted in their reliance on expert stakeholders such as themselves and subject leaders. What was also evident is that headteachers in primary schools have the power to invite stakeholders and uninvite stakeholders and to coordinate them. A reflection on this dynamic of headteacher coordination of stakeholders is also presented. The second part of Ball et al.'s enactment framework is an understanding of the type of policy roles and policy work that different stakeholders play. One of the reasons that headteachers must distribute leadership is that together with tasks specific to job roles, teachers and sometimes other stakeholders must engage in a variety of policy work. In their analysis of the stakeholders involved in policy enactment within secondary schools, Ball et al. (2011) developed a typology to describe different policy actors and policy work. As Ball (2016:552) argues, 'it is necessary to identify and explore the paths and pipelines, and nodes and activities through which policy moves, and as McCann and Ward (2012:42) put it, this means both 'following policies and "studying through" the sites and situations of

policymaking' The data from this thesis is then used to reflect upon Ball et al.'s (2011) typology of policy actors and policy work and some alterations to that are suggested.

A socio-cultural view of policymaking looks at how different educational actors inside and outside schools play an active role in the final enactment of initial stated policies (Ricento and Hornberger 1996). From this standpoint, the entire policymaking process is conceptualised as social and situated practice (Levinson et al. 2009) and a highly political, dynamic, and unpredictable process in which initial policy texts interact with the individual and collective agency of the school actors. Ball et al. (2011, 2012) observed that stakeholders in schools are positioned differently and take up different positions in relation to policy, and Skerrit et al. (2021) found in the context of evaluation policies in Irish schools that stakeholders are not equal and are not working on and with policy in similar ways. Sutton and Levinson (2001) use the concept of policy appropriation which they define as a policy text or idea that 'is circulated across the various institutional contexts, where it may be applied, interpreted, and/or contested by a multiplicity of local actors (Sutton and Levinson 2001:2). In the last twenty years, there has been a shift in conceptualising educational policy making from the 'policy implementation' (Spillane et al., 2002) to 'policy enactment' (Ball et al., 2012) which like appropriation, captures the complex process, and contested nature of policy. Like Scanlon et al. (2022), the aim is to make the complexity of policy enactment process 'clearer' and capture the 'who' of stakeholder involvement in the process of policy enactment. By exploring 'the people, places, and moments' (Prince 2010) in the interpretation and enactment of a significant policy, an understanding of why it looks the way it does in a school can be gained. As Ball (2016:552) argues, 'it is necessary to identify and explore the paths and pipelines, and nodes and activities through which policy moves, and as McCann and Ward (2012:42) put it, this means both 'following policies and "studying through" the sites and situations of policymaking'. This chapter explores the enactment of a significant new policy in schools in England to not only *tell* readers about the variability in stakeholder involvement but to also *show* the reader what the looks like through a visual representation created by the headteachers themselves. The data illustrates how headteachers maintain their power to invite stakeholders and attempt to uninvite or uninvolve those who disagree with their approach, however, stakeholders like the school's Governing Body and Multi-Academy Trust are empowered by the law to invite themselves to be involved in the interpretation and enactment of significant policies like AwL. Most of the time, these stakeholders will work on defined roles/tasks that do not reach into the headteacher's roles but what the policy of AwL reveals is that these stakeholders will invite themselves into the headteacher's remit of classroom progress capture and reporting as it also impacts upon the functioning of their role.

## 7.2 Data from the Headteachers Stakeholders Maps

The first part of this chapter presents the data from the hierarchical mapping technique (Antonucci 1986) which were used as part of the interviews with six headteachers between October 2017 and January 2018 to try and understand how headteachers created policy spaces and established a mobility for the policy of AwL from theorisation to enactment by engaging with school stakeholders. Rather than the participant being the ego, the policy of AwL was the central part of the map and headteachers were asked to place stakeholders according to their influence in the interpretation stage and also later after policy enactment. The hierarchical mapping technique was part of a semi-structured qualitative interview using a thinking-aloud approach that created visual and verbal data concurrently these are presented in this section.

Figure 25 is a scan of Head-Continuum's interpretation map which depicts the headteacher working with one particular Assistant Head to design and develop their continuum AwL system with two other Assistant Heads and parents having some input. Figure 26 is Head-Continuum's enactment map which depicts three local authority advisors and a diocese school improvement advisor whose combined feedback resulted in some alteration to the AwL system that was enacted which was not desired by Head-Continuum and the Assistant Head who designed the system. The opinion of all of these external visitors that year group attainment expectations were necessary within an AwL system influenced school governors who insisted that that was enacted (the Post-it note with the school governors details is underneath the Post-it note for church school improvement leader on the map).

Figures 25 & 26: Head-Continuum's Policy Interpretation & Enactment Stage Maps

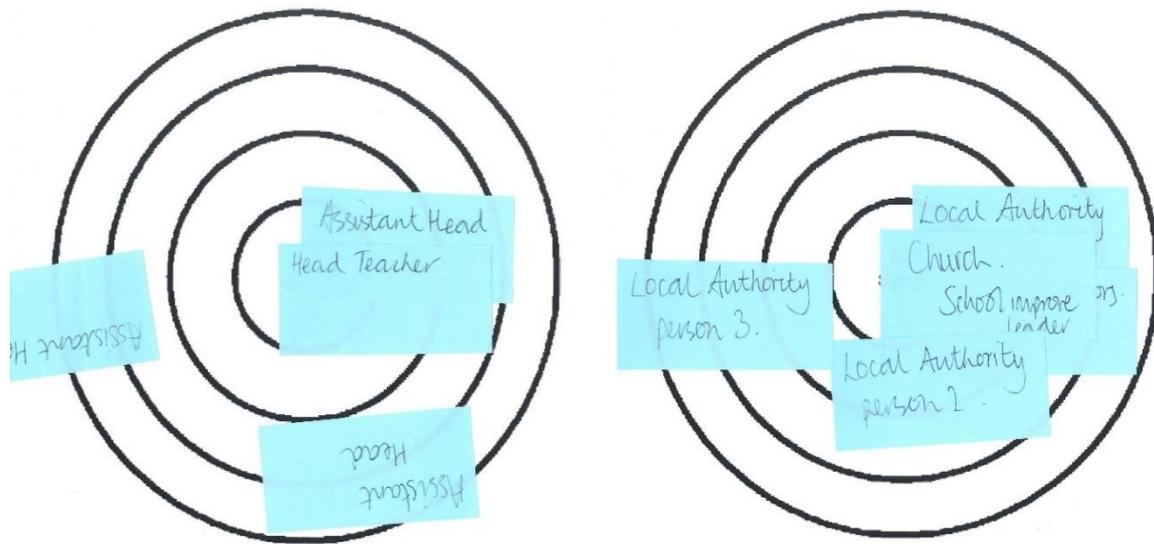


Figure 27 is Head-Qualitative's interpretation map and depicts Head-Qualitative's insistence the pupil voice was at the centre of shaping their policy together with teacher voice. They related how teachers interviewed pupils to glean what they felt should be the values behind an assessment system and the approaches to assessment which they find to be most useful. Further exploration during the interview prompted Head-Qualitative to place themselves centrally after having initially not wanted to include themselves in the map at all. School governors did not agree with the removal of classroom testing and numerical reporting and so were pushed out of the process and therefore feature on the edge of the map.

Figures 27 & 28: Head-Qualitative Policy Interpretation and Enactment Stage Maps

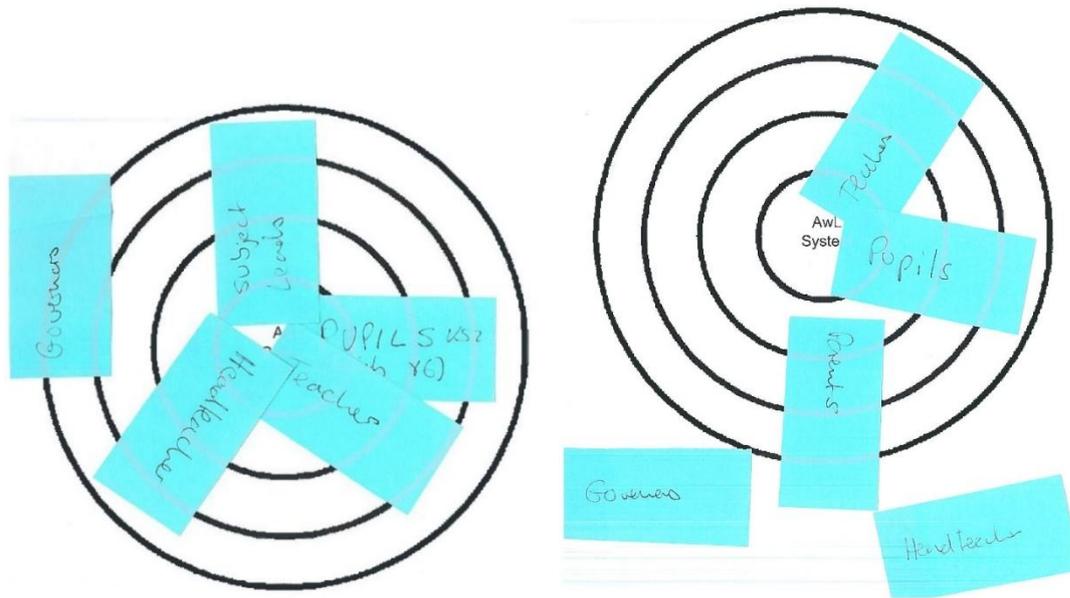
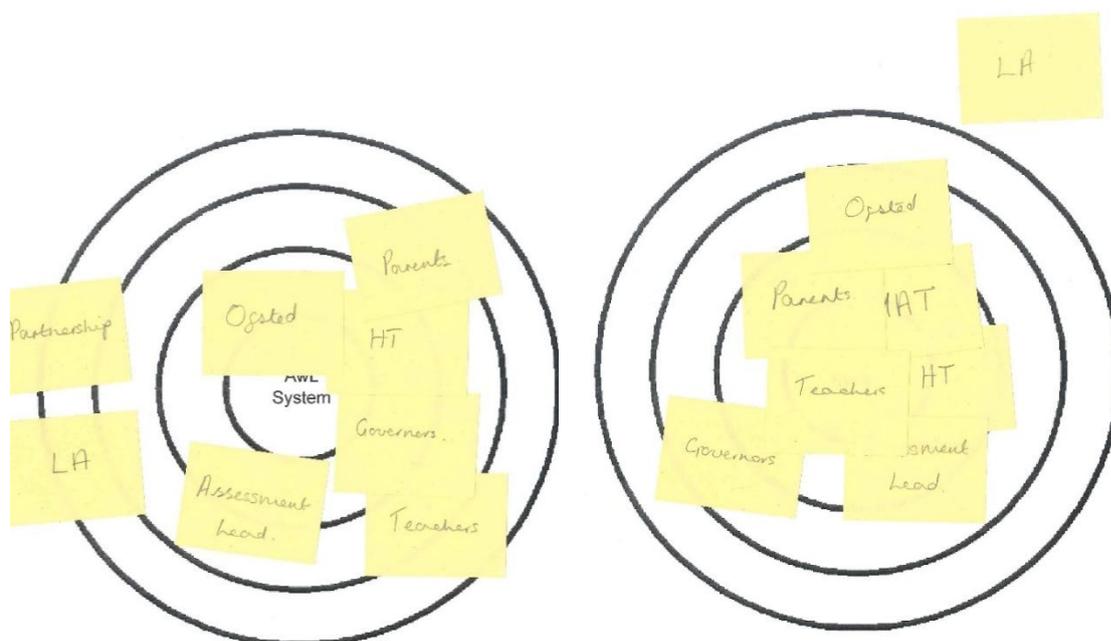


Figure 28 is Head-Qualitative's enactment map in-which they argued that the policy space was now the classroom and as teachers and pupils were the ones using it they were central with parents now also more central. Head-Qualitative placed themselves and school governors outside of the Figure as they were only involved in whole school reporting and analysis.

Figure 29 is a scan of Head-SPT's interpretation map which depicts that the headteacher and the school governors as central stakeholder as they chose the system. The other stakeholders in the Figure are psychologically present in their minds while choosing. Their most prominent consideration was what Ofsted inspectors may wish to see, and therefore, they are in the centre. The next consideration was for parental reporting and so they are depicted in the next ring. Teachers needs when using the system in the classroom were the next consideration, however, the schools partnership that they belong to which was becoming a Multi-Academy Trust (MAT) had very little input and neither did the local authority and so they are depicted on the edge of the map.

Figures 29 & 30: Head-SPT Policy Interpretation & Enactment Stage Maps



During the enactment phase, teachers and parents join Head-SPT in the centre, with school governors and Ofsted moving further out (Figure 30). Head-SPT explained that this was because they realised over time that Ofsted did not have a need for data in the way that they had initially anticipated they would need. Local authority involvement diminished after they formed the NAT who are now central as the body to which data is sent.

Figure 31 is a scan of Head-PATs interpretation map in Figure 32 is a scan of Head-PATs enactment map. Head-PAT explained that they took over the school when it was in special measures and had very few structures and processes in place and so initially Head-PAT and their deputy headteacher are central as they introduced a rudimentary assessment system after AwL was removed while they were transitioning to joining a Multi-Academy Trust. The school governing body did not exist and the local authority did not get involved because they knew the school would be joining a Multi-Academy Trust (MAT). This is the main change in the Enactment map as the MAT is depicted centrally because it provided the AwL system and headteacher and the leaders for inclusion and assessment coordinators are central within the enactment of that system. The local authority is even less involved after the school joined the MAT and the local governing board has been replaced by the formation of a School Strategy Board who receive information from the headteacher but are not involved in the formation and enactment of the AwL system.

Figures 31 & 32: Head-PAT Policy Interpretation Stage Map

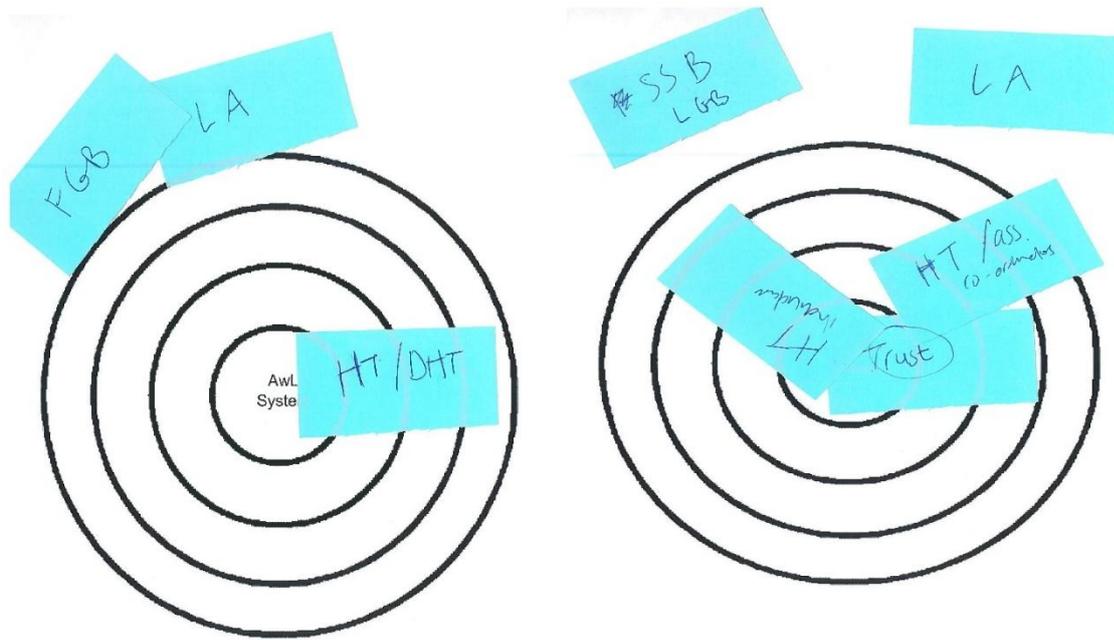


Figure 33 is a scan of Head-Points interpretation map and the Senior Leadership Team (SLT) are placed in the centre as one note as they collectively led the interpretation of the policy creation of the AwL system with other teaching staff involved one level out and non-teaching staff the children and governors being right on the outside with very little involvement. Figure 34 is Head-Points enactment map which depicts the prominence of the deputy headteacher who manages the AwL system across the school with the rest of the SLT now one ring out. Teachers, parents and school governors are now more involved and feature within the rings.

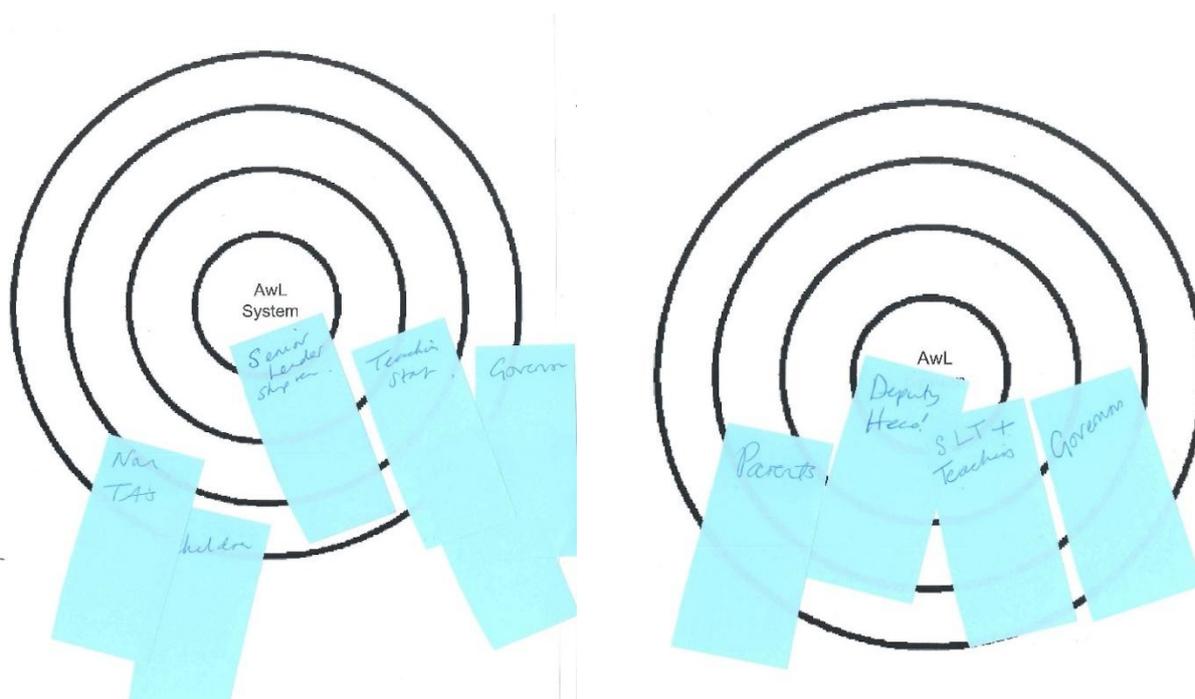
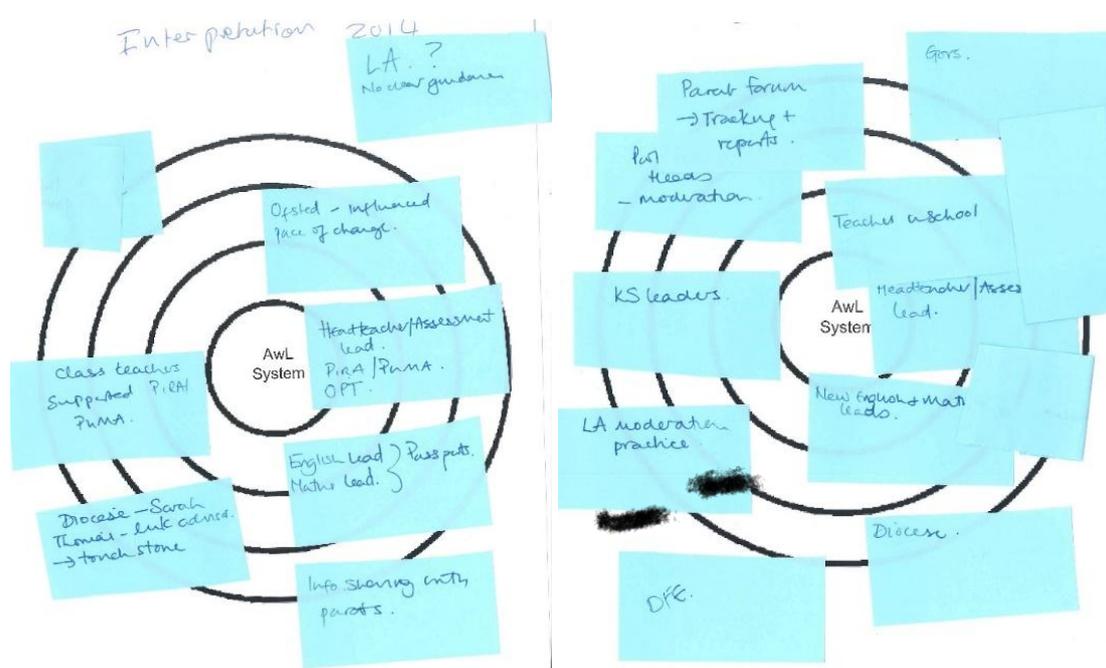


Figure 35 is a scan of Head-Passports interpretation map and Figure 36 is a scan of Head-Passports enactment map. Head-Passports included individuals, psychological pressures, permanent organisations, occasional gatherings, processes and tools as stakeholders that influence. In SchoolPassports the headteacher is the assessment lead and has been placed at the centre of the Enactment map together with some of the external testing tools that they have used to help benchmark the standards of the AwL system. English and the maths subject leads are one ring out from the headteacher and interestingly Ofsted inspectors are placed just outside them even though they did not visit the school as Head-Passports explained that anticipation of inspection drove the speed with which they enacted AwL. Then class teachers need to verify their judgements drove the purchase of in-class testing systems PUMA and PIRA and the diocese provided a mentor who Head-Passports used as a source of feedback. To emphasise that school governors and parents were not influential but that information was transmitted to them, they are placed on the outer ring. The local authority provided no guidance and so they are placed outside of the map.

<sup>17</sup> Head-Points populated both maps together and subsequently explained them together

Figures 35 & 36: Head-Passports Policy Interpretation & Enactment Stage Maps



Head-Passports Enactment Map maintains the headteacher at the centre but this time they are joined by the rest of the teaching team especially Key Stage Leads and newly created English and Maths Leads. The different AwL systems in each school have brought moderation into prominence both by the local authority and within the partnership of schools. The diocese mentor continues to support Head-Passports occasionally, and Parents and governors remain on the outside as consumers of information. Inspectors no longer have a practical or psychological impact and the DfE are depicted outside of the map to emphasise their lack of influence and involvement.

### 7.3 Interview Data explaining the Stakeholder Maps

The hierarchical maps above were used as part of the Phase 1 headteacher interviews as an aid for participants memory recall and to access their narratives on stakeholders relative importance and involvement and in this section quotes from that part of the interview are presented. Maguire et al. (2015) observed that the prominence of a policy in a secondary school is often determined by the relative power of the teacher that is responsible for it and that important policies would be led by powerful teachers. Primary schools are usually much smaller than secondary schools and so the headteacher leads on many areas of school life especially significant policy changes such as this. Also, AwL encouraged headteachers to create systems according to their assessment values and so in all six participant schools they owned and led the design of AwL, as Head-Qualitative explained:

'I suppose at the very initial Stages I was very much leading it to lead the change... I suppose I'd have to go in the middle because I've sort of tabs on everything as it's been designed... I've delegated it and I've never kept my eyes off of it because it's always been a key focus. So, therefore I've always been involved in a slightly different way' (Head-Qualitative).

Since the 1980s progress tracking approaches were provided by the DfE and so Head-Continuum explained that they relished the chance to design their own AwL system, 'we felt that this was a big opportunity and... I think we probably grabbed that with two hands' (Head-Continuum). This desire to shape the policy according to their assessment values resulted in Head-Continuum limiting the invitation to be involved in designing the AwL system to themselves and Assistant Headteachers who agreed with them,

'we had a very clear vision of what we wanted it to look like and we didn't want anybody to do what eventually happened which was to try and persuade us to do otherwise (smiles)' (Head-Continuum).

Although headteachers drove the interpretation of AwL and steered the process, Deputy and Assistant Headteachers were influential in turning ideas and ideals into a system that could record school-level pupil progress data implementing the ideas in the form of a system:

'actually without the one person who is on the post-it note (Deputy Head) who devised the Excel sheet, without her, it's pretty much nothing. So, all of the knowledge and information is contained in her brain' (Head-Continuum).

'my Deputy will take the lead in terms of developing a concrete system so on a practical level he is putting in a filtered system, he will analyse the system and then give it out to people' (Head-Points).

The relationship between headteachers and school Governing Body (GB) was significant for how AwL was interpreted and enacted. Head-SPT worked closely with governors to choose an AwL system that would meet whole-school data needs. Heads PAT and Passports invited governor involvement based on clear role boundaries, considering the design of an AwL system to be a task primarily for teaching staff, 'It's our decision - it is a management and leadership rather than the strategic overview of how it was going to happen' (Head-Passports). SchoolPoints benefitted from having assessment experts on the governing body

and so Head-Points worked with them to disseminate information and training to other governors,

'I'm lucky I've got some... my chair of governors is an Ofsted inspector and a headteacher himself, so he is very knowledgeable, and I've got someone who works at OUP and works in the assessment side, so she's very knowledgeable. So, they can also take early lead and actually skill up other people' (Head-Points).

The quotation above indicates the influence that governors with a teaching background have within the GB and how they can support the headteacher to help non-teaching governors to understand what the SLT are trying to achieve and the collegiate atmosphere that this can create between them and the headteacher. In Schools Qualitative and Continuum, the relationship between the SLT and GB was strained by the enactment of AwL as they disagreed with some parts of the approach that the headteacher was proposing. Head-Qualitative uninvited (Novak 2009) governors to be involved in the interpretation of AwL when they expressed a desire for regular test data to judge pupil progress, 'Governors were obviously involved, but they got pushed out of helping us behind the system because their idea of what they wanted did not align with our philosophy' (Head-Qualitative). A similar tension also developed in SchoolContinuum with Head-Continuum attempting to not involve the GB by exercising distinct role boundaries between the SLT and GB but they were compelled to incorporate recommendations from Local Authority Advisors due to concerns of anticipated poor inspection judgements,

'when the local authority person had been, the Chair of Governors would continue with that sort of line of questioning as to "why we don't do this one?" and "it's necessary for lots of reasons one for local authority scrutiny and another for Ofsted scrutiny because Ofsted expects" ... in the area that we are in, several local authority schools are in special measures or in some sort of classification, and that's created a bit of a furore' (Head-Continuum).

This quote illustrates how the GB can invite themselves into the interpretation process even when they are not invited to do so by the headteacher and that is likely to happen when they fear the judgement of Ofsted inspectors. This anticipation of Ofsted features in the interpretation map of Head-Passports who explained that it influenced the speed of the interpretation process within their school, 'I was likely to be receiving an Ofsted inspection and so that did galvanise me at that point to put something in place' (Head-Passports). It

also featured in the interpretation map of Head-SPT who commented, 'knowing that we would have to produce data on attainment in Year attainment was also a fairly significant driver towards that decision' (Head-SPT). Eventually, Ofsted did not visit either school when they thought that they would, and contrastingly, Head-Points was not surprised that Ofsted did not inspect them after the change to AwL due to their previous performance, 'We got 'Outstanding' so Ofsted haven't come again' (Head-Points). Head-Points did not place Ofsted on their maps because their inspection was in the first year of AwL and so they felt that 'Ofsted came too soon to understand', but they explained that parents' comments to inspectors had influenced the way in-which their AwL system reported to parents, 'That came out of our Ofsted report as to parents want to know more about how their children are progressing' (Head-Points). Parents inhabit an in-between space as they are outside of the classroom but connected with it daily through their children and informal conversations with teachers, and formally through feedback evenings and report cards. Head-SPT was the only headteacher that placed parental needs centrally when choosing School Pupil Tracker,

'The other appeal of the School Pupil Tracker was that it could be opened up for parents as well so that they could see progress and attainment that they hadn't been able to do with any of the other systems and wouldn't have been able to with the system that we had built internally' (Head-SPT).

Enactment phase maps show that parents had an increased presence in most settings as the AwL system was used in the classroom. This seems to be linked to the enactment of a stated AwL policy intention to help parents know the next step of learning for their child so that they can help them outside of school, 'On a termly basis we are providing parents with next steps, so they understand where the gaps are. So, in that way we are now (during the enactment phase) involving them much more' (Head-Qualitative).

Interpretation maps show some schools involved class teachers in the sensemaking of the new policy, 'I would say the teaching staff because they are the people on the ground using it ... they were involved there in the interpretation but now they've become more influential' (Head-Points). For example, Head-Passports explained that teachers collectively agreed the meaning of the standards for each Year group, 'we spent a lot of staff meetings looking at what was Emerging, Expected - what that might look like. We looked at exemplar materials, we looked at work, we used moderation of our own work samples' (Head-Passports). Contrastingly, Heads Continuum and PAT took a transmissional approach in training teachers to use the AwL system rather than involving them in the design of it, 'I would say for the initial developments we said to teachers that "here is how we are going to do

assessment” (Head-Continuum). The enactment maps convey that those teachers with specific responsibilities for assessment, numeracy, literacy and special educational needs and disabilities (SEND), were particularly influential in how AwL was enacted. Head-PAT explained that teachers were involved in the adjustments made to the AwL system after it had been used for over one year.

The placement of Multi-Academy Trusts (MATs) at the centre of the enactment maps of Schools SPT and PAT, conveys the changing landscape of school governance structures during this period. These two MATs were also structured differently from each other, as SchoolSPT’s MAT was created by a collaboration of local schools whereas SchoolPAT joined a MAT with schools spread across several counties. These MATs were also different in their approach to AwL as SchoolSPT’s MAT did not provide or proscribe an AwL system meaning that schools in the MAT interpreted AwL differently and were using different systems to each-other. This MAT invited themselves into the AwL process through requiring compliance with data generation requirements to enable moderation between schools using differing systems, ‘What has crept in, not crept in, waded in, who wasn’t there before, is the Multi-Academy Trust who are asking for data frequently and at quite a level of detail’ (Head-SPT). Head-SPT was of the view that their MAT would increase the standardisation and perhaps allocate one system for all their schools. SchoolPAT did not have a choice of AwL system as their MAT provided Pupil Asset Tracker to all schools together with a support structure to aid enactment which is why they are central in Head-PAT’s interpretation map. Head-PAT explained that while mandating the use of the same AwL system, their MAT permitted each school to choose which elements of Pupil Asset Tracker to use in their AwL system according to local preferences and needs, ‘there is a common language across all the schools, but there’s a really strong ethos that each school has its own individual needs... so I can still adapt it and make my own targets and use the system in a way that will best support me in my school’ (Head-PAT). In their Phase 2 interview, Head-PAT conveyed how they had stopped using many of the features of Pupil Asset Tracker in the year since their Phase 1 interview and tailored the system further to the needs of their context. The contrasting approaches of the MATs also convey the challenges that AwL poses between autonomy for individual schools and standardisation across the MAT particularly for the purposes of moderation. Only Head-Qualitative’s interpretation map placed pupils near the centre,

‘I would start with the children - we started questioning them about assessments and that was just general in-class assessment with Assessment for Learning. So, I would say the children and the teachers would definitely be the first ones in the

middle... It was Key Stage 2 (pupils), mainly Year 6 to begin with I have to say, because we were doing a lot of interviews with them, talking with them about what works best, and what they felt worked best' (Head-Qualitative).

In Phase 2 of data collection, a group discussion with the Year 6 pupils who had been involved in the conversations that Head-Qualitative is referring to several years as Year 1 pupils was conducted and they remembered discussing the values that should underpin their assessment.

#### 7.4 Chapter 7 Discussion: Stakeholder Analysis of AwL Enactment in Primary Schools

This chapter explores why AwL was enacted differently in each of the six participant primary schools, firstly by accessing headteacher narratives on the stakeholders that influenced the initial design of the AwL approach and also the stakeholders that shaped the AwL system after it was enacted; and secondly, by interrogating interview data from all participants to glean the influences on teacher policy sensemaking. Stakeholder analysis originated in organisational management practice as a process for providing insights into, and understanding of, the interactions between a project and its stakeholders (Grimble and Wellard 1997). It is usually used at the planning stage of a project to identify and prioritise stakeholders who can have an impact on project success and the type of influence individuals have and in what way they might be an asset or hindrance to achieving successful outcomes. The stakeholder analysis adopted in this thesis is retrospective to explore who were the influential and important stakeholders during the enactment of AwL (Kennon et al. 2009). In contrast to much of the policy interpretation literature that tends to consider stakeholders to be equal and working on and with policy in similar ways, Ball et al. (2011, 2012), found that stakeholders in schools are positioned differently and take up different positions in relation to policy. This was supported by Skerrit et al. (2021) found in the context of evaluation policies in Irish schools that stakeholders are not equal and are not working on and with policy in similar ways. They observed that some actors, such as those in senior leadership positions, have no choice but to comply with and enact certain policies, and Harvey (2015) found that school leaders can be more supportive and enthusiastic than teachers. Murphy (2019) presented vignettes of three teachers with one engaged in the policy, another engaged a little and a third did not have opportunities to engage. Some studies have also explored teacher motives for being involved. O'Brien et al. (2019) found that some viewed involvement with policy enactment as an opportunity for career

advancement, whereas Maguire et al. (2015) and Harrison et al.'s (2020) case study research in one particular school found that inexperienced teachers can be less involved in school decision-making than their more experienced colleagues due to a preference to focus upon their classroom practice. However, AwL in primary schools was such a significant policy that it had to involve most stakeholders. This thesis builds on this understanding of the variability in stakeholder involvement and adds the understanding of the centrality of the headteacher's coordination of which stakeholders should be involved in the policy enactment process and when and how they should be involved. Many studies on stakeholder analysis only examine one organisation at one point in time but this thesis presents data from six schools and from two snapshots of time in each, that is, twelve snapshots of stakeholder involvement in the enactment of the same policy. Tables 11 and 12 summarise the data presented in this chapter and Figures 37 to 42 present the comparative importance of stakeholders during the interpretation phase (2014-16) when AwL was first introduced and later after the AwL system was enacted (2018-19).

Table 11: Summary Table of all schools Interpretation Maps

School	Centre	Circle 1	Circle 2	Circle 3	Outside
<b>SchoolContinuum</b>	Head, Asst Head1	-	-	Asst Head2, Asst Head3	-
<b>SchoolQualitative</b>	Pupils, Teachers, Subject Leads, Headteacher	-	-	Governors	-
<b>SchoolSPT</b>	Governors, Headteacher, Ofsted	Parents, Assessment Lead	Teachers	Local Schools Partnership, Local Authority	-
<b>SchoolPAT</b>	Head, Deputy Head	-	-	-	Local Authority,
<b>SchoolPoints</b>	Senior Leadership Team	-	Teachers	Non-teaching staff, Governors, Pupils	-
<b>SchoolPassports</b>	Headteacher	-	Teachers, English & Maths Lead, Ofsted	Diocese, Local Schools Partnership, Parents	Local Authority

<b>School</b>	<b>Centre</b>	<b>Circle 1</b>	<b>Circle 2</b>	<b>Circle 3</b>	<b>Outside</b>
<b>SchoolContinuum</b>	Diocese, School Improvement Advisor, Local Authority Advisor 1,	Local Authority Advisor 2	Local Authority Advisor 3	-	-
<b>SchoolQualitative</b>	Pupils, Teachers	Parents	-	-	Headteacher, Governors
<b>SchoolSPT</b>	MAT, Head, Teachers, Parents	Ofsted, Assessment Lead	Governors	-	-
<b>SchoolPAT</b>	MAT, Head	Assessment Lead	-	-	School Strategy Board, Local Authority
<b>SchoolPoints</b>	Deputy Head	Senior Leadership Team, Teachers	Governors, Parents	-	-
<b>SchoolPassports</b>	Headteacher	Teachers, English & Maths Leads	Key Stage Leaders	Parents, Diocese, Local Authority , Local Schools Partnership, Pupils	Governors, DfE

Figure 37: SchoolContinuum Stakeholder Influence Chart

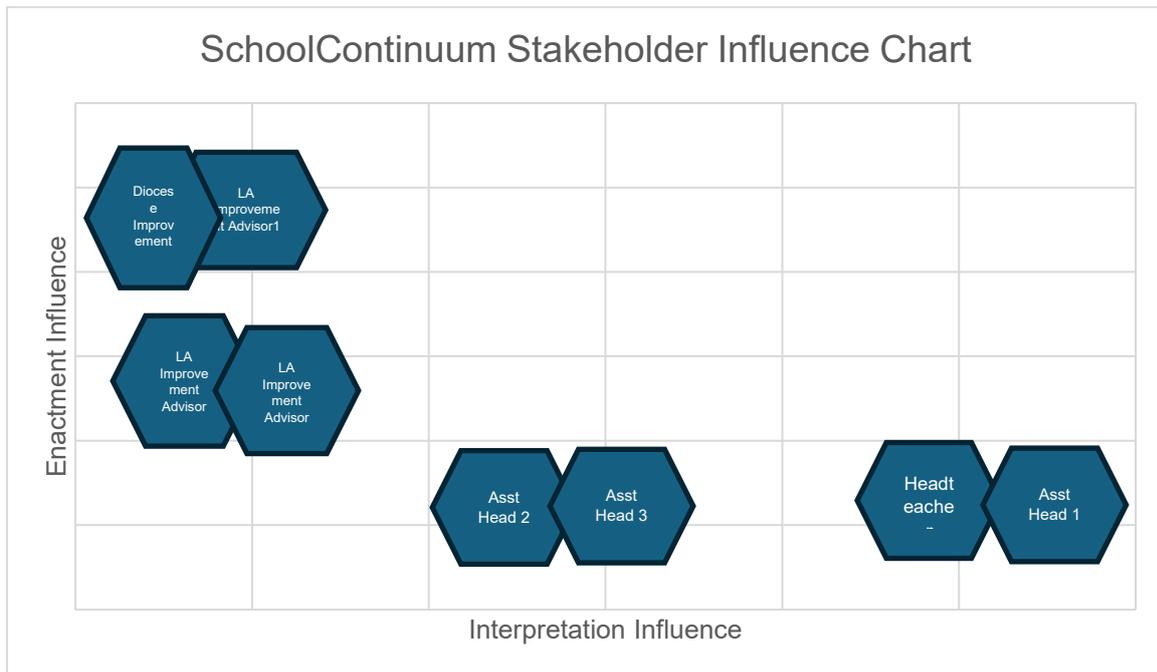


Figure 38: SchoolQualitative Stakeholder Influence Chart

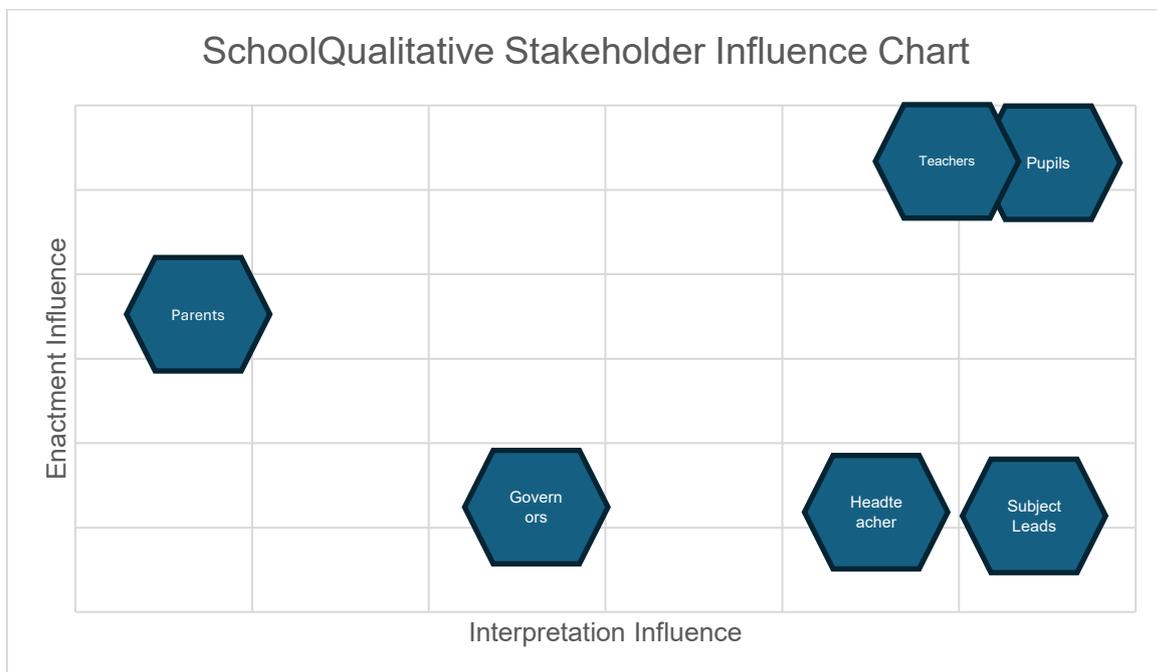


Figure 39: SchoolSPT Stakeholder Influence Chart

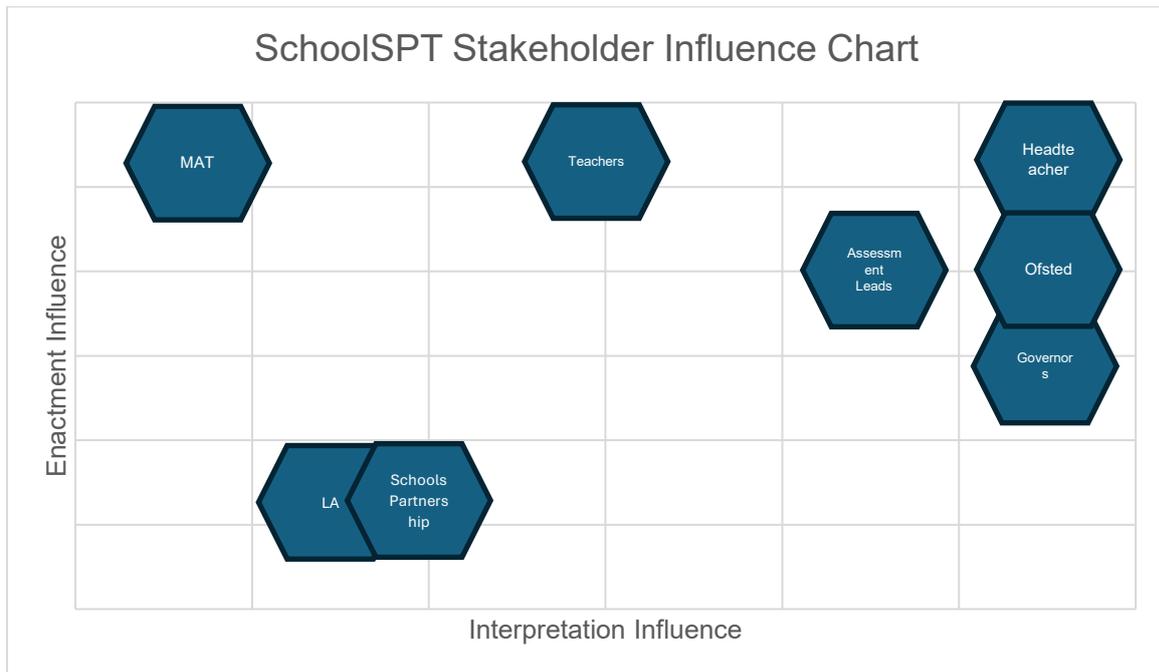


Figure 40: SchoolPAT Stakeholder Influence Chart

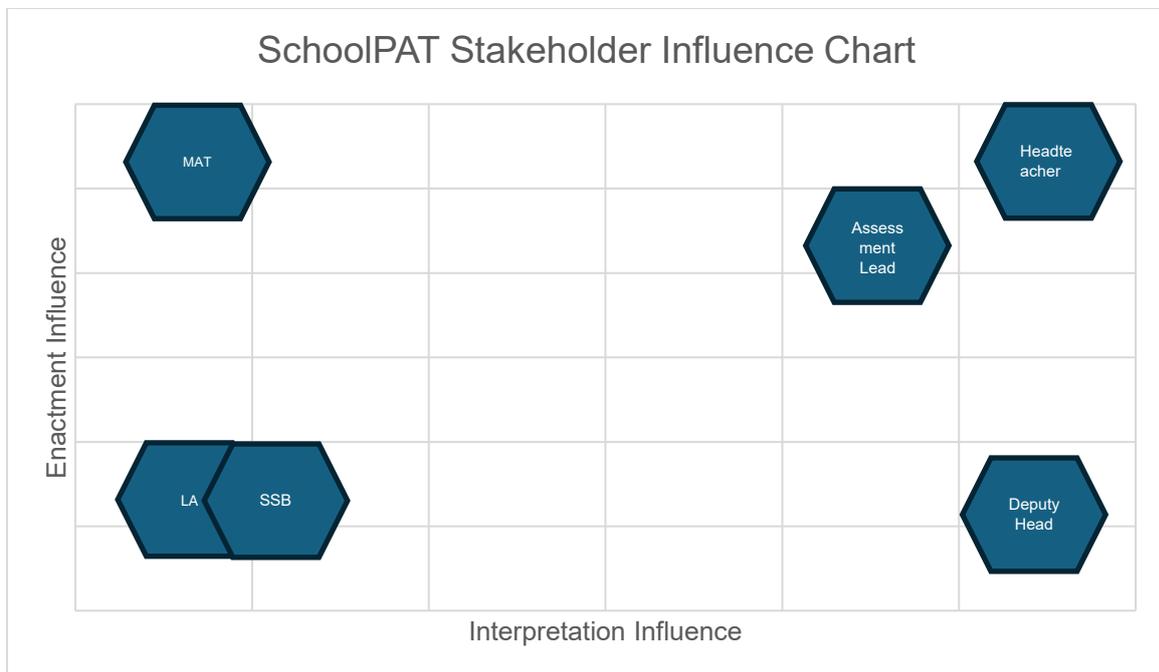


Figure 41: SchoolPoints Stakeholder Influence Chart

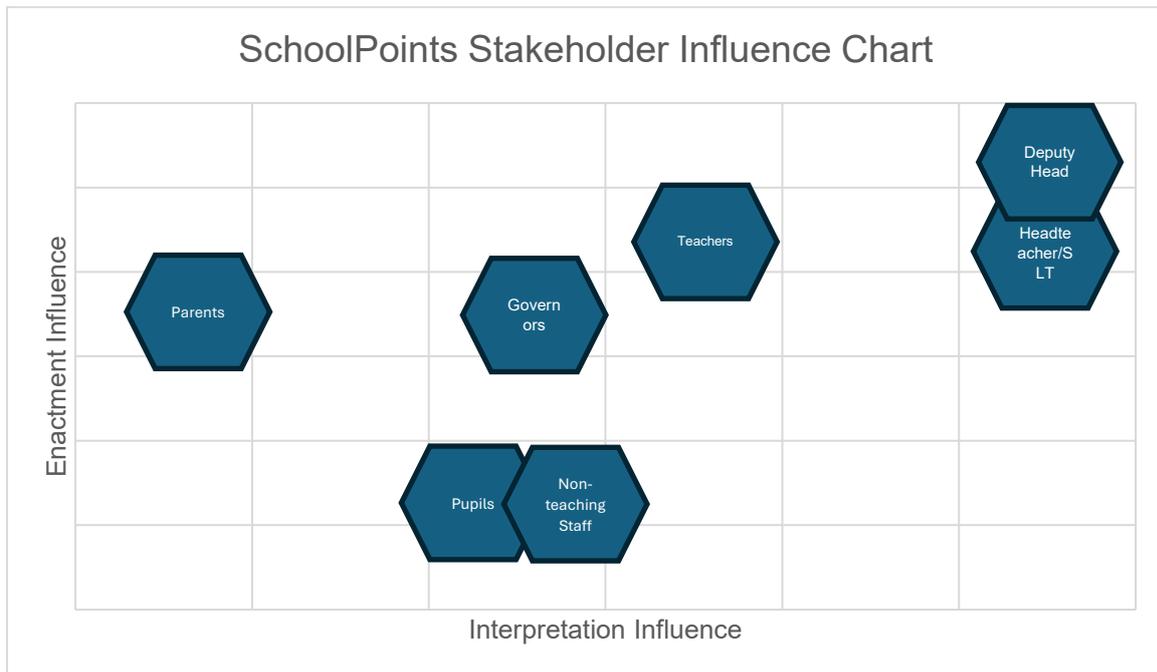
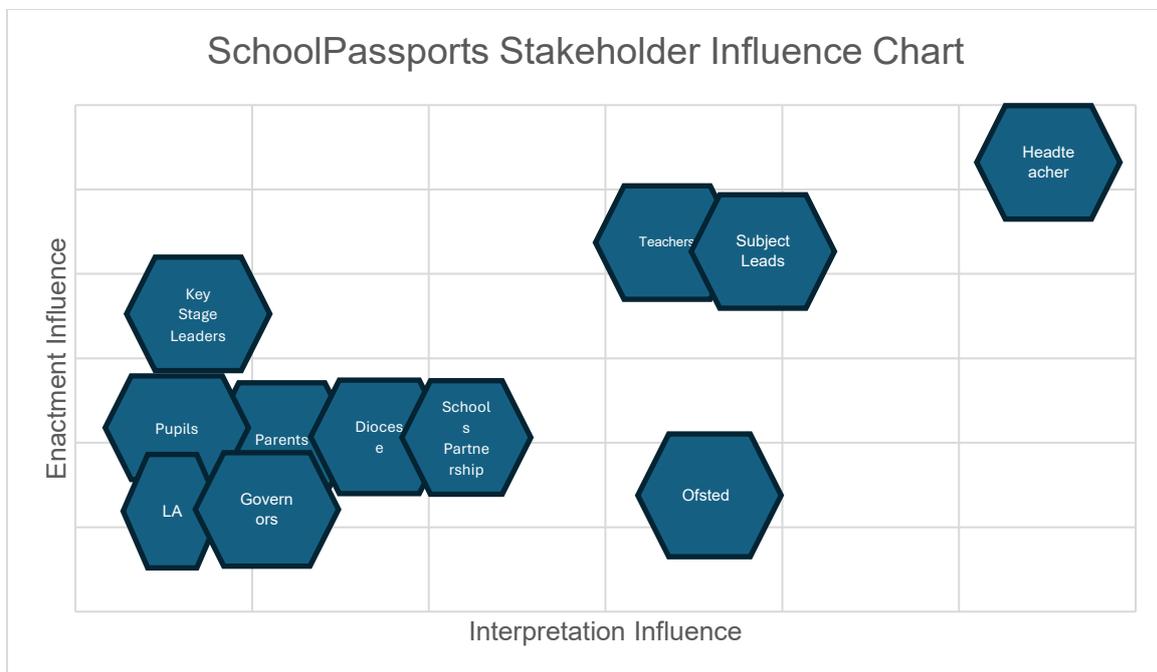


Figure 42: SchoolPassports Stakeholder Influence Chart



All schools have primary stakeholders who have legal and operational authority over the running of schools such as headteachers, governors and MATs, and secondary stakeholders who are not experts in teaching and assessment like pupils and parents who

'are diverse and include those who are not directly engaged in the organization's economic activities but are able to exert influence or are affected by the organization' (Savage et al. 1991:62). The tables and Figures above illustrate how the influence of these primary and secondary stakeholders changed over time differently in each location. A quick view of the hierarchical maps can give the impression that secondary stakeholders such as parents and pupils become equally influential as the policy is enacted but taking an observational approach to Figures 37-42 to aggregate the positions of the stakeholders in the hierarchical maps (Table 13) reveals the dominance of expert stakeholders in all six participant schools (Bijlsma et al. 2011).

Table 13: Aggregated Influence of the Stakeholders in each school

<b>Aggregated Influence Rank</b>	<b>Continuum</b>	<b>Qualitative</b>	<b>SPT</b>	<b>PAT</b>	<b>Points</b>	<b>Passports</b>
<b>1</b>	Headteacher, Asst Head1, Diocese Advisor, Local Authority Advisor 1	Teachers, Pupils	Headteacher	Headteacher	Deputy Head	Headteacher
<b>2</b>	Asst Heads 2 & 3, Local Authority Improvement Advisors 2 & 3	Subject Leads	Ofsted, Assessment Leads	Assessment Lead	Headteacher, Senior Leadership Team	Teachers, Subject Leads
<b>3</b>		Headteacher	Governors, Teachers	MAT, Deputy Head	Teachers	Key Stage Leaders, Diocese Advisor, School Partnership, Ofsted
<b>4</b>		Parents	MAT	SSIB, Local Authority	Governors	Pupils, Parents
<b>5</b>		Governors	Local Authority, Schools Partnership		Parents, Pupils, Non-teaching Staff	Local Authority, Governors

Chapters 4, 5 and 6 of this thesis explain and illustrate that the new curriculum based on mastery teaching and AwL involved many complex concepts that needed experts to interpret what to do and how to do it (Anderson and Taner 2022). Further, the stakeholders presented in this DPhil study have had to negotiate the uncertainty of the significant change to teaching and assessment that AwL introduced in a short time with very little guidance other than for headteachers to commence the interpretation with their assessment values. Koppenjan and Klijn (2004), distinguished between substantive, strategic, and institutional uncertainty. They

defined substantive uncertainty as a lack of knowledge about the subject matter either because there is no information or because the available information is too abundant or is ambiguous or conflicting. Strategic uncertainty refers to a lack of knowledge about how stakeholders will anticipate and respond to each other's actions; and they refer to institutional uncertainty as a lack of knowledge about procedures, processes and priorities in the different organisations of each stakeholder. The interview data presented in this chapter from the six headteachers alludes to all three types of uncertainty being present in the enactment of AwL and so teachers have had to use their expertise to bring clarity to the enactment of AwL. While the data reveals that the interpretation of AwL was concentrated in the hands of 1-2 experts in some schools, it also shows that power was distributed among expert teachers. This sharing of power among the teaching team is also indicative of the typical organisational culture in primary schools. Hammersley-Fletcher and Strain (2011) explain that distributed leadership was adapted from notions of 'distributive' leadership by the National College for Teaching and Leadership of Schools. 'In consequence former expectations of one leader taking full responsibility were replaced by arrangements whereby a distributed leadership could facilitate staff working together, taking responsibility for particular initiatives and periodically working as a leader for limited periods and purposes' (Hammersley-Fletcher and Strain 2011:873). Grint (2005) explains that this model of leadership draws on the unique skills of people within the organisation and allows them to act flexibly, taking on leadership and working on tasks as and when appropriate, and sometimes reverting to a non-leadership role when the task is complete.

## 7.5 Chapter 7 Discussion: Policy Actors and Policy Work

One of the reasons that headteachers must distribute leadership is that together with tasks specific to job roles, teachers and sometimes other stakeholders must engage in a variety of policy work. In their analysis of the stakeholders involved in policy enactment within secondary schools, Ball et al. (2011) developed a typology to describe different policy actors and policy work (Table 14). They labelled the first type of actor as 'narrators' who are usually school principals or senior leaders who interpret policy and plan what needs to be done. Skerit et al. (2021) observed that an important part of this role is the creation of institutional narratives about how their school interacts with policy, and how the school operates and functions, and how they strive for improvement. The second type of actor are 'entrepreneurs' who champion a policy and convince others to perceive it their way. The third type are outsiders as although the majority of policy actors are based inside schools, some outsiders such as consultants, advisers and trainers are also involved in policy. Skerit et al. (2021) point out that teachers and governors often value the interpretations of external

professionals and adjust their own policy translations to accommodate their views and this is supported by the experiences of Head-Continuum in this study. The fourth type of policy actor are called ‘transactors’ who account, report and monitor work according to targets and procedures. The fifth type are called ‘enthusiasts’ who are both engaged with a policy themselves and also encourage conversations on the potential of a policy in both informal and formal meetings. The sixth type are called ‘translators’ who create who enact a policy. The seventh type are ‘critics’ who scrutinise the enactment of policies and provide alternative ideas when a policy is being designed and also subsequently when the implications of the enactment become clearer. The eighth type are ‘receivers’ who are more junior policy actors that are not expected to interpret a policy or translate the policy into their own tools or activities but have this done for them by more senior members.

Table 14: Ball et al. (2011:626) Typology of Policy Actors and Policy Work in a School

<b>Policy Actors</b>	<b>Policy Work</b>
Narrators	Interpretation, selection and enforcement of meanings, mainly done by headteachers and the SLT
Entrepreneurs	Advocacy, creativity and integration
Outsiders	Entrepreneurship, partnership and monitoring
Transactors	Accounting, reporting, monitoring
Enthusiasts	Investment, creativity, satisfaction and career
Translators	Production of texts, artefacts and events
Critics	Union representatives: monitoring of management, maintaining counter-discourses
Receivers	Mainly junior teachers and teaching assistants: coping, defending and dependency

Skerrit et al. (2021) utilised Ball et al.’s (2011) typology to analyse their interview data on secondary school self-evaluation in Ireland and found that it was shaped by various, mainly internal, stakeholders. In line with Maguire et al. (2015) they also found that the SLT led on policy work and experienced teachers were more confident than junior teachers to discuss the policy. They commented that,

‘entrepreneurs are not necessarily found in every school while critics can be marginal (Ball et al. 2012) and outsiders sparse, but we did find strong evidence in particular of narrators, transactors, enthusiasts and translators, as well as receivers’ (Skerrit et al. 2021:17).

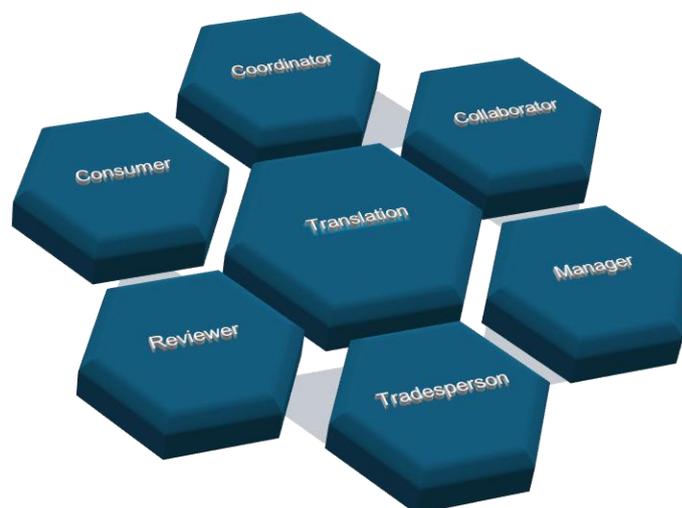
Ball et al.’s (2011) typology is useful when trying to understand the policy work that different stakeholders engage with, however, the data in this thesis suggest that there are some

limitations. Firstly, several policy actors and policy work should be merged, for example, narrators and translators do the same thing as they interpret policy and create systems, tools and activities that realise the policy in a practical form. For example, Head-SPT and School Governors interpreted the AwL policy and purchased School Pupil Tracker which they then trained staff to use; and in SchoolPoints, the SLT developed their Points system and also purchased maths problems packages which they were embedding into teaching across the school. Entrepreneurs and enthusiasts are also very similar as the former just seem to be the initial enthusiast who encourages others to become enthusiastic. For the enthusiast role the word enthuse should be added to policy work and career should be revisited because the data in this thesis does not convey that enacting AwL is motivated by career development and it is suggested that Ball et al. (2011) were too focused on teachers on the Future Leaders programme who were focused on career development more generally, rather than examining enthusiasm displayed by all teachers in their school. Narration, entrepreneurship and advocacy are also very similar roles and should be simplified and merged. For receivers, Ball et al.'s (2011) typology uses the language of weakness such as, coping, defending and dependency because they perceive junior teachers and teaching assistants as being continually overwhelmed by their job and incapable of being involved in policy-making. The data in this thesis, challenges this within the primary school context as all headteachers regarded teachers to be influential in either the interpretation or enactment phases and some placed them in both. The other main receivers of AwL are pupils and parents who are also influential in shaping some parts of AwL, such as the report card in SchoolPoints or the values behind the AwL system in SchoolQualitative. In all their work on policy enactment, Ball, Braun and Maguire give insufficient attention to the role of pupils and parents in policy sensemaking and so this thesis adds this to our understanding of enactment in schools.

Ball et al.'s (2011) typology also suggests that policy roles are played by separate policy actors, however, this is almost never the case with the same actor playing several roles and this is conveyed using the heading 'policy roles' rather than 'policy actors'. For example, Head-Continuum and Assistant Head-Continuum 1 relished the opportunity to be translators and design their AwL Continuum which they then promoted to school governors and enthused staff (receivers) to use it. Assistant Head-Continuum 1 then became the transactor by developing an assessment data capture and analysis system that they ran to monitor progress and attainment and provide teachers with information on the gaps in understanding that their pupils had. In SchoolContinuum, the main critics were also outsiders in the form of three Local Authority Advisors who all recommended that Year group expectations were used as a benchmark for judging pupil progress and attainment rather than Key Stage

expectations. Another problem with Ball et al.'s (2011) typology is the presentation of policy actors and policy work in specific rows and columns which suggests that certain types of policy work is the domain of certain policy actors. Further, the term policy actor suggests that it is an identity rather than depicting that policy actors (stakeholders) play several roles and do several types of policy work sometimes simultaneously. This error is probably due partly to their research being almost entirely in secondary schools that are usually bigger and require stakeholders to have more defined roles; and partly because they did not explore how stakeholders roles change over time as a policy is enacted this thesis does. The data in this thesis suggests that the term policy actor refers to a stakeholder and actually what Ball et al. (2011) are referring to in the left column of their table should be termed 'policy roles' that stakeholders play to conduct tasks in the right column which they correctly term as policy work. This policy work is done at different times by different stakeholders or some stakeholders may do multiple policy work at the same time and so this policy work should not allocated to particular roles as Ball et al. (2011) have done. Reflecting on the data presented above from this thesis, it is accepted that the policy work listed by Ball et al. (2011) in the right column of their typology is correct for primary schools too.

Figure 43: A Typology of Policy Work Roles in the Translation of a Policy



Secondly, labelling one type of actor as a translator seems to contradict the theoretical definition of policy translation. Ball et al. (2011, 2012) themselves define translation as the process of putting policies into action and assigning them symbolic value (Ball et al. 2011, 2012) and so all stakeholders that are involved with interpretation and enactment of a policy

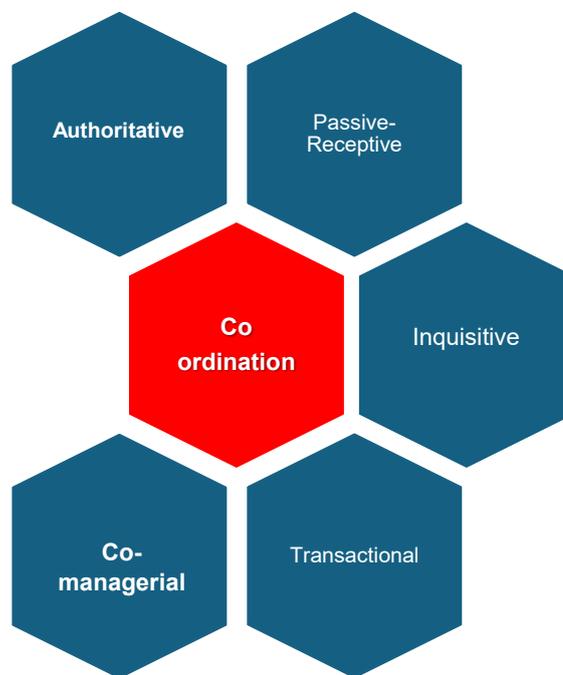
are part of the process of translating it. Further, to have translated a policy means to have enacted it and therefore, together with being a term to describe the overall process of policy sensemaking, translation is also a policy objective, that is, policy work aims to translate a policy idea or instruction into a practical useful form. The data in this thesis conveys the presence of a policy enactment coordinator who is usually the headteacher in a primary school. In all schools, this coordination commenced with selecting collaborators among the stakeholders who worked together to interpret the policy and design the new AwL system, such as, Head-Continuum working with one of their Assistant Heads, and Head-Qualitative involved pupils, teachers and subject leads. After the AwL system was enacted, they were sometimes managed by Assistant Headteachers as with Schools Points and Continuum or the SEND Coordinator as with SchoolPAT. These roles often involved design and running of pupil progress and attainment data systems that for the use of teachers to plan lessons and support for all students, and also for stakeholders who had roles connected to holding the school accountable that can be termed reviewers. A combination of school Governing Boards (GB), local authority advisors, and MATs played the role of reviewers together with an anticipation of what Ofsted inspectors may want to see. For example, Head-SPT commenced their sensemaking conscious of the needs of the GB and these two stakeholders collaborated to envisage what Ofsted inspectors may need. In their enactment map, Head-SPT explained that the formation of a MAT had resulted in that stakeholder becoming a reviewer with needs for data reporting. Ultimately, AwL progress data is consumed by teachers so that they know which topics need reinforcing, and by pupils and parents to know what their next step in learning is. Figure 43 offers an alternative representation of policy roles in school enactment with translation as the central focus and coordination as the initial role together with some of Ball et al.'s (2011) roles condensed into a fewer number of roles.

## 7.6 Chapter 7 Discussion: The Headteacher as Coordinator of Management Approaches towards Stakeholders

Ball et al. (2011, 2012) observed that stakeholders in schools are positioned differently and take up different positions in relation to policy; and in the context of evaluation policies in Irish schools, Skerrit et al. (2021) found that stakeholders are not equal to each-other, and work on, and with, the same policy differently. This chapter builds on this understanding of the variability in stakeholder involvement and adds the understanding of the centrality of the headteacher's coordination of which stakeholders should be involved in the policy enactment process and when and how they should be involved. This chapter presents data from six schools in two snapshots of time in each, that is, twelve snapshots of stakeholder

involvement in the enactment of the same policy. These snapshots help tell the story of headteachers adopting different approaches to the management of different stakeholders during the initial interpretation of a new policy and also 3-4 years later. In their study of deer park management in New York, Chase et al. (2000) developed a typology to explain the management approaches that wildlife officers took when involving stakeholders and the typology is useful for understanding how headteachers in this study have engaged stakeholders. Figure 44 uses the management approaches that Chase et al. (2000) list and after reflecting upon the data presented above,

Figure 44: Management Approaches of School Stakeholders during Policy Enactment (Adapted from Chase et al.'s (2000) management approaches towards stakeholders)



Chase et al. (2000) explain that the authoritative approach occurs where leaders prioritise their own expertise and make all decisions and taking all actions without consulting stakeholders, much like Head-Continuum and Head-PAT did. Professionals' interpretations of their mandate together with their expertise and knowledge about stakeholders' beliefs, attitudes, and opinions are used to make judgements in both schools. In the passive-receptive approach, managers do not actively seek stakeholder input or invite stakeholder input but are open to it when stakeholders take the initiative to be heard. Contrastingly, the inquisitive approach occurs when managers actively invite input from stakeholders, such as Head-Qualitative organising group discussions with pupils to explore the values that should underpin their assessment system. A transactional approach affords much greater power to

stakeholders compared to the previous approaches, allowing them to become involved in making decisions, rather than merely supplying input, such as Head-SPT deciding to purchase School Pupil Tracker based on conversations with Governors and parents. Chase et al. (2000) explain that co-management differs from the transactional approach by allowing stakeholders to be involved in multiple stages of the management process from goal setting to evaluation and not just in decision-making. For the enactment of AwL, the data suggests that the only stakeholders involved in co-management were the Senior Leadership Team, and as time proceeded, the new entrants into co-management of AwL in Schools SPT and PAT were the Multi-Academy Trusts. Reflecting on data from both the initial interpretation phase and the later enactment phase it is evident that headteachers employed a combination of these management approaches over that period and how they did that differed in each school and they are content with employing a fluid approach to this. The multiple management approaches used by each headteacher are consistent with Hammersley-Fletcher and Strain. (2011) who found that in a primary school the role of middle leadership 'is still fluid, reforming and arguably seeking to give renewed emphasis to teaching, learning and management structures' that effectively support each individual school in a manner that is needed at that time.

## Chapter 8: Insights into the Wider Paradigm of Policy Enactment in Schools

### 8.1 Introduction to Chapter 8

This thesis has sought to document differences and synthesise similarities on the enactment of AwL in the six participant primary schools and to treat this as an indicative case of how schools respond to and enact a significant policy that reforms aspects of schooling such as curricula, pedagogy and assessment. This chapter begins by making the case that this study has been successful in operationalising the research design and then proceeds to add to the scholarship on sensemaking and policy enactment. To aid an understanding of sensemaking, it is argued that headteachers together with other stakeholders such as teachers and the school's Governing Body, have their interpretations framed by three influences: the social, which includes factors such as colleagues' views and opinions shared in teaching communities; the headteachers' own values; and the accountability processes that pupils and schools are subjected to. A heuristic of 'Frames of Sensemaking' is offered. Based on the data generated in the six participant schools in this study, this thesis concludes that AwL is a case of successful policy enactment as all of the policy intentions identified in Chapter 2 seem to have been implemented. This seems to be because of the clear communication of policy intentions, together with the latitude offered to school leaders to design their AwL systems, coupled with teachers experiencing pedagogical improvements in the classroom, and the strong national accountability processes being maintained. A heuristic capturing these four aspects is offered.

### 8.2 Benefits of a Single Critical Paradigmatic Case Study Approach

This thesis utilises in-depth interview data generation with multiple participants and provides a detailed description and analyses themes inductively before connecting them with theory (Cresswell 2013). Using the enactment of the policy of AwL as the single case study with the six participant primary schools as sub-units that are located within a larger case (Yin 1989) has provided a nuanced exploration of the policy, the participant schools and of some of the main themes in the reform of pedagogy and assessment in schooling in England during the first two decades of the 21<sup>st</sup> century. The research design has preserved the benefits of conducting a single case study while also gaining the benefits of a multi-case study. Gustafsson (2017) argued that the focus that a single case study provides elicits a deeper understanding of the case and assists in the questioning of old theoretical relationships and the exploration of new ones; and Gerring (2004) posited that case studies have a double function as they are a study both of an individual unit and also a larger group of units. In this

DPhil study, the individual unit is the exploration of the enactment of AwL but rather than follow the policy enactment process broadly across the many locations as Poet et al. (2018) helpfully did, this DPhil has used a group of just six units in the form of six primary schools to take a deep-dive into how the enactment of a policy that reformed in-school assessment practices and pedagogy and an analysis of the data both within each situation and across situations has been presented (Yin 1989). This has enabled this study to have many of the cited benefits of a multiple case study design such as the opportunity to contrast similarities and differences in the data (Vannoni 2014), and an exploration of the research questions in multiple contexts (Eisenhardt and Graebner 2007).

The research design has achieved this breadth and has used it to aid the deep exploration of the enactment of AwL rather than having to compromise on either. This thesis conveys the locatability of individual voices and specific details that are a strength of a single case study and presents observable 'concrete interconnections between actual properties and people within an actual concrete setting' (Stoecker 1991:95). The thick description that qualitative studies promise has been achieved through the closeness of the case study to real-life school situations and some of the considerable wealth of details that these contain has been conveyed particularly about AwL and AfL. This thesis immerses the reader in the technical details of AwL and the stories and experiences surrounding enactment because the researcher has immersed themselves both in the topic and in two of the school contexts. Giddens (1982) claimed that it was such involvement in the social context that produced valid descriptions and advanced understandings, which are enhanced by the ability of participants to clarify meanings and correct understandings during the interview process creating 'mutual knowledge', that is shared by the researcher and participants (Giddens 1982:15).

This process has been enhanced by the use of the dynamic of time into the research design. By collecting data at two points or waves (Audulv et al. 2022) in the enactment process, this DPhil study has captured some of the benefits associated with a longitudinal study by exploring how people's experiences and perspectives evolve or stay the same over time (Hermanowicz 2016). Much of the data collection, particularly with headteachers, sought to access memories of the interpretation of AwL when the policy was introduced and in the early stages of implementation together with an exploration of the enactment at the time of the interview and considerations for future enhancements and so it had a retrospective-prospective element (Neale 2021). Adding time as part of the research approach has aided an understanding of transitions across the stages of a policy moving from an initial interpretation and early enactment stage to a more embedded stage, and how individuals

adapt to changes and it helps to assess the success of the implementation of the new policies. Young, Savola, and Phelps (1991) defined Qualitative Longitudinal Research as involving at least two waves of data collection over a year and so this DPhil study could be classified as such. This approach has enabled an exploration of the nature and intensity of the change processes, whether one-off or recurring (Neale 2021).

Ultimately, the key decision in case study research is the choice of case study. Mitchell (1983:203-4) asserted that a case should be chosen for 'Its explanatory power rather than for its typicality', but this thesis has achieved both. Using Flyberg's (2006) test of 'a case of strategic importance', the enactment of AwL is a critical case, and this is supported by using his 'most likely' test, as the sub-units of the case, that is, the six participant schools are indicators of what was happening in other primary schools in England between 2014 and 2019. These sub-units have made the analysis in this single case study more convincing as it grounds it in several sources of empirical evidence and tells the complex, sometimes conflicting stories, of situational actors in real-world primary school scenarios. The case is also a paradigmatic case study as the enactment of AwL has reformed assessment approaches and embedded AfL into pedagogy in schools across the country and the analyses presented attributes to national accountability processes a significant influence upon how AwL was enacted. Huitt (2018) argued that a paradigm may be thought of as a model of how processes are structured and also how they function in a particular context or time and Capra (1996:6) defined a paradigm as 'a constellation of concepts, values, perceptions and practices shared by a community, which forms a particular vision of reality that is the basis of the way a community organizes itself'. Thus, a paradigm can be seen as a combination of practices and the world views of important stakeholders. May (2001) emphasises the need to examine practices particularly habitual one and patterns of practices in order to determine what a paradigm is as one's actions say more about one's true self. Huitt (2018) felt that accountability practices were particularly important in understanding a paradigm as they drive curriculum choices and pedagogy. He was particularly concerned that this paradigm used to create the desired outcomes could lead to the selection of different goals and objectives than those actually needed for success in a particular location or time frame. The other part of the school paradigm is the worldview of the stakeholders, particularly the headteacher's which is a set of constructed perceptions and ideas about how the world works (Huitt 2018) or what Hatcher calls a 'map of reality' (Hatcher 1990:22). According to Aerts et al. (1994), a worldview is a "coherent collection of concepts and theorems that must allow us to construct a global image of the world, and in this way to understand as many elements of our experience as possible." This construct provides a

frame of reference that guides one's understanding of reality and provides the foundation by which one gives meaning to experiences and thoughts. One's worldview paradigm is used to successfully navigate through the challenges and obstacles of one's professional life and guides one's interpretation of the possible structure and functioning of both visible and invisible reality (Thulasidas 2008).

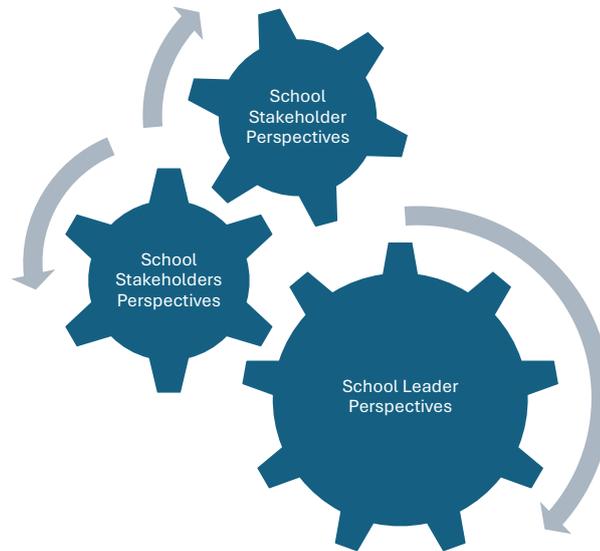
## 8.2 Frames of Sensemaking

This sensemaking does not exist within a vacuum but is influenced by factors within the stakeholder and factors within their environment. To fully grasp why a policy is enacted in a particular way it is necessary to explore both school leader and school stakeholders perspectives and then analyse how these perspectives interact (Figure 45). Sensemaking exists in the interaction between these stakeholders. As Maitlis (2005:22) argues,

‘relatively little is known about the dynamics of sense making when different parties engage simultaneously or reciprocally in such activities, or about the ways in which the accounts that they generate are reconciled – or are not reconciled.’

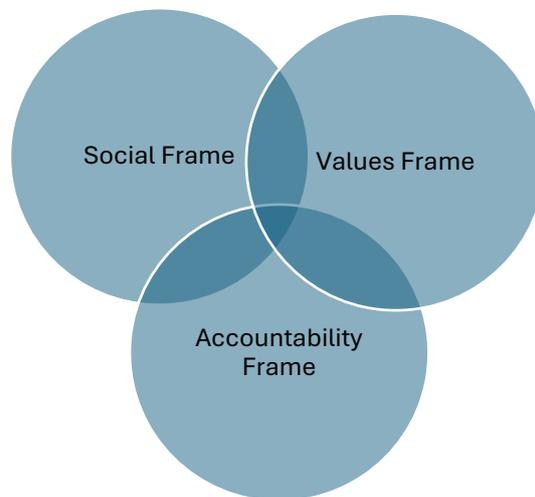
By exploring both stakeholder and leader sensemaking, this thesis builds a picture of the processes of communication undertaken when giving sense to AwL together with relationship dynamics between leaders and other stakeholders. Stakeholder’s sensemaking can also affect strategic decision making, and when combined with leader sensemaking it has been shown to shape the processes and outcomes of organisational sensemaking (Maitlis and Lawrence 2007). Gioia and Chittipeddi’s (1991) study emphasised the significant role that leader sensemaking plays in effecting major change, and Maitlis (2005), and Maitlis and Lawrence (2007), observed that leader sensemaking is also important during times of stability. Together with leader sensemaking it is necessary to understand how other stakeholders engage in sensemaking within schools.

Figure 45: The interaction between school stakeholder perspectives



Although much has been written about sensemaking and also that it is situated, less has been written on the specifics of the influences within those situations. In this thesis, the enactment of AwL is the vehicle that enables a detailed analysis of the influences upon headteachers and teachers within the six participant schools and these are presented in this chapter. Even where scholars have explored the influences on sensemaking none seem to have considered whether these influences can be categorised into frames. Zuckerman (2019) explains that In the social movement literature, collective action frames refer to the outcome of negotiating shared meaning during group sensemaking. For Snow and Benford (1992), a frame is an interpretative schemata comprising objects, situations, events, experiences, and sequences of action in one's present or past environment for the purpose of clarifying the agency that one has to act. They explain that frames are derived from master frames which are broader societal narratives, for example, the injustice of workers not sharing in a company's profits. Gahan and Pekarek (2013) explain that as cognitive schema, frames not only work to highlight the features of a social situation but also help to articulate that in the form of a grievance and the solution that is needed. Snow and Benford (1988) identify three interrelated purposes of framing: diagnostic, prognostic, and motivational framing or in other words, identify that these is a problem, explain what is causing it, and make a plan to solve the problem. These processes are likely to be contested, particularly in a strategic action field where there are multiple counter-frames and so Snow et al. (1986) argued that the frame must align. Using this concept of the frame and an inductive approach to analysing interview data, the findings below are presented in three frames of sensemaking: social, values, and accountability (Figure 46).

Figure 46: Frames of Sensemaking in a School



### 8.2.1 Social Frame

It has been established in the literature that sensemaking is a social process and that communication is a central component (Weick et al. 2005). The social influences come from multiple actors some of whom are local to the school, such as nearby schools with whom the participant school has had a long running relationship. Coburn (2001) focused on how teachers' networks and contacts shaped their reactions to policy and Vangrieken et al. (2017) argued that teacher communities were central to professional development. With the enactment of AwL, the absence of national guidance and local authority support, prompted schools such as SchoolPoints to use their existing local school partnership as their first reference point for their own conceptualisation of what an AwL system should look like and to benchmark standards, 'this is in the early days and there is another school down the road that is using a similar type of idea, and we talk to them about what they are doing, and we try to get to what equates' (Head-Points). Headteachers such as Head-Passports who developed an AwL system early felt that they were able to influence headteachers in their school partnership, 'I went quick, so it was more about me imparting what I was doing to them rather than them influencing what I was doing' (Head-Passports). Similarly, Headteachers who purchased an AwL system such as Head-SPT were able to provide leadership and guidance to schools in their newly formed MAT, 'Because we bought the system earlier than several other schools, I had a better handle on it, so I was probably the

cheaper resource to go to other schools to explain to their staff how to use it, and how we were using it' (Head-SPT).

The social influences on school enactment also involves conversations with teachers who are far removed from the school's locality, partnership or academy chain and can be part of teaching alliances that have been running information workshops for teachers such as the Learning First events which were held between 2016-18 following tweets from senior teachers like Dame Alison Peacock who was part of the 2011 Expert Panel on Curriculum review that recommended the adoption of a mastery curriculum and AwL. This community of teachers uses social media platforms such as Twitter to form online communities of practitioners and this exposes teachers to assessment approaches that may not be used by local schools. Both Head-Continuum and Head-Qualitative were headteachers that created AwL systems that had elements that were unconventional and they seemed to have been emboldened to try a novel approach by engagement with these teacher groups. Head-Qualitative explains that the combination of exposure to new ideas, encouragement from a respected school leader, and training in specific assessment tools, gave them the confidence to be unconventional if they felt that that was best for pupils, 'I think it was kind of working within a community, that community on Twitter, and going to Wroxham School (Alison Peacock's old school) for training, is kind of where we were given sort of the tools, and then I guess you would say the courage to actually step outside of the box and actually do what we felt was right for the kids' (Head-Qualitative). Head-Continuum also benefitted from meeting Peacock and then having conversations with teachers in schools that were also intending to be novel, 'we met with Dame Alison Peacock and we had a few meetings where she was talking; and we talked to a few other schools that had a similar sort of plan' (Head-Continuum).

The social influence on policy sensemaking to AwL also involved increased dialogue between school teachers and existing school partnerships and through the formation of new ones such as MATs. Headteacher dialogue also occurs nationally through teacher training forums and conversations on Twitter create online communities to exchange ideas which exposes teachers to new ideas from a wider geographical spread, and inspiring them to try unconventional ideas not used in neighbouring schools. AwL is a movement away from a technocratic approach to progress measurement using a common national rubric to teachers assessing pupils according to curriculum constructs which is addressed in the next section.

### 8.2.2 Values Frame

Siraj et al. (2014) argued that there was a sense of professional purpose wedded to primary teaching as a value-led profession (Siraj et al. 2014) and therefore, to understand the process of sensemaking, interviews with leaders and class teachers are important as the data reveals the values and priorities that frame their sensemaking. The Commission on Assessment without Levels Report (CAWL 2015) encouraged teachers to use their values and principles on the purposes of assessment as a starting point in the process of AwL sensemaking, as Head-Qualitative commented, 'it actually helped us to rethink what the purpose of everything was'. This autonomy to design an AwL system has enabled headteachers to prioritise values that benefit both their teachers and pupils.

The first value relates to increasing the integrity of teacher assessment judgement. As detailed in the literature review chapter, a key rationale behind the removal of levels was that they obscured teachers understanding of what pupils had mastered and what their next steps of learning should be (Expert Panel 2011). Previously, teachers were interpreting the differing language of a levels rubric and curricula objectives whereas under AwL, teachers just assess against curriculum statements and so have greater confidence in their own judgements about pupil progress. As teachers have used their system, they have identified key curriculum statements for each topic which aids sensemaking as it focuses teachers attention further on what needs to be taught to meet National Curriculum Key Stage objectives. Head-Qualitative explains that AwL involves a straightforward 'can do' judgement of pupil attainment against the curriculum statements rather than an interpretation against a system of progressive levels, 'subject leaders basically take the curriculum and create a system that is not a system, because it's a curriculum that they assess against' (Head-Qualitative). SPAT-Y6T described the process of identifying learning and skills objectives in the curriculum statement, spending two weeks teaching that knowledge and giving pupils practise activities so that they acquire those skills before judging evidence of attainment in a summative piece of class work or homework,

'So, we would pick a piece of writing based upon the text that we are looking at and we would know that that piece of writing would have this, this and this objective out of the national curriculum and then when we are looking at that final published piece of writing which the children would do which probably would take 2 weeks minimum, then we are looking to see if there is evidence of those skills being able to be used in a purely independent piece of writing' (SPAT-Y6T).

The identification of the constructs within a curriculum statement together with end of Key Stage objectives framed teachers' sensemaking of what an AwL progress tracking system should look like. Head-PAT contrasts this with levels where rubric statements often did not precisely match curriculum statements resulting in teachers feeling that judgements on pupil progress lacked integrity, 'levels were made-up. We had some levels which were a bit of nonsense' (Head-PAT). This confirms a finding from the Expert Panel Report in 2011, that levels were not sufficiently specific to identify the extent to which a pupil was secure in concept or skill, 'If you had a 4b or c there's no way to actually track whether a child is all there but this is a very clear tracking system so you can literally see them moving through them' (Head-PAT). SQ-Y4T explained that this lack of clarity was compounded by Standard Attainment Tests (SATs) using a different rubric to levels, 'The assessment criteria on the SATs did not always match the school assessment, so, trying to balance those two was always quite tricky, because the criteria was very different' (SQ-Y4T). They point-out that rubrics are used occasionally during half-termly assessments and so teachers familiarity with it was limited, and contrastingly assessing against curriculum statements has improved integrity by increasing the alignment between what is taught in the classroom and how progress is measured, 'it is also much more close to what we are delivering in class, and we're using it all the time, and so that is where the familiarity with it comes from, and it is not something that I pull out every six weeks' (SQ-Y4T). Rather than judging progress numerically teachers were making sense of pupil progress by judging the evidence from their work as to how secure they were in a concept or skill, 'we would simply tick boxes about statements of curriculum areas, and it worked, so colour coded. I think it was yellow or orange and green or something like that, with obviously green being that they were completely secure in that particular statement' (SPAT-Y4T). The simplicity of judging against a curriculum statement has reduced the cognitive load on teachers and enabled them to record judgements in a simple and clear format which enhances the class teachers, SLT and parents' ability to identify skills and knowledge that need further teaching.

Integrity is also being enhanced through increased dialogue between teachers within the same school to establish what quality should look like. As SQ-Y4T explains, teachers are actively seeking feedback on their interpretation of a pupil's attainment from a teacher of another Year group<sup>18</sup> so that there is a shared understanding of progression across a Key Stage, 'So, I would take a writing book to another teacher - to the Year 5 or Year 6 teacher - and we would look together at that piece of writing to check that we all were happy with where it was and the progress that they have made' (SQ-Y4T). They observed a change of

---

<sup>18</sup> This school uses a one class Year group

approach within staff meetings after AwL to increased dialogue between teachers of teaching approaches that are working in their classroom, 'so there is a lot of shared discussion about what we are doing' (SQ-Y4T). In the same school, the deputy headteacher, SQ-Y6T, explained that the SLT felt that integrity would be enhanced through a culture where teachers were interested in the progress of all children within the school rather than just the pupils in their class. They wanted to establish relationships of trust between teachers where the dialogue involved learning from each-other and a preparedness to challenge each other, 'Your information is not as robust as you think it is; so why don't you know about the flaws in your kids... people are now asking the question about what we're gonna do about that? Because you've got that change, people are now thinking deep' (SQ-Y6T). This culture of increased dialogue was corroborated by data from teachers in SchoolPAT. HPAT described the change in moderation process from the SLT transmitting judgements about pupil progress, to class teachers being involved in creating those judgements. 'When we do monitoring, we try to do it with teachers, so we have more or less done-away with senior leadership team doing everything on our own, and then just telling teachers what the feedback is' (HPAT). SQ-Y4T's observations in SchoolQualitative that AwL encourages teachers to reflect deeply on a pupil's fluency or mastery on a topic or skill was echoed in SchoolPAT by SPAT-Y4T who described the increased rigour and challenge that they were experiencing, 'I think that is where the pupil progress meetings that I mentioned that we have with the SLT comes in. There is a lot of rigour in those meetings and you're challenged' (SPAT-Y4T). SQ-Y6T conveyed that their experience under levels had been one of unspoken frustration that baseline tests continually revealed that some pupils entering Year 6 did not have the skills that they needed, 'I have been raising the question for a long time, why kids arrive in Year 6 and they can't read fluently. And we've never had a dynamic system that can deal with it because it's accusatory' (SQ-Y6T). They perceived that for many years that the pressure to make 17 sub-levels of progress during the course of a Key Stage resulted in a lack of integrity in class teacher reporting of pupil progress, 'I would say, this child is not going to make 3a because they're barely going to make 3b; so they (headteacher) would say, but that can't be true because they came out of Year 4 with 3a and now they've plateaued. "No, they haven't, the data has been massaged". "Are you saying?" ... and we'd have this kind of thing... "Are you saying that there is no way that this child is 4c, because really that's going to be seven children that are not 4c if that's true?" And then there would then be, "can you...?" - the pressure was on' (SQ-Y6T). They claim that they would resist pressure to enter a grade that they did not believe was true, 'And it's only because I'm so belligerent and pig-headed and long-in-the-tooth, that I thought, I don't wanna... I'm not lying' (SQ-Y6T). They argue that some teachers may not feel that they could or wanted to resist pressure and so the system with levels lack integrity,

'a stronger Head, a more powerful governing body, it may be a weaker teacher, and you can see that they're going to stand there and say, "well what do you want it to look like? - I'll put that as 4c"; and then they (the pupil) don't perform in the SATs you can then put it down to a bad day. And that was largely the culture that we were in, which was awful' (SQ-Y6T).

These comments confirm conclusions reached by The Expert Panel Report (2011) when they referred to the pressure to move quickly through the levels resulted in a lack of clarity of what pupils knew and their next steps of learning and they align with observations in 2015 in two London primary schools prior to AwL made by Braun and Maguire (2018).

The second value expressed by headteachers in this study related to a desire to make teachers experience that they were trusted professionals as Head-Passports explained, 'The way that I am doing it in my school which is to believe teachers and to put professionalism back into the profession (Head-Passports). Since the Education Reform Act 1988, literature has documented the reduced agency that teachers have had together with the feeling that their professionalism was being undermined (Ball 2012) resulting in the impression that they were not trusted by policy makers (Eaude 2018). So, headteachers have attempted to enact in AwL relating to teachers is the restoration of the feeling of trust in teacher judgement, 'it was really about getting professional trust back in teaching and allowing teachers to assess in their own way in the classroom' (Head-Qualitative). Over the last 20 years researchers have documented impacts of managerial controls on the autonomy of teachers to make judgements on what to teach, how to teach it and how to assess (Arnott 2000; Simkins 2000; O'Brien and Down 2002) and recent research has connected managerialism with teacher identity and teacher wellbeing (Skinner et al. 2021). Headteachers in this study seem to have viewed the capture of pupil progress as being based upon teachers professional judgement rather than on any system whether that was with or without levels, 'day-by-day good teachers do that (assess pupils), and they don't need a system; they don't need to write it down - they know' (Head-SPT). AwL has allowed headteachers to create a system that is based upon trusting their colleagues, 'What I strive for as a Head is to make sure that I have accurate, robust data where I allow my teachers' competence and professionalism to be at the centre' (Head-Passports).

A third value guiding headteachers' sensemaking was improving teacher wellbeing by reducing workload pressures, 'I haven't tied them to a cumbersome work-life balance that is impossible to meet' (Head-Passports). Dissatisfaction with workload was cited frequently by

teachers leaving the profession and the resulting shortage of teachers had prompted a national conversation on teacher workload (Higton et al. 2017). Braun and Maguire (2018) found that primary school teachers were dissatisfied with a large proportion of their time being spent upon data-collection and analysis activities and this was expressed by headteachers in this thesis. Head-Passports explained that they wanted teachers to use their time for planning and delivering teaching with assessment of what pupils need being embedded into the classroom interactions rather than making them 'spend hours on the computer to please me by making the boxes go purple!' (Head-Passports). In Phase 2 of data collection in this thesis, SQ-Y5T exemplified the issue by related an experience of how administrative burdens had effected their wellbeing at a previous school when they were a newly qualified teacher approximately 2 years previously,

'One (school) was, very, over-managed - there was lots of pressure from above (SLT) for books to look a certain way, and this was all for Ofsted, and I did not know that that was the norm, because that was my first year of teaching, so when this started happening I was actually really stressed - I lost two stone in weight through stress, which for me is - there's not much of me! So, it was quite a stressful time' (SQ-Y5T).

SQ-Y5T expressed greater satisfaction under the AwL system enacted in SchoolQualitative. To enhance teacher wellbeing, Head-Qualitative provided a checklist for discerning whether an assessment or data collection activity should be undertaken:

- 1) What's the purpose of reporting - why do we do what we do?;
- 2) What impact is it going to have on our children?;
- 3) What impact is it going to have upon staff well-being and workload?

They explained that only if the SLT were convinced that the assessment activity was going to enhance pupil learning would teachers allocate time to it.

In-line with Day and Gu (2018), a fourth value expressed by headteachers in this study was to prioritise the best learning experience for pupils. In Schools Continuum and Qualitative this entailed challenging normative approaches to tracking and reporting pupil progress. As detailed in the Literature Review Chapter, the Expert Panel Report (2011) had lauded the theories of a growth mindset posited by Dweck (2006) and Head-Continuum interpreted this to mean that pupils in the same school Year group will learn at different speeds but with effort and guidance all pupils would attain key learning outcomes. Their logic was that judging pupil progress against general expectations for that Year group was to ignore the theories of a growth mindset, 'we weren't going to have any levels - we weren't going to

have anything that dictated how this should be in this Year' (Head-Continuum), and so they created an AwL system that had age-related expectations (ARE) over a Key Stage, for example, between Year 3 and Year 6, 'all children could sit anywhere on that continuum' (Head-Continuum). They felt that they had latitude to do this as AREs are set out in the National Curriculum across a Key Stage rather than at the end of each school Year and that this would remove comparison between children as each pupil was progressing on their own continuum, 'and it wouldn't necessarily mean that if your child was in Year 3 and they only got 6 in handwriting, it wouldn't mean that they were behind or any other child' (Head-Continuum). Although it seems as though the headteacher values were compatible with the policy intentions, this approach was criticised by three local authority advisors who insisted the Year-group expectations were necessary for progress and attainment tracking, 'over about a Year we had some local authority people who came in who chose to say that this just won't work for any purpose for collection of data because it doesn't give a Year group or an age-related score' (Head-Continuum). After resisting for almost two years, eventually Head-Continuum felt compelled by school governors to compromise their values, 'we were sort of forced to break it down into Years' (Head-Continuum).

Head-Qualitative perceived that pupils were over-tested and that teachers were over-reliant on test data to discern strengths and weaknesses of their pupils, 'we were a school where every turn, we were giving sort of practice SATs papers (in Years 2 and 6) to assess the children and determine where they were. We took all of that testing away' (Head-Qualitative). SQ-Y6T explained further that classroom teaching was designed to be formative so that pupils were gaining practise in answering questions on the topics that would be covered in the SATs papers rather than using SATs papers to practise and consequently making pupils focused on passing a test paper rather than learning, 'we are doing SATs, but it's not SATs, we are not *doing* SATs. We do the arithmetic, because we are looking at the different ways of doing it, and I'm trying to train them to do questions in 45 seconds, but we're doing it from a "understand your learning" point of view' (SQ-Y6T). They proceeded to explain that they experienced resistance from parents who expected their children to have a similar amount of test practise as their friends' children may be receiving in other schools and that this put pressure on the willingness to stick to their values, 'it takes a lot of courage when you've got parents hammering the door down from September (the start of Year 6) ...the pressure comes from everywhere else, and that is hard, because they will say in that school, my friend, they've been doing SATs since the middle of Year 5' (SQ-Y6T). The culture that the SLT were trying to establish in SchoolQualitative was for teachers throughout the Key Stage feeling that they are preparing pupils to be ready for taking SATs through all learning that takes place in the school rather than just being the Year 6 teachers'

responsibility and consequently 'Year 6 would become a point of refinement, rather than trying to fill the gaps and then trying to make do' (SQ-Y6T). The idea seems to be that greater dialogue about pupils between teachers in different Year groups encourages authentic judgements and genuine reporting rather than reporting to avoid disapproval by the headteacher.

This removal of classroom testing to determine termly progress had consequences for reporting to parents and governors could not be numerical but consisted of teacher perceptions, 'So, we were adamant that we did not want to measure pupil progress by percentage – that we didn't want to have to give governors data and numerically provide progress data, but rather that we wanted to give a summary of progress' (Head-Qualitative).' School governors resisted this change as they were trained to hold the school accountable based on this test data. However, the values and principles underpinning their AwL system were precious to Head-Qualitative and they were unwilling to compromise, so in-contrast to SchoolContinuum, it was the governors who had had to compromise after several years of disagreement resulting in some governors deciding to leave their roles. In their group discussion, school governors explained that they struggled with being able to discern pupil progress due to the lack of numerical data, 'how do we hold the school, as governors... how do we do it?; how do we make our responsibilities... how do we tick the box for Ofsted?... So, there was a difficult time for, maybe, probably, 18 months' (SQ-GA). This concern about providing sufficient data for accountability processes is another part of the frame of sensemaking.

### 8.2.3 Accountability Frame

Although AwL has allowed schools to exercise a high level of autonomy, it is tempered by engagement in accountability processes and also thoughts about what accountability processes may require. Troman (1997:349) argued that 'inspectors are the absent presence in the school'. In this thesis, headteachers conveyed that sensemaking the anticipation of an Ofsted inspection, drove the pace of their adoption of an AwL system, that 'I had to go early on that because I was expecting Ofsted... I didn't want to be moving into a new system or be caught down the gap of no system when I was likely to be receiving an Ofsted inspection' (Head-Passports); and guided their initial sensemaking to what AwL should look like, and also determined the type of system that they selected, 'knowing that we would have to produce data (for Ofsted) on attainment in Year attainment was also a fairly significant driver towards that decision' (Head-SPT). The absent presence of Ofsted is not only experienced through a school's activities but is also felt through the accountability experiences of

neighbouring schools. Head-Continuum had wanted to adopt an unconventional system that utilised end of Key Stage expectations as a benchmark for pupil progress rather than Year group expectations but was thwarted by opposition from the School Governing Board (GB). Head-Continuum felt that part of the reticence from the Chair of the GB to allow the school to deviate from norms was a fear of receiving a poor judgement from Ofsted inspectors as some other local schools had done, 'in the area that we are in, several local authority schools are in special measures or in some sort of classification, and that's created a bit of a furore (among the school governors' (Head-Continuum). This risk averse approach to anticipated Ofsted judgements is common among GBs (Wilkins 2015) and influences sensemaking within primary schools. This absent presence of Ofsted can also be positive in the form of confidence if the inspection was favourable as Head-Qualitative expressed, 'the inspector said the assessment system that we were using, and the feedback, and the way that we were using feedback was absolutely the right thing' (Head-Qualitative). However, generally, a positive judgement such as a school being 'outstanding' results in headteachers having the assurance that another inspection will not occur soon and so Ofsted's presence in a minor one, 'We got 'Outstanding' so Ofsted haven't come again' (Head-Points).

Although AwL systems are now different in each school, national testing in the form of the Standard Attainment Tests (SATs) for pupils in Year 2 and Year 6 remain a common accountability process for judging primary schools' performance and this contributes to the sensemaking process to AwL. Teachers in this study felt that this resulted in a new system that mixed the previous system of levels with AwL: 'it's the bit of levels that never went - it's the radiation left-over from the extraction. That's what they (the government) have to have' (SQ-Y6T). Head-PAT expressed their frustration that SATs continue to measure separate learning outcomes from the National Curriculum that is the basis of AwL progress measurement, 'I've got no problems with Assessing without Levels, it's given us more responsibility for how we say we are tracking and how we are monitoring, and what we call it and what we do about it, but it doesn't tally at all with SATs and SATs are archaic' (Head-PAT). In Phase 1 of data collection, Heads Points and Passports explained that they met this challenge by identifying objectives from the curriculum that also prepare pupils for SATs and then make them non-negotiable objectives in both SATs Years, that is, Years 2 and 6, and the Years previous to these, that is, Years 1 and 5, so that pupils are ready for SATs Years. In Phase 2 of data collection, Head-PAT explained that they had also adopted that approach. In all participant schools, the AwL system was not used to moderate in Years 2 and 6 but the SATs Interim Framework was used both within the school and when moderating with partner schools. In Schools Points and Passports optional SATs papers

were used to verify pupil progress judgements made by teachers as they were gaining confidence in their new system.

Therefore, schools needed to ensure that their AwL systems would prepare pupils to pass SATs together with meeting National Curriculum outcomes. Initially, SchoolQualitative's AwL system seemed to effect pupil performance in the SATs maths test which prompted the teachers to consider how the two systems should complement each other, 'we felt that we should find a way that the SATs fitted into this' (Head-Qualitative). Their alterations had the desired effect, 'we've gone from bottom 3% to top 7%' (Head-Qualitative). They attribute improvement in the SATs results to three features of their AwL system. Firstly, they believe that their AwL system has higher expectations for each Year group compared with the previous system of levels, 'we automatically had higher expectations which has helped us' (Head-Qualitative). Secondly, their AwL system was designed to establish a greater alignment of learning objectives in previous Year groups with the outcomes that need to be achieved in the SATs in Year 6 resulting in pupils being better prepared for entering Year 6, 'it's focused everybody's work on Year 6 and it's meant that we don't have to have a narrowed curriculum in Years 3, 4, 5' (Head-Qualitative). Thirdly, the AwL system identifies pupils who need additional support in Year 1 and 2 and by ensuring that they do not fall behind there is then less pressure to provide additional support to them in Year 6, 'if we identify children early on, we hope that through this system it sets a high enough standard that we don't need to pull a huge amount of support into Year 6 for instance, to ensure that we get better grades' (Head-Qualitative). Several headteachers pointed-out that aligning AwL systems with SATs outcomes was a challenge as SATs objectives were different to the National Curriculum objectives. The influence of accountability on teachers' sensemaking of AwL is illustrated by this quotation from Head-PAT who believes that SATs should be reformed if schools are to have true autonomy to enact AwL precisely how they think it should be,

'at the end of the day they haven't given us any more responsibility at all because it's still all measured in a punitive way in that test on the day so the two systems aren't working together and they never will unless they take SATs away; or they think of a different way of doing it' (Head-PAT).

Although AwL provides headteachers with autonomy to design systems according to their values, their expression of their values is constrained or tempered by the frame of accountability whether that is in the form of Ofsted judgements or SATs results which would be published in school league tables. As detailed in the literature review chapter, this tension

between autonomy and accountability is a dominant theme in the schools policy literature. The introduction of the National Curriculum, SATs, school league tables and Ofsted inspections during the Conservative governments of the 1980s and 1990s was for the purpose of curtailing school autonomy. These accountability technologies still frame policy sensemaking within primary schools even for a policy such as AwL that is not directly connected with accountability processes.

Categorising interview data in the form of social, values and accountability frames provides adds to the explanations offered in Chapter 7 on why AwL was enacted differently in each school and also provides a useful model for understanding how a curriculum or assessment policy may be received in future. Although the written data is presented separately under the category of each frame, Figure 46 displays these frames in a Venn diagram emphasising the overlap that occurs between social influences, teacher values and accountability frames when policy sensemaking occurs within a school. These frames are likely to be contested and in tension with each-other particularly when deciding on a strategy for enactment as was illustrated by the tensions between the School Governors in School Qualitative who wanted more of a numerical measure of progress and the values which caused tension with Head-Qualitative's values. Benford and Snow (2000) call such challenges from within an organisation, frame disputes, and they term challenges external to the organisation as framing contests, such as the Local Authority Advisors who challenged the Head-Continuum's interpretation that AwL required the removal of year-group expectations. In order to neutralize a counter-frame, the challenged party seeks to 'reframe' or displace it as the dominant way of seeing which Head-Qualitative succeeded in doing but Head-Continuum did not. In this way, frames change and develop in the ongoing interaction. Snow et al. (1986) argued that the frame must align and identified four frame alignment processes: bridging, amplification, extension, and frame transformation. Bridging involves building connections between two distinct frames at both the organizational (aligning vision and goals) and individual (sentiment) levels, such as, Head-SPT altering the feedback to parents about their child's progress when they realised that they wanted to know if they were meeting expectations or below or above them. Frame amplification involves the clarification and invigoration of values or beliefs important to the target audience which was illustrated by Head-Passports who explained that the expectation that Ofsted inspectors would visit drove the speed of their adoption of an AwL system and also some of the content of it. Frame extension involves a group drawing connections between its own primary interests and values and those of other groups by framing them as mutually compatible sometimes adopting new values, meanings, understandings and beliefs (Benford and Snow 2000). This is particularly evident in the two schools joining Multi-Academy Trusts (MATs), especially

SchoolPAT whose MAT provided the same AwL system to all schools in the Trust while allowing each school to personalise some elements of it to suit their needs.

### 8.3 Chapter 8 Discussion: AwL as a case of successful policy enactment

Together with an understanding of sensemaking, the exploration of the paradigm of schooling also requires an examination of the practices that policy makers and policy enactors engaged in. The data from this thesis suggests that this has been a case of successful policy implementation achieved through clear communication of policy intentions, allowing school leaders to have autonomy to design their systems, ensuring that teachers experience pedagogical improvements in the classroom, and with strong national accountability processes being maintained (Figure 47).

Figure 47: The Four Keys to the Successful Policy Implementation of AwL



AwL has succeeded due to the clarity in communication of policy intentions which started with the Expert Panel Report (2011), particularly Chapter 8 (Popat 2015), and was consolidated in the Commission for Assessment without Levels (CAWL) Report (2015). The House of Commons Report on Primary Assessment (2017) concluded that schools were not provided with sufficient clarity, support or training from the Department of Education and that schools were just left to identify what they wanted to do, but all six school in this thesis have enacted the intentions of the policy which suggests that they were clear as to what AwL was meant to be. Firstly, these reports justified the removal of levels by explaining the problems that they caused such as being representative of attainment with broad descriptions but

without specifically identifying which aspects of the curriculum had been mastered by the pupil and which curriculum constructs the pupil was still weaker in. This caused particular problems when pupils moved to the next Year group as the receiving teacher would often disagree with the level that had been assigned to them after conducting diagnostic tests. However, as levels were presented as a ladder the mindset that pupils, teachers and parents had was that good progress equated to rapid progress up the levels rather than on building a secure understanding of curriculum constructs. This also encouraged streaming of pupils as inevitable some pupils learnt more quickly than others which resulted in pupils not covering less of the curriculum than others and labelling themselves and being labelled as less intelligent or less able than quicker learners (CAWL 2015).

The CAWL Report (2015) argued that AwL should comprise of classroom formative assessment, in-school summative assessment and national tests. CAWL said that day-to-day in-school formative assessment is integral to the ongoing process of teaching and learning within the classroom and through homework. They listed activities that fall into this category such as question and answer sessions during class, marking work, conducting observational assessments and topic recap quizzes. The CAWL Report (2015) proceeded to outline that day-to-day in-school formative assessment served different purposes for different stakeholders but for formative assessment the purposes are similar for all of them. Teachers need to gauge the knowledge and understanding of a pupil relative to learning objectives and broader outcomes, thereby identifying areas requiring targeted improvement and any adjustments needed to instruction, and leaders need to know that teachers know this. Parents and pupils need to have information from teachers on their next step of learning so that there is a collaborative effort between home and school to support the child's educational development. CAWL urge the government to not interfere with this process and reiterates that the Ofsted framework for 2015 states that the specifics of marking and feedback are at the discretion of the school. The CAWL Report (2015) explained that together with formative assessment, in-school summative assessments should be conducted at the end of a topic or half-term, term of school Year to gauge student learning. They felt that this would benefit teachers, pupils and parents for formative purposes in a similar way to formative assessments, and school leaders can use the data to monitor cohort performance.

### 8.3.1 Allowing Headteachers to have Latitude to Design AwL Systems

The second key to the implementation of AwL was that the Department for Education did not proscribe the specific details of what an AwL system should look like but gave headteachers latitude to design systems according to their values and the contextual needs which seems

to have resulted in greater satisfaction. The results of Federici's (2013) survey study revealed a positive relationship between school headteacher self-efficacy and perceived job autonomy. Headteacher self-efficacy and perceived job autonomy were positively related to job satisfaction and negatively related to contextual constraints. Contextual constraints to autonomy were negatively related to job satisfaction (Federici 2013). The CAWL Report (2015) emphasised such headteacher autonomy while providing some scaffolding in the form of reflection questions. The first question asked headteachers to consider what insights the assessment should provide to discern the security of the students' knowledge and understanding of the subject matter, concept, or skill and whether they can apply that with a growing independence. To ensure inclusivity for pupils with Special Educational Needs and Disabilities (SEND), thought should be given on how they can demonstrate their learning. Secondly, headteachers were asked to reflect upon how they would communicate the purpose of an assessment and also the assessment data with pupils so that they know their next steps in learning. Thirdly, headteachers were asked to encourage teachers to reflect on how they could adjust teaching to improve pupil understanding and to consider whether some pupils need extra teaching or practice. Fourthly, headteachers were asked to consider how to support progression where learning is secure in quicker learners as they should not be taught outside of the topic being taught to the rest of the class. This might involve assessing whether pupils can apply the concept in different contexts or to explore a concept in greater depth before moving on to new material. Fifth, headteachers were asked to consider how to record and depict progress. None of the headteachers in this study have resisted implementing AwL according to these principles even if they resisted numerical reporting and classroom testing as in the case of Head-Qualitative or an attempt to resist Year-group expectations as in the case of Head-Continuum. Practices such as formative assessment, adapting teaching approaches, and conveying progress and next steps of learning to pupils and parents, are accepted parts of schooling and so headteachers never had autonomy to remove them and it can be viewed that the reflection questions provided by the CAWL Report (2015) emphasise that. What AwL provided was latitude to enact these teaching functions according to a headteachers values and to meet the needs of the stakeholders in that school context which seems to have worked by providing sufficient autonomy to enable headteacher satisfaction while maintaining a large degree of standardisation between schools.

### 8.3.2 Teachers Experiencing Improved Pupil Outcomes

AwL has been enacted together with a new National Curriculum that expects teachers to teach for mastery and the data in this thesis from interviews with class teachers in Schools

PAT and Qualitative demonstrates that this has improved the experience of teaching as they now focus on planning teaching according to constructs in curriculum statements and engage in continuous classroom conversations with pupils and with the work that pupils produce to discern the extent to which they have mastered the learning goal. This has changed teachers relationship with pupils as no child can be allowed to stagnate as they must be set a challenging task which also challenges the teacher to think how best to support that pupil to achieve them. Butler and Shibaz (2014) explain that teachers who adopt a mastery approach, enhance their own knowledge and skills as they themselves are challenged by the challenging tasks that they set pupils and the progress monitoring that they focus on. Under levels, teachers would stream pupils into ability groups and so their expectations on the possibilities of progress varied according to the pupil which resulted in some pupils being challenged less. Teachers in this DPhil study also explained that AwL had simplified progress monitoring as it involves tracking of evidence that a pupil has demonstrated understanding of the curriculum construct rather than having to interpret a rubric as they did under levels. This results in teachers seeing pupils achieve their learning goals on a daily basis which has increased enthusiasm in the pupils, teachers and parents as they also get involved in helping their children achieve their next step of learning when they are at home. This is in contrast to Braun and Maguire's (2018) finding in primary schools under levels that there were 'emotional costs expressed by teachers about enacting policies that they do not believe benefit all their children' (Braun and Maguire 2018:11). The feeling that primary school teaching had replaced its emphasis on caring for all pupils with a focus on achieving a target that these authors reported, seems to have been changed under AwL as the learning goal is individualised for the pupil.

### 8.3.3 High Stakes Standardised Accountability Systems

National accountability testing in Years 2 and 6 in primary schools are implemented to ensure uniformity in evaluating student performance across the nation (CAWL 2015) carry high stakes for all stakeholders in a school as they are used to compile league tables and inform inspections by Ofsted. This has prompted all schools in this study to use Year 2 and Year 6 Interim Standards to benchmark the standards in their AwL system and also to benchmark with partner schools. This approach is in-line with observations by Braun and Maguire (2018) under the levels system that despite disapproval of these extensive testing regimes, active opposition was unusual as school leaders and teachers embody pragmatic compliance (Keddie 2018). Ball (2003) noted that these accountability mechanisms prompt leaders and teachers to organise themselves in response to inspections and test performance (Ball 2003). Bevan and Hood (2006) observed that across public services in

England, accountability process combined targets with 'an element of terror' (Bevan and Hood 2006:517); and Braun and Maguire (2018) explained that teachers had a defence mechanism against a fear of failure by attempting to anticipate what inspectors may want to see. These behaviours were still apparent under AwL. In the early stages of interpreting AwL, Head-Passports and Head-SPT felt that they had to anticipate what Ofsted may want to see in an AwL system when choosing despite Ofsted saying that they would not be looking for anything specifically. The fear of what inspectors may expect also prevented Head-Continuum from enacting an AwL system without Year-group expectations.

## Chapter 9: Conclusions and Recommendations for Future Research

### 9.1 Chapter 9 Introduction

This is an ambitious DPhil study that seeks to use a small interview study to provide insights into the wider paradigm of school policy enactment. The combination of the inductive approach of learning from policy documents and academic research together with a deductive approach to the exploration of the topic that framed the co-production of data with participants, followed by the inductive approach to comprehending interview data and then a deductive approach of using academic theory to reconstitute that primary data into theoretical arguments has captured the nuance involved in understanding both the 'how' and the 'why' of the enactment of AwL. Just as the inductive and deductive approaches have interacted somewhat like a dance of approaches, the two research questions of how AwL was enacted and why it was enacted that way have also interacted throughout the data rather than being two distinct queries – to understand the *how* one must also understand the *why*. The flexibility and adaptability of the interview process and the ability to modify questions to the specific experiences and backgrounds of participants and in response to participant responses lead to more relevant and meaningful data. This adaptability is crucial in uncovering unexpected insights and new areas of interest, which can significantly inform future research endeavours. Furthermore, the thesis underscores the importance of contextual understanding gained through interviews. By exploring the context in which participants' experiences and behaviours occur, the research has developed theories and hypotheses grounded in real-world settings. This contextual knowledge enhances the relevance and applicability of future studies. The in-person interviews mostly conducted at the schools enhanced participant engagement as the rapport established between researcher and participants fostered honest and open responses, and more nuanced and thoughtful answers which contributed to a richer dataset. The exploratory nature of a small-scale interview study such as this is its ability to identify key themes, patterns, and variables that warrant further investigation, thereby laying the groundwork for robust research questions and the design of larger-scale studies. This thesis also effectively identifies gaps and opportunities in the existing literature by highlighting underexplored topics, emerging trends, and potential research opportunities, which can shape the direction of future studies. Additionally, the thesis illustrates how small-scale interviews can serve as a pilot test for larger research projects. The insights gained from these initial interviews help refine interview techniques, question design, and data analysis methods, improving the overall quality and effectiveness of future research.

## 9.2 This Thesis' Contribution to Policy Implementation Theory

Previous research has indicated that successful education policy implementation is a multifaceted process influenced by several critical factors such as, smart policy design, inclusive stakeholder engagement, a conducive context, and a coherent implementation strategy. Firstly, policies must be clear, coherent, and aligned with the educational goals they aim to achieve. A well-designed policy should include specific, measurable objectives and provide a clear roadmap for implementation. It should also be flexible enough to allow for adjustments based on feedback and changing circumstances (Viennet and Pont 2017). The clarity and feasibility of the policy design significantly impact its acceptance and execution by educators and administrators. Secondly, engaging stakeholders at all levels is crucial for the successful implementation of education policies. This includes teachers, school leaders, parents, students, and community members. Effective stakeholder engagement ensures that the policy is understood, accepted, and supported by those who are directly affected by it. It also provides valuable insights and feedback that can help refine and improve the policy. According to Viennet and Pont (2017), inclusive engagement fosters a sense of ownership and commitment among stakeholders, which is essential for sustained implementation efforts. Thirdly, the broader context in which a policy is implemented plays a significant role in its success. Policies are more likely to succeed in contexts where there is strong institutional support, adequate funding, and a stable political climate (Viennet & Pont, 2017). Contextual factors can either facilitate or hinder the implementation process, and understanding these factors is essential for effective policy execution. Fourthly, continuous professional development and support for leaders and teachers are also essential components of a successful implementation strategy (Fullan 2016). Fourth, a clear strategy, coupled with regular monitoring and evaluation, is essential for the successful implementation of education policies. These elements help identify challenges and areas for improvement, allowing for timely adjustments to the implementation plan. According to Viennet and Pont (2017), continuous monitoring and evaluation are crucial for ensuring that policies are effectively executed. This process involves systematically assessing the progress of policy implementation, understanding its effectiveness, and identifying any unintended consequences. By doing so, policymakers can make necessary adjustments to improve outcomes and ensure that the policy remains relevant and effective in achieving its goals (Williams 2003). Fifth, clear and consistent communication is vital for the successful implementation of policies (Albright et al. 2022). Stakeholders need to be well-informed about the policy's objectives, processes, and expected outcomes to ensure alignment and support. Effective communication strategies help build a shared understanding and commitment among stakeholders, which is essential for the smooth execution of the policy.

The data from this thesis adds to the literature by suggesting that this has been a case of successful policy implementation achieved through clear communication of policy intentions, allowing school leaders to have autonomy to design their systems, ensuring that teachers experience pedagogical improvements in the classroom, and with strong national accountability processes. The clear communication of policy intentions began with the Expert Panel Report (2011), particularly Chapter 8 (Popat, 2015), and was consolidated in the Commission for Assessment without Levels (CAWL) Report (2015). Despite the House of Commons Report on Primary Assessment (2017) concluding that schools lacked sufficient clarity, support, or training from the Department of Education, the six schools in this thesis have effectively enacted the policy's intentions, indicating a clear understanding of AwL's objectives. Secondly, the CAWL Report (2015) advocated for headteachers to have the latitude to enact assessment functions according to their values and the needs of their context. The results of Federici's (2013) survey study revealed a positive relationship between school headteacher self-efficacy and perceived job autonomy. Headteacher self-efficacy and perceived job autonomy were positively related to job satisfaction and negatively related to contextual constraints. Contextual constraints to autonomy were negatively related to job satisfaction (Federici 2013). The CAWL Report (2015) emphasised such headteacher autonomy while providing some scaffolding in the form of reflection questions.

Thirdly, the implementation of Assessment without Levels (AwL) alongside a new National Curriculum has led to improved teaching experiences, as evidenced by interviews with teachers in Schools PAT and Qualitative. Teachers now focus on planning according to curriculum constructs and engage in continuous classroom interactions to assess mastery of learning goals. This shift has altered teacher-student relationships, ensuring that no child stagnates and that each is set challenging tasks, which in turn challenges teachers to support their students effectively. Butler and Shibaz (2014) note that a mastery approach enhances teachers' knowledge and skills as they are challenged by the tasks they set and the progress they monitor. Previously, under the levels system, teachers streamed pupils into ability groups, leading to varied expectations and less challenge for some students. Teachers in this study reported that AwL simplifies progress monitoring by tracking evidence of understanding rather than interpreting rubrics, resulting in daily achievements that boost enthusiasm among pupils, teachers, and parents. This contrasts with Braun and Maguire's (2018) findings that the levels system imposed emotional costs on teachers who felt the policies did not benefit all children. AwL, by individualizing learning goals, appears to have shifted the focus back to caring for all pupils. Fourthly, national accountability testing in Years 2 and 6 remains high stakes, influencing schools to use interim standards for

benchmarking within their AwL systems and with partner schools. Despite disapproval of extensive testing regimes, pragmatic compliance persists among school leaders and teachers (Keddie 2018).

### 9.3 Further Research on Feedback

One of the contributions of this thesis is interview data on primary pupils' experiences of classroom assessment. Future research could explore the impact of continuous dialogue and verbal feedback in the classroom, as exemplified by Schools PATs and Qualitative AwL approaches. These involve a fluid, fast-paced interaction between teachers and pupils, focusing on small daily objectives to ensure clear understanding and studies could investigate how this method affects pupil engagement and learning outcomes, particularly for those with Special Educational Needs and Disabilities (SEND). In classroom discussions, pupils are encouraged to articulate their thoughts and respond to their peers by either supporting, expanding upon, or challenging their ideas. This interactive dynamic fosters a debate-like environment where pupils learn to be original in their responses. Further research could explore the benefits of such interactive discussions on critical thinking, communication skills, and collaborative learning.

The desire for challenge is instilled in pupils from Year 1, following the approach advocated by Pardoe (2009) who suggests that pupils learn best when they feel comfortable being challenged and challenging themselves, a concept known as the 'danger zone'. This approach is effective only if teachers encourage it and provide feedback that praises the effort in tackling challenges rather than merely completing tasks correctly. Future studies could investigate the effectiveness of this approach in different educational settings, its impact on student motivation and resilience, and how it shapes students' attitudes towards learning and perseverance. Pupils in both Schools PAT and Qualitative expressed an understanding that they should not fear mistakes as the school values learning from them. The tone of adult responses to mistakes is crucial, as it impacts pupils' psychological traits. Pupils appreciate when teachers provide constructive feedback without anger, often using humour to address mistakes. Further studies could explore the effects of different feedback styles on student learning, emotional well-being, and the development of a positive classroom climate. How these then combine to create a school culture that promotes a growth mindset, its effects on student achievement, and how it influences students' attitudes towards failure and success. Duckworth (2016) introduced the concept of GRIT, which stands for Growth mindset, Resilience, Independence, and Teamwork. Pupils explained that

they are taught these traits and recognize different types of mistakes, such as silly mistakes and mathematical errors. They use mistakes as tools for SATs preparation by formulating questions with common errors for their classmates to identify and challenge. Research could be conducted to assess the impact of teaching GRIT on students' academic and personal development, their persistence in the face of challenges, and their long-term educational outcomes. Pupils in both Schools PAT and Qualitative are encouraged to express when they find something challenging and seek ways to make it manageable rather than overwhelming. This approach discourages labelling tasks as simply right or wrong, or easy and hard. Research could be conducted to examine how this mindset influences students' problem-solving skills, their ability to cope with academic challenges, and their overall academic self-efficacy. Future research could also explore the impact of pupils feeling secure in their understanding and the role of teacher insights into pupils' thought processes.

From an early age, pupils learn to assess the difficulty of their work and take responsibility for their role in it. Pupils also practice their writing and mathematics skills in different books, with one for main work and another for skills practice. This process helps them plan and refine their work, ensuring continuous improvement. Further studies could investigate the benefits of this dual-book system on student organization, skill development, and the ability to transfer skills across different contexts. Investigating the effectiveness of pupil reflections, written during or after lessons, could provide valuable data on how self-review influences learning and identifies areas needing support. This practice encourages independent reflection, which is crucial for developing self-regulated learners. Year 6 pupils demonstrated their use of self-evaluation sheets, which they primarily used to reflect on written work across various subjects. This practice underscores the importance of self-assessment in fostering reflective learning. Further research could explore the long-term impacts of these self-evaluation sheets on academic performance, personal development, and how it influences students' self-regulation and metacognitive skills. Further studies could examine the concept of Independent Working as understood by pupils, particularly how working alone or with minimal assistance enhances learning outcomes. Research could also explore the balance between independent work and collaborative learning, assessing how peer assistance without direct teacher involvement impacts student autonomy and engagement.

Understanding how greater student autonomy in learning and assessment leads to improved outcomes and higher engagement could inform teaching practices and curriculum design. These insights could help develop strategies to promote a sense of responsibility for learning among pupils, ultimately enhancing their educational experiences. Year 4 pupils explained that science lessons required them to use higher-order skills like evaluation and self-assessment. They documented their experiments, including conclusions and methods, which fostered critical thinking. Further research could explore the impact of structured science

lessons on cognitive development. Furthermore, the use of home/school diaries for continuous communication between teachers, pupils, and parents could be examined. Research could explore how this practice fosters a community of learning and enhances parental involvement in their children's education.

One pupil highlighted that the lack of feedback in core subjects sometimes led to reduced motivation among classmates. They emphasized the importance of feedback for maintaining motivation and developing transferable skills. Research could examine the relationship between feedback, motivation, and skill development. Pupils noted that feedback in PE often focused on behaviour, while during the Music class, they received constructive feedback on technique. They expressed a desire for more feedback across subjects to enhance their skills, particularly in Art. This suggests a need for comprehensive feedback processes across non-core subjects. Another pupil mentioned that Geography feedback often focused on literacy rather than content, indicating a gap in subject-specific feedback. Further research could explore the balance between literacy and subject-specific feedback and its impact on comprehension and interest; and the effects of consistent feedback on student engagement and skill acquisition could be explored.

Particularly in SchoolQualitative, verbal feedback groups were a central part of pedagogy, replacing written feedback on homework in some lessons. Pupils participated in feedback groups where they critiqued each other's work and learn from teacher feedback given to others. This dialogue helps clarify the next steps in learning and fosters a sense of progress. Research could examine the impact of verbal feedback groups on student engagement, learning outcomes, and the development of a collaborative learning environment. Pupils generally expressed a preference for the combination of both verbal and written feedback with a desire for written feedback to be explained verbally, and so the relationship between active participation in learning and the effective communication of feedback warrants investigation. Future research could investigate the effectiveness of combining written feedback with subsequent teacher attention in enhancing pupil understanding. This approach addresses the issue of written feedback sometimes being unclear and ensures consistency in comprehension when multiple pupils struggle with a concept. Studies could explore how this combination impacts learning outcomes and student confidence. Research could assess the specific benefits of written comments in highlighting areas for improvement and serving as a reference for future learning. The suggestion by some pupils of having a separate book for recording teacher feedback across all subjects presents another area for investigation. Future studies could evaluate the practicality and effectiveness of this approach in helping pupils reflect on feedback at home, and its impact on their overall

learning and retention. Research could assess how verbal feedback clarifies written comments and how written feedback serves as a reference for future work. The practice of pupils writing responses to teacher feedback and the use of specific feedback codes (e.g., GWT, GWAT, IW) could also be explored to determine their impact on student learning and self-assessment skills.

Teaching Assistants (TAs) are now seen as professionals who regularly meet with teachers to understand learning expectations and identify gaps in pupil understanding. They share the teaching of pupils with teachers to ensure all pupils are included in the whole-school conversation on progress. Future research could investigate the impact of this collaborative approach on student learning, TA professional development, and the overall effectiveness of classroom support structures. Extra teaching is provided to reinforce learning, with some pupils receiving additional instruction on fundamental concepts during assembly times. This approach helps pupils grasp new topics and continue working with the rest of the class. Studies could explore the effectiveness of such interventions in closing learning gaps, promoting academic success, and ensuring that all pupils progress at a similar pace.

#### 9.4 Further Research on Key Performance Indicators

The implementation of Key Performance Indicators (KPIs) in primary schools is a multifaceted issue that requires careful consideration of various factors to ensure their effectiveness. One of the primary concerns is the impact of KPIs on teacher workload and well-being. Head-Passports, for instance, emphasized the importance of balancing teachers' workload with the need for detailed objectives. They argued that overly detailed KPIs could overwhelm teachers and detract from core teaching and learning processes. This highlights the need for research into how KPIs affect teacher workload and well-being, and the optimal level of detail that supports effective teaching without causing burnout. Additionally, the different attainment thresholds set by Head-Passports (85%) and Head-Points (80%) suggest a need to investigate the effectiveness of these benchmarks in accurately reflecting pupil progress and readiness for the next Year's curriculum. Understanding the impact of these thresholds can help in setting appropriate and realistic benchmarks that ensure pupils are adequately prepared for future academic challenges. Setting realistic and achievable targets is crucial for maintaining teacher morale and ensuring that pupils are not unfairly penalised by overly ambitious benchmarks. The direct impact of KPIs on student outcomes is another crucial area for research. This includes studying how different KPIs influence student performance, engagement, and overall well-being. Studies should also explore the

role of various stakeholders, including teachers, parents, and students, in the development and implementation of KPIs. Understanding their perspectives can help in creating more effective and inclusive KPI systems. Long-term studies are needed to understand the sustained impact of KPIs on student progress and institutional performance, tracking the progress of students over several years to identify trends and patterns. Future research should focus on developing comprehensive frameworks that include a wide range of KPIs to provide a holistic view of student progress and institutional performance. This includes academic, behavioural, and socio-emotional indicators.

Further research into the implementation of KPIs, such as ensuring accurate and consistent data collection, aligning KPIs with the strategic goals of the educational institution, and allocating adequate resources to monitor and analyse KPIs effectively. The accuracy and practicality of various KPI tracking systems warrant further investigation. Head-Points' experience with inaccurate percentage mapping due to broad statements underscores the need for reliable and user-friendly tracking methods. Comparing electronic and manual tracking systems could provide insights into the most effective approaches for monitoring pupil progress that can provide accurate and timely data to inform teaching practices. Furthermore, the role of KPIs in standardisation and inter-school comparisons would be interesting. The use of nationally standardised KPIs from the Year 2 and Year 6 National Curriculum Statements as benchmarks for inter-school moderation warrants an exploration on how KPIs contribute to or hinder subsequent in-school standardisation. Additionally, examining the alignment of KPIs with National Curriculum objectives and their long-term impact on pupil outcomes could inform best practices for KPI development and implementation, ensuring that KPIs support pupils' academic trajectories effectively. Such guidance on selecting and operationalising KPIs for each curriculum subject would be very valuable especially for non-core subjects.

The connection between formative and summative assessment and KPIs presents an opportunity for further research. Research could examine how the choice and enactment of KPIs to inform and prepare for formative and summative assessments and also how KPIs affect pupil learning and engagement. Additionally, the communication of KPIs to parents and their understanding of these indicators is another area worth exploring. Different ways of presenting pupil progress, such as using terms like 'Better than Expected Progress' are linked with the attainment of KPIs but it is unclear whether parents understand this. Researching how effectively KPIs are communicated to parents and their understanding of these indicators could enhance parental engagement and support for their children's education.

## 9.5 Further Research on Mastery Learning

The data in this thesis illustrates several key points about mastery-based learning. This educational approach allows students to progress through a curriculum at a variable pace, advancing only when they have demonstrated a thorough understanding of the material. It is grounded in the belief that all students can achieve high levels of understanding if given the appropriate time and support. This approach emphasises personalised, student-centred learning, where instruction is tailored to meet the individual needs of each student. The data conveys the shift in instructional practices that teachers have adopted to implement mastery learning, and the willingness that they seem to have shown to adapt their teaching methods based on student needs and to breakdown learning objectives into smaller, manageable units and the use of formative assessments to monitor student progress. When students struggle with a particular concept, they receive targeted instruction and additional practice until they achieve mastery. The data also illustrates that mastery learning can lead to improved student self-efficacy and motivation, as students experience success and build confidence in their abilities which can improve students' attitudes towards course content and instruction. In both Schools PAT and Qualitative, mastery learning helps students develop essential skills such as creative engagement, innovation, self-regulation. Student motivation and engagement are crucial for the success of mastery learning and so research should focus on strategies to enhance student motivation and engagement in mastery learning environments. For instance, studies could investigate the role of self-regulation and metacognitive skills in student success within mastery learning frameworks. Additionally, research should explore how different instructional strategies, such as goal-setting, feedback, and collaborative learning, impact student motivation and engagement in English primary schools. Understanding the factors that influence student motivation and engagement can help educators design more effective mastery learning environments that support student success.

This thesis has also shown that standards-based grading is an assessment method that aligns closely with mastery learning. Students are evaluated based on their ability to meet specific learning standards rather than on a traditional grading scale so that grades reflect students' actual understanding of the material and so it can provide a more accurate representation of student learning. The data suggests that standards-based grading can help teachers identify specific areas where students need improvement, allowing for more targeted interventions and support and that it offers clear learning targets and provides meaningful feedback to students and parents. There is a need to update scholarship on mastery learning that provided insights into schools prior to AwL. For example, curriculum

design principles are crucial for effective mastery learning and so research should focus on developing and validating curriculum design principles that support mastery learning. For instance, studies could investigate how to balance the need for individualised pacing with the requirements of standardised curricula that promotes deep understanding and application of knowledge. Teacher training programmes are essential for the successful implementation of mastery learning and future research should focus on developing and evaluating professional development programs that prepare teachers to implement mastery learning effectively. These programmes should provide teachers with the knowledge and skills needed to design and deliver instruction that meets the diverse needs of students. Additionally, research should investigate the long-term impact of these training programs on teacher practices and student outcomes. For example, studies could update scholarship on how teacher training programmes influence teachers' use of formative assessments, differentiation strategies, and technology integration in mastery learning environments. The data in this thesis conveys that several forms of assessment are used in primary classrooms and so research should explore and update the impact of different types of formative assessments on student learning and mastery, such as peer assessments, self-assessments, and technology-enhanced assessments.

Equity and access are significant considerations in implementing mastery learning. There is a lack of research investigating the impact of mastery learning on different student populations, including those from disadvantaged backgrounds. Studies should explore strategies to ensure equitable access to the resources and support needed for mastery learning. For example, research could examine the effectiveness of targeted interventions, such as tutoring and additional instructional time, in supporting students who may struggle to achieve mastery. Additionally, studies should investigate how mastery learning can address disparities in educational outcomes and promote equity in education. Longitudinal studies are essential for understanding the long-term effects of mastery learning on student achievement and development. Research should focus on conducting longitudinal studies that track students' progress over time and examine the impact of mastery learning on their academic and career success. For example, studies could investigate how mastery learning influences students' performance in higher education and their career trajectories. Additionally, longitudinal studies can provide valuable insights into the sustainability of mastery learning practices and their long-term impact on student outcomes. Technology integration is another critical area for future research. Technology can support mastery learning by providing personalized learning experiences, adaptive learning platforms, and real-time feedback. For example, adaptive learning platforms can adjust the difficulty of tasks based on a student's performance, ensuring that each student is challenged at an

appropriate level. Research should investigate the effectiveness of various technological tools in facilitating mastery learning and explore how these tools can be integrated into classroom instruction. Additionally, studies should examine the impact of technology on student engagement, motivation, and learning outcomes in mastery learning environments. Developing and testing different models of mastery learning implementation is essential for identifying best practices. Research should focus on investigating various implementation models to determine what works best in different educational contexts. For example, studies could examine the scalability of successful mastery learning models across different school settings. Additionally, research should explore the factors that influence the successful implementation of mastery learning, such as leadership, teacher collaboration, and community support. Understanding these factors can help educators design and implement effective mastery learning models that support student success.

#### 9.6 Further Research on Narrowing of the Curriculum

During Phase 2 of data collection in February and March 2019, the headteachers in Schools PAT and Qualitative were making plans with their colleagues in the Senior Leadership Team to respond to the changes that Ofsted had indicated they would make to the inspection framework to motivate schools to focus upon subjects that were not tested in national tests. The introduction of the Education Inspection Framework (EIF) in September 2019 marked a significant shift, emphasizing the quality of education, including the breadth and balance of the curriculum. Inspectors were asked to assess how well schools provided a diverse and comprehensive curriculum that extended beyond core subjects. This framework reduced the emphasis on test results and standardized testing, instead highlighting the importance of pupils receiving a well-rounded education that encompassed a variety of subjects and learning experiences (Ofsted 2019). Additionally, the EIF included separate judgements for personal development and behaviour and attitudes, (something that both schools had introduced into parent report cards already) as Ofsted wanted to recognise the importance of pupils broader development and their opportunities to grow as active, healthy citizens. Inspectors were told to evaluate the intent, implementation, and impact of the curriculum, examining how well it is planned and sequenced, how effectively it is taught, and the outcomes it achieves for all pupils. This EIF urged schools to divert time and resources away from test preparation and towards non-core subjects (Ofsted 2019). Concerns about the narrowing of the curriculum in primary schools have been raised by various educational researchers and institutions since national tests were introduced. Critics argued that the curriculum often focused too heavily on reading, writing, and mathematics, neglecting other subjects such as arts, humanities, and physical education. This overemphasis on core

subjects can limit students' overall development and creativity (Wyse and Torrance 2009). The pressure to perform well in standardised tests, such as SATs led schools to prioritise test preparation over a broader educational experience, resulting in reduced time for non-core subjects. There was also a concern that the curriculum did not provide a well-rounded education, sidelining subjects that foster critical thinking, problem-solving, and creativity (Berliner 2011). Critics argued that a narrow curriculum may not adequately prepare students for the complexities of the modern world, where a broad set of skills and knowledge is essential (Brundrett, Duncan and Rhodes 2010).

An important area for future research is the effects of granting schools more autonomy in designing their own curricula. Investigations should focus on the role of school leaders in driving curriculum innovation and research could examine how effective leadership practices facilitate the implementation of a broad and balanced curriculum, including the strategies they use to balance core and non-core subjects, and the impact of leadership on school culture and student achievement. Studies should also explore the professional development needs of school leaders to equip them with the skills and knowledge necessary to lead curriculum innovation. This could include case studies of schools that have successfully tailored their teaching to meet the diverse needs and interests of their students, providing a more balanced educational experience. Studies should also assess the effectiveness of providing teachers with the necessary training and resources to deliver a broad curriculum. This includes evaluating professional development programmes that focus on innovative teaching methods and interdisciplinary approaches, and identifying best practices for teacher training. Investigating the relationship between teacher autonomy and curriculum breadth is also crucial for understanding how empowering teachers can lead to a more balanced educational experience. Future research should explore whether there are different degrees of teacher autonomy and how this may influence the inclusion of non-core subjects and innovative teaching methods. Case studies of schools which give teachers varying degrees of autonomy can shed light on the benefits and challenges of granting teachers more control over curriculum and pedagogical decisions. Future research should examine how curriculum breadth influences teacher job satisfaction, stress levels, and professional fulfilment. Longitudinal studies that track the long-term impact of a narrowed curriculum on students' academic and personal development are essential. These studies should examine how a limited focus on core subjects affects students' critical thinking, creativity, and overall preparedness for secondary education and beyond. By following cohorts of students over several years, researchers can gain insights into the lasting effects of curriculum narrowing and identify key areas for intervention. Also, comparative studies that analyse curriculum practices in different countries can provide valuable insights into effective strategies for

maintaining a broad and balanced curriculum. By comparing the educational outcomes of students in countries with diverse curriculum approaches, researchers can identify best practices and potential pitfalls. These studies should consider cultural, economic, and policy differences to provide a comprehensive understanding of how various factors influence curriculum design and implementation. Future research should investigate the impact of curriculum breadth on the wellbeing of both students and teachers. Studies should explore how a diverse and engaging curriculum affects students' motivation, engagement, and overall mental health. One area of interest is the long-term benefits and practical implementation strategies of this shift towards arts, humanities, physical education, and other creative subjects into the daily school schedule. This research could provide insights into how such integration impacts student engagement and overall educational outcomes and how the AwL progress system incorporates these subjects. Understanding the relationship between curriculum design and wellbeing can inform policies and practices that support a healthy and productive learning environment. Additionally, investigations are needed into the outcomes of reducing the focus on standardised tests, such as SATs. This research could explore alternative assessment methods and their effectiveness in promoting a broader range of subjects and learning experiences, as well as their impact on student stress and performance. Future research should also evaluate the effectiveness of formative assessment strategies in non-core subjects and could include studies on how formative assessments can be implemented to improve educational progress and provide a more accurate measure of student learning. Research should focus on the effectiveness of formative assessment strategies in promoting a broad and balanced curriculum.

Future research should include a comprehensive policy analysis to identify the strengths and weaknesses of current curriculum policies. Researchers should evaluate the impact of national and local policies on curriculum breadth and propose recommendations for policy reform. This analysis should consider the perspectives of various stakeholders, including educators, policymakers, parents, and students, to ensure that proposed changes are practical and effective. Furthermore, research should explore the potential benefits of advocating for policy changes at the national level to support a more holistic approach to education. This could involve examining the impact of revised national curriculum and inspection guidelines that emphasise the importance of a well-rounded education on student outcomes and overall educational quality as the 2019 EIF has done. Future research should examine the role of parental and community involvement in supporting a broad curriculum. Studies should explore how schools can effectively engage parents and community members in curriculum planning and implementation. By understanding the impact of parental and community involvement on curriculum breadth, researchers can identify

strategies for fostering strong partnerships outside of the classroom to enhance the educational experience for pupils.

### 9.7 Further Research on Relationships between Headteachers and School Governing Bodies

This thesis demonstrates that the relationship between headteachers and school Governing Bodies (GB) significantly influenced the interpretation and implementation of AwL systems. Head-SPT collaborated closely with governors to select an AwL system that would address the data needs of the entire school. In contrast, Heads PAT and Passports involved governors based on clearly defined role boundaries, viewing the design of an AwL system as primarily the responsibility of teaching staff. Head-Passports emphasised that the decision was a matter of management and leadership rather than a strategic overview. SchoolPoints benefitted from having assessment experts on the governing body, which allowed Head-Points to work with these experts to disseminate information and provide training to other governors. This collaboration was facilitated by the expertise of the Chair of the GB, who was an Ofsted inspector and a headteacher, as well as another governor with experience in assessment at Oxford University Press. This expertise enabled these governors to take an early lead and skill up other members of the governing body. The influence of governors with a teaching background within the GB is evident, as they can support the headteacher in helping non-teaching governors understand the goals of the Senior Leadership Team (SLT). This support fosters a collegiate atmosphere between the headteacher and the governors. However, in Schools Qualitative and Continuum, the relationship between the SLT and GB was strained due to disagreements over the AwL approach proposed by the headteacher. Head-Qualitative excluded governors from the interpretation of AwL when they expressed a desire for regular test data to judge pupil progress, as their ideas did not align with the school's philosophy. Similarly, in SchoolContinuum, Head-Continuum attempted to maintain distinct role boundaries between the SLT and GB but was compelled to incorporate recommendations from Local Authority Advisors due to concerns about anticipated poor inspection judgements. The chair of governors continued to question the headteacher based on local authority and Ofsted expectations, which created tension within the school. This situation illustrates how the GB can insert themselves into the interpretation process even when not invited by the headteacher, particularly when they fear the judgement of Ofsted inspectors. The dynamics between headteachers and governing bodies are complex and can be influenced by the expertise of governors, role boundaries, and external pressures. Understanding these relationships is crucial for the effective implementation of AwL systems and overall school governance.

Future research could explore the impact of governor expertise on school leadership is a significant area for future research. The presence of governors with specialised knowledge, such as Ofsted inspectors or assessment experts, can profoundly influence decision-making processes and outcomes. Investigating how these experts contribute to the development and implementation of school policies, and identifying the training and development needs of governors to support headteachers effectively, could provide valuable insights into optimising governor recruitment. Secondly, the dynamics of role boundaries between headteachers and governors merit closer examination. It is uncertain whether clear role boundaries can facilitate effective collaboration and decision-making and blurred boundaries may lead to conflicts and misunderstandings or whether the opposite can be true. Research should explore how different approaches to defining and maintaining these boundaries impact the implementation of school policies and initiatives. Understanding the balance between leadership and strategic oversight can help in designing governance frameworks that promote harmonious and productive relationships. Another critical area for investigation is the involvement of governors in the design and implementation of assessment systems. Different models of governor involvement can lead to varying levels of acceptance and success among teaching staff. Analysing these models and their effectiveness can shed light on best practices for engaging governors in meaningful ways. Additionally, assessing the impact of governor involvement on the overall success of assessment systems can inform strategies for improving educational practices. Conflict and resolution strategies between headteachers and governors also require attention. This thesis documents that conflicts often arise from differing perspectives on assessment and accountability measures and research should identify other common sources of conflict and explore effective strategies for resolution. Understanding how to foster a collaborative environment despite these differences can enhance the overall functioning of school leadership teams and governing bodies. The influence of external pressures, such as local authority recommendations and Ofsted inspections, on the relationship between headteachers and governors is another important area for study. These pressures can shape decision-making processes and affect the dynamics between school leaders and governors. Investigating how headteachers and governors navigate these external demands while maintaining their school's vision and goals can provide insights into effective leadership and governance practices. Finally, conducting case studies of schools with particularly strong and effective relationships between headteachers and governors can offer valuable lessons. Identifying best practices and key factors that contribute to successful partnerships can inform the development of strategies to enhance collaboration in other schools. These case studies can serve as models for fostering positive and productive relationships between school leadership and governing bodies.

## 9.8 Further Research on Using Visual Methods during Elite Interviews

The interactive visual method employed in this thesis based on Antonucci's (1986) hierarchical mapping technique offered a novel approach to accessing headteachers' memories on how policies are interpreted and enacted over time by various stakeholders. This method, which involved headteachers using post-it notes on concentric circles to represent the influence of different stakeholders, served as both a memory recall aid and a stimulus for conversation. Understanding the cognitive processes involved in using visual methods is an important area for further research. Visual methods may help participants to access and articulate their memories and experiences in ways that traditional interviews do not. Researchers could investigate the cognitive mechanisms that underlie the use of visual methods, such as how visual aids trigger memory recall and how participants organize and interpret their thoughts visually. This line of research could provide insights into the psychological processes that make visual methods effective and inform the development of new visual techniques for research interviews. While this study demonstrated the effectiveness of visual methods in accessing headteachers' memories and generating rich data, it would be valuable to compare the outcomes of visual methods with those of traditional interviews. Researchers could conduct studies where participants are interviewed using both visual methods and traditional techniques to determine if there are significant differences in the depth and quality of the data collected. This comparison could help to establish the strengths and limitations of visual methods in different research contexts. Further research could also explore the impact of visual methods on participant engagement and the quality of the data collected. Visual methods, such as the one used in this study, may enhance participants' engagement by providing a more interactive and dynamic way to express their thoughts and experiences and so researchers could investigate whether participants find visual methods more engaging and whether this increased engagement leads to more detailed and insightful data. Additionally, studies could examine how different types of visual methods (e.g., drawings, diagrams, maps) influence participant engagement and data quality. Further research could also use visual methods to assess the long-term impacts of policy implementation and leadership approaches. This study focused on the initial and later stages of policy implementation, but it would be valuable to investigate how visual methods influence leaders' ongoing interpretation and enactment of policies over time. Researchers could conduct longitudinal studies to track how leaders' perceptions and actions evolve and how visual methods contribute to sustained policy implementation and organizational change. Future research could explore the application of visual methods in diverse educational contexts and with various types of leaders. While this study focused on headteachers and the implementation of a specific policy, subsequent studies could

investigate how visual methods can be utilised in interviews with other educational leaders and policymakers, in different educational settings. Additionally, researchers could examine the use of visual methods when conducting interviews with participants in other sectors, such as healthcare, business, and government, to understand how leaders in these fields interpret and enact policies.

## References

- Acocella, I., 2012. The focus groups in social research: advantages and disadvantages. *Quality & Quantity*, 46(4), pp.1125-1136.
- Adams, P. and Anderson, J., 2019. Moderation and the primary school context. *Education 3-13*, 47(1), pp.1-17.
- Adler, K., Salanterä, S. and Zumstein-Shaha, M., 2019. Focus group interviews in child, youth, and parent research: an integrative literature review. *International Journal of Qualitative Methods*, 18, p.1609406919887274.
- Aerts, D., Apostel, L., De Moor, B., Hellemans, S., Maex, E., Van Belle, H., Van Der Veken, J., 1994. *Worldviews: From fragmentation to integration*. Brussels, Belgium: VUB Press. Retrieved from <http://www.vub.ac.be/CLEA/pub/books/worldviews.pdf>
- Åkerblad, L., Seppänen-Järvelä, R., & Haapakoski, K., 2021. Integrative Strategies in Mixed Methods Research. *Journal of Mixed Methods Research*, 15(2), 152-170.
- Albright, K., Navarro, E.I., Jarad, I., Boyd, M.R., Powell, B.J. and Lewis, C.C., 2022. Communication strategies to facilitate the implementation of new clinical practices: a qualitative study of community mental health therapists. *Translational behavioral medicine*, 12(2), pp.324-334.
- Alexander, R., 2004. 'Still no pedagogy? Principle, pragmatism and compliance in primary education.' *Cambridge Journal of Education*, Vol. 34, No. 1, March 2004
- Allal, L., 2010. 'Assessment and the Regulation of Learning'. Available at: <http://unige.ch/fapse/people/allal/doc/Allal2010.pdf> (Accessed: 14 August 2017).
- Allal, L., 2013. 'Teachers' professional judgement in assessment: A cognitive act and a socially situated practice.' *Assessment in Education: Principles, Policy & Practice*, v20 n1 p20-3
- Alonzo, A. C. and Steedle, J. T., 2009. 'Developing and assessing a force and motion learning progression', *Science Education*. Wiley Subscription Services, Inc., A Wiley Company, 93(3), pp. 389–421. doi: 10.1002/sce.20303.
- Anderson, J. and Taner, G., 2022. Building the expert teacher prototype: A metasummary of teacher expertise studies in primary and secondary education. *Educational Research Review*, p.100485.

Andrade, H.L., 2013. Classroom assessment in the context of learning theory and research. SAGE handbook of research on classroom assessment, pp.17-34.

Antonucci, T. C. 1986. Measuring social support networks: Hierarchical mapping technique. *Generations*, 3, 10–12.

Apple, M.W., 2004. Schooling, markets, and an audit culture. *Educational Policy*, 18(4), pp.614-621.

Arnott, M.A., 2000. Restructuring the governance of schools The impact of 'managerialism' on schools in Scotland and England<sup>1</sup>. *The Governance of Schooling: Comparative Studies of Devolved Management*, p.52.

Arriaza, P., Nedjat-Haiem, F., Lee, H. Y., & Martin, S. S., 2015. Guidelines for conducting rigorous health care psychosocial cross-cultural/language qualitative research. *Social Work in Public Health*, 30(1), 75-87. <https://doi.org/10.1080/19371918.2014.93839>

Audulv, Å., Hall, E.O., Kneck, Å., Westergren, T., Fegran, L., Pedersen, M.K., Aagaard, H., Dam, K.L. and Ludvigsen, M.S., 2022. Qualitative longitudinal research in health research: a method study. *BMC Medical Research Methodology*, 22(1), p.255.

Auerbach, C., and Silverstein, L. B., 2003. *Qualitative Data: An Introduction to Coding and Analysis*, New York University Press. ProQuest Ebook Central.

Bailey, W.J., 1997. *Organizing schools (Vol. 4)*. R&L Education.

Baird, J., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T. and Daugherty, R., 2011. Policy effects of PISA.

Baird, J.A. and Elliott, V., 2018. Metrics in education—control and corruption. *Oxford Review of Education*, 44(5), pp.533-544.

Baird, J.A., Andrich, D., Hopfenbeck, T.N. and Stobart, G., 2017. Assessment and learning: Fields apart?. *Assessment in Education: Principles, Policy & Practice*, 24(3), pp.317-350.

Ball S. J., 2016. Following policy: Networks, network ethnography and education policy mobilities. In: *Globalisation and education*, pp. 126-143. Routledge.

Ball S. J., Maguire M, Braun A, Hoskins K., 2011. Policy actors: Doing policy work in schools. *Discourse: Studies in the cultural politics of education*, 32(4), pp.625-639.

Ball, S. J., Maguire, M., & Braun, A. 2012. *How Schools do Policy: Policy Enactments in Secondary Schools*. Routledge.

- Ball, S. J. and Bowe, R., 1992. Subject departments and the 'implementation' of National Curriculum policy: an overview of the issues. *Journal of curriculum studies*, 24(2), pp.97-115.
- Ball, S.J., 2003. The teacher's soul and the terrors of performativity. *Journal of education policy*, 18(2), pp.215-228.
- Barlow CM, Jolley RP, Hallam JL., 2011. Drawings as memory aids: Optimising the drawing method to facilitate young children's recall. *Applied Cognitive Psychology*, 25(3), pp.480-487.
- Barthes, R., 1974. *Le Plaisir du Texte* (translated by Miller, R. as *The Pleasure of the Text*', 1976). London: Cape.
- Battistone, M.J., Kemeyou, L. and Varpio, L., 2023. The Theoretical Integrative Review—A Reader's Guide. *Journal of Graduate Medical Education*, 15(4), pp.449-452.
- Baxter, P. and Jack, S., 2008. Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, 13(4), pp.544-559.
- Beard, K.S., 2025. Sensemaking and engagement in ethical decision-making (SEED): A United States decision-making framework for educational administration and school leadership. *Educational Management and Leadership*
- Beaton, A. E., 1996. *Handbook of Classroom Assessment, Handbook of Classroom Assessment*. doi: 10.1016/B978-012554155-8/50019-3.
- Becker, L.C., 1996. Trust as noncognitive security about motives. *Ethics*, 107(1), pp.43-61.
- Benford, R.D. and Snow, D.A., 2000. Framing processes and social movements: An overview and assessment. *Annual review of sociology*, 26(1), pp.611-639.
- Berkowitz, S., 1997. *Analyzing qualitative data. User-friendly handbook for mixed method evaluations*, pp.4-1.
- Berliner, D., 2011. Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge journal of education*, 41(3), pp.287-302.
- Berryhill, J., Linney, J.A. and Fromewick, J., 2009. The effects of educational accountability on teachers: are policies too stress provoking for their own good?. *International Journal of Education Policy and Leadership*, 4(5).

- Bevan, G., Hood, C., 2006. What's measured is what matters: targets and gaming in the English public health care system. *Public Adm.* 84(3), 517-538.
- Bijlsma, R.M., Bots, P.W., Wolters, H.A. and Hoekstra, A.Y., 2011. An empirical analysis of stakeholders' influence on policy development: the role of uncertainty handling. *Ecology and society*, 16(1).
- Black, P. and Wiliam, D., 1998. Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), pp.7-74.
- Bloom, B. S., 1971. Mastery learning. In J. H. Block (Ed.), *Mastery learning, theory and practice* (pp. 47-63). New York: Holt, Rinehart, and Winston.
- Bloom, B. S., Madaus, G. F., & Hastings, J. T., 1971. *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.
- Bloom, B.S., 1976. *Human characteristics and school learning*. McGraw-Hill.
- Bloom, B.S., 1984. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*, 13(6), pp.4-16.
- Bourque, L.B. and Fielder, E.P., 1995. Overview of self-administered questionnaires. In *How to conduct self-administered and mail surveys* (pp. 1-22). Sage Publication Inc., London.
- Bourque, L. B., Fielder, E. P., 2011. Overview of Telephone Surveys In: *How to Conduct Telephone Surveys*. Access Date: February 3, 2017. SAGE Publications, Inc. Print ISBN: 9780761925910. Online ISBN: 9781412984423. DOI: <http://dx.doi.org/10.4135/9781412984423>
- Bowe, R., Ball, S. J. and Gold, A., 1992. *Reforming Education and Changing Schools*. London: Routledge.
- Bozhovich, L. I., 2009. The social situation of child development. *Journal of Russian and East European Psychology*, 47, 59–86.
- Bozhovich, L. I., 2010. Zone of proximal development. The diagnostic capabilities and limitations of indirect collaboration. *Journal of Russian and East European Psychology*, 47,48–69.
- Braun, A. and Maguire, M., 2018. 'Doing without believing – enacting policy in the English primary school', *Critical Studies in Education*, 61(4), pp. 433–447. doi: 10.1080/17508487.2018.1500384.

- Braun, V. and Clarke, V., 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4), pp.589-597.
- Brighouse, T., 2011. Decline and fall: are state schools and universities on the point of collapse? Oxford Education Society First Annual Lecture ,Friday 16th September 2011, Lady Margaret Hall.
- Brinkmann, S. and Kvale, S., 2018. *Doing interviews (Vol. 2)*. Sage.
- Brinkmann, S., Kvale, S., 2015. Seven Stages of an Interview Investigation (Part II), in: Brinkmann, S. (Ed.), *InterViews: Learning the Craft of Qualitative Research Interviewing*. Sage, Los Angeles, London, New Delhi, Singapore, Washington DC.
- British Educational Research Association., 2011. *Ethical Guidelines for Educational Research*. London.
- Brown, M., Taggart, B., McCallum, B., Gipps, C., 1996. The impact of Key Stage 2 tests. *Educ. 3-13*, 24(3), 3-7.
- Brown, N. J. S. and Wilson, M., 2011. 'A Model of Cognition: The Missing Cornerstone of Assessment', *Educ Psychol Rev*, 23, pp. 221–234. doi: 10.1007/s10648-011-9161-z.
- Brundrett, M., Duncan, D. and Rhodes, C., 2010. Leading curriculum innovation in primary schools project: an interim report on school leaders' roles in curriculum development in England. *Education 3–13*, 38(4), pp.403-419.
- Bryan, H. and Burstow, B., 2018. Understanding ethics in school-based research. *Professional development in education*, 44(1), pp.107-119.
- Butler, R. and Shibaz, L., 2014. Striving to connect and striving to learn: Influences of relational and mastery goals for teaching on teacher behaviors and student interest and help seeking. *International journal of educational research*, 65, pp.41-53.
- Byrne, D., 2022. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & quantity*, 56(3), pp.1391-1412.
- Commission of Assessment without Levels (CAWL) Report., 2015. *Final report of the Commission on Assessment without Levels*.
- Caldwell, B. J., 2008. Reconceptualizing the self-managing school. *Educational Management, Administration and Leadership* 36(2): 235–252.

Calfee, R. C. and Masuda, W. V., 1996. 'Chapter 4 – Classroom Assessment as Inquiry', in Handbook of Classroom Assessment. doi: 10.1016/B978-012554155-8/50006-5.

Capra, F., 1996. The web of life: A new scientific understanding of living systems. New York, NY: Anchor Books.

Chandler, P. and Sweller, J., 1991. Cognitive load theory and the format of instruction. *Cognition and instruction*, 8(4), pp.293-332.

Chase LC, Schusler TM, Decker DJ., 2000. Innovations in stakeholder involvement: what's the next step?. *Wildlife Society Bulletin*, pp.208-217.

Chatterji, M., 2013. Insights, emerging taxonomies, and theories of action towards improving validity. In M. Chatterji (Ed.), *Validity and test use: An international dialogue on educational assessment, accountability and equity* (pp. 273–308). Bingley: Emerald Group.

Chowdry and Sibieta., 2011. Trends in education and schools spending. Institute for Fiscal Studies October 2011. Accessed Online 27th January 2017, <https://www.ifs.org.uk/bns/bn121.pdf>

Clarke, S. 2008. *Active learning through formative assessment*. London: Hodder Education

Coburn, C.E., 2001. Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational evaluation and policy analysis*, 23(2), pp.145-170.

Coburn, C.E., 2005. Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational policy*, 19(3), pp.476-509.

Coiffiat, L., 2015. Chapter in *The Gove Legacy and the Politics of Education after 2015 (3): Can the Legacy Endure?* 139

Coldron, J., Crawford, M., Jones, S. and Simkins, T., 2014. The restructuring of schooling in England: The responses of well-positioned headteachers. *Educational management administration & leadership*, 42(3), pp.387-403.

Colebatch, H. K., 2006. Thinking about policy: Finding the best way. Paper given at GovNet International Conference, Australian National University, Canberra.

Colman, A., 2021. School leadership, inspection and visibility: Panopticism and post-panopticism in an English coastal area of deprivation. *British Journal of Educational Studies*,

pp.1-17. Created from bham on 2020-12-23 05:55:01. Copyright © 2003. New York University Press.

Conley, L., 2020. *Mastery Learning: A Mixed-Methods Study of Teacher Efforts Implementing Key Components of the Mastery Learning Framework*. Gardner-Webb University.

Converse, J.M., & Presser, S., 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Thousand Oaks: SAGE Publications

Corcoran, T., Mosher, F. A. and Rogat, A., 2009. 'Learning Progressions in Science An Evidence-based Approach to Reform Nondiscrimination Statement'. Available at: [http://www.cpre.org/images/stories/cpre\\_pdfs/lp\\_science\\_rr63.pdf](http://www.cpre.org/images/stories/cpre_pdfs/lp_science_rr63.pdf) (Accessed: 14 August 2017).

Cresswell, J., 2013. *Qualitative inquiry & research design: Choosing among five approaches*.

Cundiff, P.R., McLaughlin, O., Brown, K. and Grace, K., 2020. In search of greater understanding: The impact of mastery learning on social science education. *Teaching Sociology*, 48(2), pp.107-119.

D'Souza, S., 2013. Commentary: Parent Feedback: A Critical Link in Improving Our Schools. *Public Administration Review*, 73(3), pp.413-414.

Department of Children, Schools and Families., 2009. *Getting to Grips with Assessing Pupils Progress*. Crown copyright 2009.

Department of Children, Schools and Families., 2010. *Independent Review of the Primary Curriculum: Final Report*. Available at <http://www.educationengland.org.uk/documents/pdfs/2009-IRPC-final-report.pdf> (Accessed on 14 November 2017)

Davies, J., Howells, K. and Jones, L., 2007. Evaluating innovative treatments in forensic mental health: A role for single case methodology?. *The Journal of forensic psychiatry & psychology*, 18(3), pp.353-367.

Davis, H. and Parker, T., 2014. *Rising Stars Assessment Progression Frameworks*.

Day, C. and Gu, Q., 2018. How successful secondary school principals in England respond to policy reforms: The influence of biography. *Leadership and Policy in Schools*, 17(3), pp.332-344.

Day, T., 2005. Teachers' craft knowledge: a constant in times of change? *Ir. Educ. Stud.* 24(1), 21-30.

De Decker, E., De Craemer, M., De Bourdeaudhuij, I., Wijndaele, K., Duvinage, K., Koletzko, B., Grammatikaki, E., Iotova, V., Usheva, N., Fernández-Alvira, J.M. and Zych, K., 2012. Influencing factors of screen time in preschool children: an exploration of parents' perceptions through focus groups in six European countries. *Obesity reviews*, 13, pp.75-84.

De Vaus, D. A., 2013. *Surveys In Social Research* (6th ed.). Hoboken: Taylor and Francis

Dennis, J., 2016. Michael Gove: the government's reformer in chief?,

Denzin, N.K., 1989. *Interpretive biography* (Vol. 17). Sage.

Department for Education., 2010. Academies Act. Department of Education. Accessed Online, 27th January 2017, <http://www.legislation.gov.uk/ukpga/2010/32/contents>

Department for Education., 2010b. Press release, Gove: 'Teachers, not politicians, know how best to run schools'. Accessed online on 22nd October 2017 <https://www.gov.uk/government/news/gove-teachers-not-politicians-know-how-best-to-run-schools>

Department for Education., 2015. Performance Descriptors for Key Stage 1 and 2 Statutory Teacher Assessment Government Consultation Response February 2015. Accessed Online, 27th January, [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/407178/Performance-descriptors-consultation-government-response.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/407178/Performance-descriptors-consultation-government-response.pdf)

Department for Education., 2017. Open Academies and Academy Projects awaiting approval, September 2017. Accessed online on 22nd October 2017 [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/646827/Open\\_academies\\_and\\_academy\\_projects\\_Sept\\_2017.ods](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/646827/Open_academies_and_academy_projects_Sept_2017.ods)

Department for Education., 2015. Common inspection framework: education, skills and early years from September 2015. Available at: <https://www.gov.uk/government/publications/common-inspection-framework-education-skills-and-early-years-from-september-2015>

Department of Children, Schools and Families & Qualifications and Curriculum Authority., 2010. The National Curriculum Level descriptions for subjects. First published in February 2010 © Crown copyright 2010. © Qualifications and Curriculum Authority

Donaldson, M.L., Mavrogordato, M., Youngs, P., Dougherty, S. and Al Ghanem, R., 2021. "Doing the 'real'work": How superintendents' sensemaking shapes principal evaluation policies and practices in school districts. *AERA Open*, 7, p.2332858420986177.

Drew, S. and Guillemin, M., 2014. From photographs to findings: Visual meaning-making and interpretive engagement in the analysis of participant-generated images. *Visual Studies*, 29(1), pp.54-67.

Duckworth, A.L., 2016. *Grit: The power of passion and perseverance*. Simon & Schuster.

Dweck CS., 2006. *Mindset: The New Psychology of Success*. New York: Balantine Books.

Dyer Jr, W.G. and Wilkins, A.L., 1991. Better stories, not better constructs, to generate better theory: A rejoinder to Eisenhardt. *Academy of management review*, 16(3), pp.613-619.

EES for Schools (undated). Target Tracker The Complete Assessment Solution. <http://www.eesforschools.org/targettracker>

Earle, S., 2019. Assessment in the Primary Classroom: Principles and practice. *Assessment in the Primary Classroom*, pp.1-120.

Eaude, T., 2015. Primary education: Can we escape the legacy of elementary education?. In *A Generation of Radical Educational Change* (pp. 43-54). Routledge.

Eaude, T., 2018. Re-humanising primary education—placing trust in teachers, learning from the legacy of Christian Schiller. *Education in Britain after the Coalition*, Ed Finn, M. Palgrave Macmillan.

Educational Endowment Foundation., 2017. *Mastery learning Moderate impact for very low cost, based on moderate evidence*.

Eisenhardt, K.M. and Graebner, M.E., 2007. Theory building from cases: Opportunities and challenges. *Academy of management journal*, 50(1), pp.25-32.

Eleventh Report of Session 2016–17. Accessed Online 30th May 2017 <https://www.publications.parliament.uk/pa/cm201617/cmselect/cmeduc/682/682.pdf>

Expert Panel Report., 2011. *The Framework for the National Curriculum. A report by the Expert Panel for the National Curriculum Review*. London: Department for Education.

- Eysenck, H. J., 1976. Introduction. In H. J. Eysenck (Ed.), *Case studies in behaviour therapy* (pp. 1-15). London: Routledge
- Feagin J, Orum AM, Sjoberg G (Eds.). 1991. *A Case for the Case Study*. University of North Carolina Press, Chapel Hill, NC.
- Federici, R.A., 2013. Principals' self-efficacy: Relations with job autonomy, job satisfaction, and contextual constraints. *Eur. J. Psychol. Educ.* 28, 73-86.
- Fenwick, T., 2010. Normalising standards in educational complexity. In: *Complexity Theory and the Politics of Education*, pp. 57-67.
- Filer, A., 2014. *Assessment: Social Practice and Social Product*, edited by Ann Filer. Taylor and Francis.
- Finn, M., 2015. 'The politics of education revisited: Anthony Crosland and Michael Gove in historical perspective', *London Review of Education*, 13(2).
- Fleer, M., 2015. Developing an assessment pedagogy: The tensions and struggles in re-theorising assessment from a cultural–historical perspective. *Assessment in Education: Principles, Policy & Practice*, 22(2), pp.224-246.
- Flyvberg, F., 2006. 'Five Misunderstandings About Case-Study Research'. *Qualitative Inquiry* Volume 12 Number 2 April 2006 219-245. Sage Publications  
10.1177/1077800405284363
- Flórez Petour, M. T., 2015. *Systems, ideologies and history: a threedimensional*
- Forester, J., 2012. On the theory and practice of interpretive policy analysis: From the micropolitics of practice to interpretive analysis and theorizing in action. Paper presented at the Interpretive Policy Analysis Conference, Tilburg.
- Frowe, I., 2005. Professional trust. *British Journal of Educational Studies*, 53(1), pp.34-53.
- Fullan, M., 2016. *The New Meaning of Educational Change* (5th ed.). Teachers College Press.
- Gahan, P. and Pekarek, A., 2013. Social movement theory, collective action frames and union theory: A critique and extension. *British Journal of Industrial Relations*, 51(4), pp.754-776.
- Gerring, J., 2004. What is a case study and what is it good for?. *American political science review*, 98(2), pp.341-354.

Gibb, N., 2015. Assessment after levels. From: Department for Education and The Rt Hon Nick Gibb MP. Part of: School and college qualifications and curriculum. Published: 25 February 2015

Gibb, N., 2015a. Reforming education through international evidence. From: Department for Education and The Rt Hon Nick Gibb MP. Part of: School and college qualifications and curriculum

Gibb, N., 2016. What is a good education in the 21st century? From: Department for Education and The Rt Hon Nick Gibb MP Part of: School and college qualifications and curriculum and Teaching and school leadership Published: 5 February 2016 Delivered on: 4 February 2016

Gibb, N., 2017. England's education reforms. From: Department for Education and The Rt Hon Nick Gibb MP. Part of: School and college qualifications and curriculum. Published: 11 April 2017

Gibb, N., 2017a. The importance of knowledge-based education. From: Department for Education and The Rt Hon Nick Gibb MP. Part of: Teaching and school leadership Published: 19 October 2017 Delivered on: 19 October 2017 (Original script, may differ from delivered version)

Gibson, J.E., 2012. Interviews and focus groups with children: Methods that match children's developing competencies. *Journal of Family Theory & Review*, 4(2), pp.148-159.

Giddens, A., 1982. *Profiles and critiques in social theory*. Univ of California Press.

Gioia, D.A. and Chittipeddi, K., 1991. Sensemaking and sensegiving in strategic change initiation. *Strategic management journal*, 12(6), pp.433-448.

Glegg, S.M., 2019. Facilitating interviews in qualitative research with visual tools: A typology. *Qualitative health research*, 29(2), pp.301-310.

Glesne, C., & Peshkin, A., 1992. *Becoming qualitative researchers: An introduction*. White Plains, NY: Longman Publishing Group.

Goldinger, S.D., Kleider, H.M., Azuma, T. and Beike, D.R., 2003. "Blaming the victim" under memory load. *Psychological Science*, 14(1), pp.81-85.

GOV.UK., 2015. Assessment Package for Primary Schools in Hackney. London Borough of Hackney. <https://www.contractsfinder.service.gov.uk/Notice/5fbaf114-5251-49c6-81a6-7a46688a2052?p=@RRPT0=NjJNT08=UFQxUI>

Grafanaki, S., 1996. How research can change the researcher: The need for sensitivity, flexibility and ethical boundaries in conducting qualitative research in counselling/psychotherapy. *British Journal of Guidance and Counselling*, 24(3), pp.329-338.

Greany, T., 2016. England's performance in TIMSS 2015: a 20 year story of improvement?

Grimble, R. and Wellard, K., 1997. Stakeholder methodologies in natural resource management: a review of principles, contexts, experiences and opportunities. *Agricultural systems*, 55(2), pp.173-193.

Grint, K., 2005. Problems, problems, problems: The social construction of 'leadership'. *Human relations*, 58(11), pp.1467-1494.

Greckhamer, T. and Cilesiz, S., 2014. Rigor, transparency, evidence, and representation in discourse analysis: Challenges and recommendations. *International journal of qualitative methods*, 13(1), pp.422-443.

Groundwater-Smith, S. and Sachs, J., 2002. The activist professional and the reinstatement of trust. *Cambridge journal of education*, 32(3), pp.341-358.

Gummesson, E. 1988. *Qualitative methods in management research*. Lund, Norway: Studentlitteratur, Chartwell-Bratt

Guskey, T.R., 2010. Lessons of mastery learning. *Educational leadership*, 68(2), pp.52-57.

Guskey, T.R. 2015. Mastery learning. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, pp. 752-759.

Gustafsson, J., 2017. Single case studies vs. multiple case studies: A comparative study.

Hamilton, L.S., Stecher, B.S. & Yuan, K., 2008. Standards-Based Reform in the United States: History, Research, and Future directions. Paper commissioned for the Center on Education Policy, Washington, D.C. For its project on Rethinking the Federal Role in Education.

Hammersley-Fletcher L, Strain M., 2011. Power, agency and middle leadership in English primary schools. *British Educational Research Journal*, 37(5), pp.871-884.

Hands, T., 2015. The Gove Legacy in Independent Schools: The Making and Unmaking of a Supreme Goviet. In *The Gove Legacy Education in Britain after the Coalition*, Ed Finn, M. Palgrave Macmillan

Hargreaves, A., 1994. Restructuring restructuring: Postmodernity and the prospects for educational change. *J. Educ. Policy*. 9(1), 47-65.

Hargreaves, D. H., 2014. A self-improving school system and its potential for reducing inequality, *Oxford Review of Education*, 40:6, 696-714.

Harlen, W., 2007. *Assessment of learning*.

Harrison, K., McNamara, G., & O'Hara, J., 2020. Evaluating the Level of Student and Teacher Involvement in Second-Level School Processes and Participation in Decision Making: An Irish Case Study. *Italian Journal of Sociology of Education*, 12(1), 102–121.

Hartley, J.F., 1994. Case studies in organizational research. *Qualitative methods in organizational research: A practical guide*, pp.208-229.

Harvey, G., 2015. The evolving model of school self-evaluation in Ireland: How a person's perception of purpose and power determines practice (Doctoral dissertation, National University of Ireland Maynooth).

Hatcher, W., 1990. *Logic and logos: Essays on science religion, and philosophy*. Oxford: George Ronald Press.

Hattie, J., 2009. *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J. and Timperley, H., 2007. The power of feedback. *Review of educational research*, 77(1), pp.81-112.

Hay, A., 2025. What may be: policy enactment in education, a new conceptual framework with actor-network theory. *Journal of Education Policy*, 40(2), pp.179-198.

Hawkes, T., 1977. *Structuralism and Semiotics*. London: Methuen.

Heale, R. and Noble, H., 2019. Integration of a theoretical framework into your research study. *Evidence-Based Nursing*, 22(2), pp.36-37.

Heifetz, R. and Linsky, M., 2002. *Leadership on the line: Staying alive through the dangers of leadership*. Cambridge, MA: Harvard Business School Press.

Heitink, M.C., Van der Kleij, F.M., Veldkamp, B.P., Schildkamp, K. and Kippers, W.B., 2016. A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational research review*, 17, pp.50-62.

- Hermanowicz, J.C., 2016. Longitudinal qualitative research. Handbook of the life course: Volume II, pp.491-513.
- Heyvaert, M., Hannes, K., Maes, B. and Onghena, P., 2013. Critical Appraisal of Mixed Methods Studies, *Journal of Mixed Methods Research*, 7(4): 302-327
- Higham, R. and Earley, P., 2013. School autonomy and government control: School leaders' views on a changing policy landscape in England. *Educational Management Administration & Leadership*, 41(6), pp.701-717.
- Higton, J., Leonardi, S., Choudhoury, A., Richards, N., Owen, D. and Sofroniou, N., 2017. Teacher workload survey 2016.
- Hogan S, Pink S, 2012. Visualising interior worlds. *Advances in visual methodology*, 230.
- Hood, C., 1991. A public management for all seasons?. *Public administration*, 69(1), pp.3-19.
- Hopfenbeck, T., Tolo, A., Florez, T. and El Masri, Y., 2013. Balancing trust and accountability? The assessment for learning programme in Norway: A governing complex education systems case study.
- Hopfenbeck, T. N., Flórez Petour, M. N., & Tolo, A., 2015. Balancing tensions in educational policy reforms: large-scale implementation of Assessment for Learning in Norway, *Assessment in Education: Principles, Policy & Practice*, 22:1, 44-60, DOI:
- Hopkins, P.E., 2007. Positionalities and knowledge: Negotiating ethics in practice. *ACME: an international journal for critical geographies*, 6(3), pp.386-394.
- House of Commons Education Committee Primary Assessment., 2017. Eleventh Report of Session 2016–17 Report, together with formal minutes relating to the report. HC 682 Published on 1 May 2017 by authority of the House of Commons.
- Hu, W., Adey, P., Jia, X., Liu, J., Zhang, L., Li, J. and Dong, X., 2011. Effects of a 'Learn to Think' intervention programme on primary school students. *British journal of educational psychology*, 81(4), pp.531-557.
- Huitt W., 2019. Analyzing paradigms used in education and schooling. *Educational Psychology Interactive*, 1–7.

Huitt, W., 2018. Understanding reality: The importance of mental representations. In W. Huitt (Ed.), *Twelve core ideas supporting holistic education* (pp. 65-81). La Vergne, TN: IngramSpark.

Hyman, L. Lamb, J. and Bulmer, M., 2006. The Use of Pre-Existing Survey Questions: Implications for Data Quality. *Proceedings of the 2006 European Conference on Quality in Survey Statistics*

Irving, A., 2011. Strange distance: Towards an anthropology of interior dialogue. *Medical anthropology quarterly*, 25(1), pp.22-44.

Irwin, L.G. and Johnson, J., 2005. Interviewing young children: Explicating our practices and dilemmas. *Qualitative health research*, 15(6), pp.821-831.

James, M., 2011. Assessment for learning: Research and policy in the (Dis) United Kingdom. *Assessment reform in education: Policy and practice*, pp.15-32.

Keddie, A., 2019. Context matters: Primary schools and academies reform in England. *J. Educ. Policy*. 34(1), 6-21.

Keddie, A., 2018. Supporting disadvantaged students in an English primary school: Matters of entrepreneurial and traditional professionalism. *Camb. J. Educ.* 48(2), 197-212.

Kennedy, M.M., 1979. Generalizing from single case studies. *Evaluation quarterly*, 3(4), pp.661-678.

Kennon, N., Howden, P. and Hartley, M., 2009. Who really matters?: A stakeholder analysis tool. *Extension Farming Systems Journal*, 5(2), pp.9-17.

Kezar, A. and Eckel, P.D., 2002. The effect of institutional culture on change strategies in higher education: Universal principles or culturally responsive concepts?. *The journal of higher education*, 73(4), pp.435-460.

Kim, J.H., 2016. In: *Understanding Narrative Inquiry: The Crafting and Analysis of Stories as Research*.

King, A. and Crewe, I., 2014. *The blunders of our governments*. Simon and Schuster.

Kington, A., Reed, N. and Sammons, P., 2014. 'Teachers' constructs of effective classroom practice: variations across career phases', *Research Papers in Education*. Routledge, 29(5), pp. 534–556. doi: 10.1080/02671522.2013.825309.

Kirton, A., Hallam, S., Peffers, J., Robertson, P. and Stobart, G., 2007. Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools. *British educational research journal*, 33(4), pp.605-627.

Klenowski, V. and Wyatt-Smith, C.M., 2013. *Assessment for education: Standards, judgement and moderation*.

Kline, P., 1998. *The new psychometrics: Science, psychology and measurement*. London: Routledge

Knight, J., 2008. 'The assessment for learning strategy', Department for Children, Schools and Families

Koppenjan, J.F.M. and Klijn, E.H., 2004. *Managing uncertainties in networks: a network approach to problem solving and decision making*. Psychology Press.

Kruckenberg, E., 2024. *Mastery Learning Approach with Formative Assessment Process to Encourage Student Success in Mathematics Classroom*.

Lamb, J., 2004. "Technical Help for researchers from ESRC and ONS", Labour

Landri, P., 2022. Waves of standardisation. In: *Educational Standardisation in a Complex World*, pp. 25-42. Emerald Publishing Limited.

Langer, E., 2002. Well Being: Mindfulness Verses Positive Evaluation. In Snyder, C.R., & Lopez, S.J. (Eds.), *Handbook of Positive Psychology* (p. 214-243). New York: Oxford University Press.

Lawn, M. & Grek, S., 2012. *Europeanizing Education governing a new policy space*

Levinson, B.A., Sutton, M. and Winstead, T., 2009. Education policy as a practice of power: Theoretical tools, ethnographic methods, democratic options. *Educational policy*, 23(6), pp.767-795.

Lewis, S., Savage, G.C. and Holloway, J., 2020. Standards without standardisation? Assembling standards-based reforms in Australian and US schooling. *Journal of Education Policy*, 35(6), pp.737-764.

Lightman, B., 2015. *The Gove Legacy in State Education*. In *The Gove Legacy*

Lunt, A., 2008. Assessment made visible: Individual and collective practices. *Mind, Culture, and Activity*, 15, 32–51.

Löfgren, H., Löfgren, R. and Pérez Prieto, H., 2018. Pupils' enactments of a policy for equivalence: Stories about different conditions when preparing for national tests. *European Educational Research Journal*, 17(5), pp.676-695.

Maguire M, Braun A, Ball S., 2015. Where you stand depends on where you sit: the social construction of policy enactments in the (English) secondary school. *Discourse: Studies in the Cultural Politics of Education*, 36(4), pp.485-499, DOI:10.1080/01596306.2014.977022.

Maitlis, S. and Lawrence, T.B., 2007. Triggers and enablers of sensegiving in organizations. *Academy of management Journal*, 50(1), pp.57-84.

Maitlis, S., 2005. The social processes of organizational sensemaking. *Academy of management journal*, 48(1), pp.21-49.

Mansell, W., 2017. Per pupil cost of MAT CEOs is eight times that of LA children's services directors. Available at <https://www.teachers.org.uk/education-policies/research/per-pupil-cost-of-mat-ceos-eight-times-more> (Accessed on 15 November 2017)

Masters, G.N., 2013. *Reforming Educational Assessment: Imperatives, Principles and Challenges*. Australian Education Review 57. Melbourne: Australian Council for Educational Research

Maxwell, G., 2002. *Moderation of Teacher Judgments in Student Assessment*. Brisbane.

May, T., 2001. *Our practices, our selves, or what it means to be human*. University Park, PA: The Pennsylvania State University Press.

May, T., and Sutton, C., 2011. *Social surveys: design to analysis*. In May, T. (ed). *Social Research*. Berkshire: Open University Press

McBride, J. R. & Wise, L. L., 2000. *Developing a Vertical Scale for the Florida Comprehensive Assessment Test (FCAT)*. Unpublished. Alexandria, VA: Human Resources Research Organization

McCann, E., and K. Ward. 2012. "Assembling Urbanism: Following Policies and 'Studying Through' the Sites and Situations of Policy Making." *Environment and Planning A* 44 (1): 42–51.10.1068/a44178

McConnell, A., 2010. Policy success, policy failure and grey areas in-between. *Journal of Public Policy*, pp.345-362.

- McCormick, R. and Murphy, P., 2008. Curriculum: The case for a focus on learning. *Learning and practice: Agency and identities*, pp.3-18.
- McDonald, R. P., 1999. *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates
- McGauren, A., 2015. Academy CEO Pay: How much do the biggest trusts pay?
- McGrew, S., Ortega, T., Breakstone, J. and Wineburg, S., 2017. The Challenge That's Bigger than Fake News: Civic Reasoning in a Social Media Environment. *American educator*, 41(3), p.4.
- Meyer, C.B., 2001. A case in case study methodology. *Field methods*, 13(4), pp.329-352.
- Miles, M.B. and Huberman, A.M., 1994. *Qualitative data analysis: An expanded sourcebook*. sage.
- Miller, M. D., Linn, R. L., & Gronlund, N. E., 2009. *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, NJ: Pearson Education
- Mitchell, J. Clyde., 1983. 'Case and situation analysis'. *The Sociological Review*.
- Murphy, G., 2019. Exploring principals' understandings and cultivation of leadership at all levels during initial teacher preparation school placement. *International Studies in Educational Administration* 47(2): 88–106.
- Murphy, P. and McCormick, R. eds., 2008. *Knowledge and practice: Representations and identities*. Sage.
- Nadoh J, Podreberšek P, Hlebec V., 2004 Cognitive evaluation of the hierarchical approach for measuring egocentered social networks. *Metodološki Zvezki*, 1, pp.379–393.
- National Association of Headteachers., 2014. *Report of the NAHT Commission on Assessment*.
- National College of Teaching and Leadership., 2014. *Beyond Levels: alternative assessment approaches developed by teaching schools Research Report* September 2014. Julie Lilly, Alison Peacock – The Wroxham transformative learning alliance Sue Shoveller, Dr d'Reen Struthers – South Farnham teaching school alliance. Crown Copyright 2014. Accessed Online 4th February 2017  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/349266/beyond-levels-alternative-assessment-approaches-developed-by-teaching-schools.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/349266/beyond-levels-alternative-assessment-approaches-developed-by-teaching-schools.pdf)

National Research Council., 2001. Knowing what students know: The science and design of educational assessment. National Academies Press. Pellegrino, J.W., Chudowsky, N., and Glaser, R. Editors

National Research Council., 2007. Taking Science to School: Learning and teaching science in grades K-8. Washington DC. National Academies Press.

Neale, B., 2021. Fluid enquiry, complex causality, policy processes: making a difference with qualitative longitudinal research. *Social Policy and Society*, 20(4), pp.653-669.

Novak, J. M., 2009. Invitational leadership. In B. Davies (Ed.), *The essentials of school leadership* (pp. 53-73). SAGE Publications Ltd, <https://www.doi.org/10.4135/9781446288290.n4>

O'Neill, O., 2013. Intelligent accountability in education. *Oxford review of education*, 39(1), pp.4-16.

O'Brien, S., McNamara, G., O'Hara, J., & Brown, M., 2019. Irish teachers, starting on a journey of data use for school self-evaluation. *Studies in Educational Evaluation*, 60, 1–13. <https://doi.org/10.1016/j.stueduc.2018.11.001>

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A., 2014. Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, 89(9), 1245e1251. <https://doi.org/10.1097/ACM.0000000000000388>

O'Brien, P. and Down, B., 2002. What are teachers saying about new managerialism?. *Journal of Educational Enquiry*, 3(1), pp.111-133.

OECD., 2014. PISA 2012 Results in Focus. What 15-year-olds know and what they can do with what they know. Available at <https://www.oecd.org/en/about/programmes/pisa/pisa-publications.html>

OECD., 2016, *Innovating Education and Educating for Innovation: The Power of Digital Technologies and Skills*, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264265097-en>. Available at <http://www.oecd.org/edu/cei/GEIS2016-Background-document.pdf>

OECD., 2016a. *Global competency for an inclusive world*. Available at <https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf>. Accessed on 14 November 2017

OECD., 2010. PISA 2009 Results: What Students Know and Can Do. Available at [https://www.oecd.org/en/publications/2010/12/pisa-2009-results-what-students-know-and-can-do\\_g1g114e2.html](https://www.oecd.org/en/publications/2010/12/pisa-2009-results-what-students-know-and-can-do_g1g114e2.html)

OECD., 2010a. PISA 2009 Results: Learning to Learn: Student Engagement, Strategies and Practices

Oates, T. 2015. Assessment: the need to 'do nothing'. Chapter in *A Generation of Radical Change: through the eyes of those who were part of it*. Ed Pring, R. & Roberts, M. Routledge, London.

Oerlemans, K., 2007. Students as stakeholders: Voices from the antipodes. *Journal of Educational Administration and History*, 39(1), pp.17-32.

Offe, C., 1999. How can we trust our fellow citizens. *Democracy and trust*, 52(1), pp.42-87.

Office for Standards in Education (Ofsted) 2019. Education Inspection Framework.

Ofsted, 2008. Assessment for learning: The impact of National Strategy support.

Olmos-Vega, F.M., Stalmeijer, R.E., Varpio, L. and Kahlke, R., 2023. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical teacher*, 45(3), pp.241-251.

Onwuegbuzie, A. J. & Teddlie, C., 2003. A framework for analyzing data in mixed methods research. In: A. Tashakkori and C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage, 351–383

Osterlind, S. J., 2010. *Modern measurement: Theory, principles, and applications of mental appraisal* (2nd ed.). Boston, MA: Pearson Education

Ozga, J. 2000. *Policy research in educational settings: Contested terrain*. Buckingham: Open University Press

Ozga, J., 2000a. Leadership in Education: the problem, not the solution?. *Discourse: studies in the cultural politics of education*, 21(3), pp.355-361.

Ozga, J., 2009. Governing education through data in England: From regulation to self-evaluation. *Journal of education policy*, 24(2), pp.149-162.

O'Brien S, McNamara G, O'Hara J, Brown M., 2019. Irish teachers, starting on a journey of data use for school self-evaluation. *Studies in Educational Evaluation* 60, 1-13. <https://doi.org/10.1016/j.stueduc.2018.11.001>.

Pal, L.A., 2005. Case study method and policy analysis. *Thinking like a policy analyst: Policy analysis as a clinical profession*, pp.227-257.

Pardoe, D., 2009. *Towards Successful Learning 2nd Edition*. A&C Black.

Parlett, M. and Hamilton, D., 1972. *Evaluation as Illumination: A New Approach to the Study of Innovative Programmes*. Occasional Paper No 9, Centre for Research in Educational Sciences, University of Edinburgh. Also published, abridged, in D. Hamilton, D. Jenkins, C. King, B. MacDonald and M. Parlett (eds) (1977) *Beyond the Numbers Game*. London:Macmillan

Department for Education., 2014. Performance Descriptors for use in key stage 1 and 2 statutory teacher assessment for 2015 /2016. Accessed online, 27th January 2017, [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/368298/KS1-KS2\\_Performance\\_descriptors\\_consultation.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/368298/KS1-KS2_Performance_descriptors_consultation.pdf)

Petersen, N. S., Kolen, M. J., & Hoover, H. D., 1989. Scaling, norming and equating. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221-262). New York, NY: Macmillan.

Poet, H., Sharp, C., Garry, J., Harland, J., Keightley, G., Kirkup, C. and Martin, K., 2018. *Assessment without levels: qualitative research*.

Popat, S., 2015. 'An Interpretive Phenomenological Analysis of Chapter 8 of the National Curriculum Review 2011'. Dissertation for the programme of Master of Arts in Educational Innovation, University of Warwick

Price Waterhouse Coopers., 2008. *Academies Evaluation: Fifth Annual Report*. Nottingham: Department for Children, Schools and Families Publications. Available at: [www.teachernet.gov.uk/publications](http://www.teachernet.gov.uk/publications)

Prince, R., 2010. Policy transfer as policy assemblage: making policy for the creative industries in New Zealand. *Environment and planning A*, 42(1), pp.169-186.

Quinn, R., Spritzer, E., 2006. 'Entering the Fundamental State of Leadership: A Framework for the Positive Transformation of Self and Others'. Working paper Series. Michigan Ross School of Business.

Reason, P., 1988. *Human inquiry in action: Developments in new paradigm research*. Sage.

Read, A. and Hurford, D., 2010. 'I know how to read longer novels'—developing pupils' success criteria in the classroom. *Education 3–13*, 38(1), pp.87-100.

- Reeves, S., Peller, J., Goldman, J. and Kitto, S., 2013. Ethnography in qualitative educational research: AMEE Guide No. 80. *Medical teacher*, 35(8), pp.e1365-e1379.
- Reinecke, J. and Ansari, S., 2015. What is a “fair” price? Ethics as sensemaking. *Organization Science*, 26(3), pp.867-888.
- Ricento, T.K. and Hornberger, N.H., 1996. Unpeeling the onion: Language planning and policy and the ELT professional. *TESOL quarterly*, 30(3), pp.401-427.
- Riese, H., Søreide, G.E., Hilt, L.T., 2022. Introduction: how to frame standards and standardisation in education. In: *Educational Standardisation in a Complex World*, pp. 3-23. Emerald Publishing Limited.
- Rignall, M. and Purnell, C. (undated). *Key Stage 3 Assessments & Target Setting – PIE*. Ilkley Grammar School.
- Rizvi, F. and Kemmis, S., 1987. Dilemmas of reform: An overview of issues and achievements of the Participation and Equity Program in Victorian schools 1984-1986. Deakin Institute for Studies in Education, Deakin University.
- Roberts, N., Bolton, P., 2017. *School funding in England. Current system and proposals for 'fairer school funding'*. House of Commons Library Briefing Paper Number 06702 11 January 2017
- Robinson, OC. and Smith, J. A., 2010b. 'A case study of pre-midlife crisis: an extraordinary transformational episode in an ordinary life', in D Robinson, P Fisher, T Yeadon-Lee, SJ
- Rogers, A., 2014. PISA, Power and Policy: the emergence of global educational governance. *International Review of Education*, 60(4), 591-596.
- Rosell-Aguilar, F., 2018. Twitter: A professional development and community of practice tool for teachers. *Journal of Interactive Media in Education*, 1.
- Rosenthal, R., & Jacobson, L., 1968. *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.
- Rotherham, A. J., & Willingham, D., 2009. *21st Century Skills: The Challenges Ahead*. Educational Leadership. September 2009 Volume 67 Number 1. Teaching for the 21st Century Pages 16-21

Rubie-Davies, C. M., Hattie, J., Townsend, M. A. R., and Hamilton, R. J., 2007. Aiming high: Teachers and their students. In V. N. Galwey (Ed.), *Progress in educational psychology research* (pp. 65–91). Hauppauge, NY: Nova.

Russell, M.K., 2010. Technology-aided formative assessment of learning: new developments and applications. In *Handbook of formative assessment* (pp. 125-138). Routledge.

Rycroft-Malone, J., Rogers, L. and Burton, C.R., 2022. Optimising the Conceptualisation of Context: Comment on " Stakeholder Perspectives of Attributes and Features of Context Relevant to Knowledge Translation in Health Settings: A Multi-country Analysis". *International Journal of Health Policy and Management*, 11(10), p.2365.

Sahlgren, G. H., 2014. *Handing Over the School Keys: The Impact of Privatisation on Education Quality*. Institute of Economic Affairs. Wiley Online

Sammons, P. & Davis, S., 2017. Mixed Methods Approaches and their Application in Educational Research Chapter 23 in Dominic Wyse, Neil Selwyn, Emma Smith and Larry E. Suter (2017) *The BERA/Sage Handbook of Educational Research Volume 1* London: Sage

Sandals, L., Bryant, B., 2014. 'The evolving education system in England: a temperature check. Isos Partnership' (2014). Accessed on 14th August 2017

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/325816/DFE-RR359.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/325816/DFE-RR359.pdf)

Santiago-Delefosse, M., Gavin, A., Bruchez, C., Roux, P., & Stephen, S. L., 2016. Quality of qualitative research in the health sciences: Analysis of the common criteria present in 58 assessment guidelines by expert users. *Social Science & Medicine*, 148, 142e151.

<https://doi.org/10.1016/j.socscimed.2015.11.007>

Savage, G.T., Nix, T.W., Whitehead, C.J. and Blair, J.D., 1991. Strategies for assessing and managing organizational stakeholders. *Academy of management perspectives*, 5(2), pp.61-75.

Sax, G., 1997. *Principles of educational and psychological measurement and evaluation* (4th ed.). Belmont, CA: Wadsworth.

Scanlon, D., 2022. *The Construction and Enactment of Curriculum Policy in Irish School Curriculum Reform: A Case Study of Leaving Certificate Physical Education* (Doctoral dissertation, University of Limerick).

Schoolsweek 5th June 2015. Accessed online on 22nd October 2017

<https://schoolsweek.co.uk/academy-ceo-pay-how-the-biggest-trusts-stack-up/>

Schreiber, J. B., 2012. Chapter Title: "Pilot Study" in *The SAGE Encyclopedia of Qualitative Research Methods* edited by Lisa M. Given. Published 2012 SAGE Publications, Inc.

Scottish Government. 2010. *Building the Curriculum 5 a Framework for Assessment: Quality Assurance and Moderation*. Glasgow: APS.

Simkins, T., 2000. Education reform and managerialism: comparing the experience of schools and colleges. *Journal of Education Policy*, 15(3), pp.317-332.

Siraj, I., Taggart, B., Melhuish, E., Sammons, P. and Sylva, K., 2014. *Exploring effective pedagogy in primary schools: Evidence from research*. London, Great Britain: Pearson (eng).

Skerritt, C., O'Hara, J., Brown, M., McNamara, G. and O'Brien, S., 2021. Enacting school self-evaluation: The policy actors in Irish schools. *International Studies in Sociology of Education*, pp.1-23.

Skinner B, Leavey G, Rothi D., 2021. Managerialism and teacher professional identity: Impact on well-being among teachers in the UK. *Educational Review* 73, 1-16.

Smaill, E., 2013. Moderating New Zealand's national standards: Teacher learning and assessment outcomes. *Assessment in Education: Principles, Policy & Practice*, 20(3), pp.250-265.

SIMS Assessment., 2017. Track pupil progress, whichever method you choose.

Snow, D.A. and Benford, R.D., 1988. Ideology, frame resonance, and participant mobilization. *International social movement research*, 1(1), pp.197-217.

Snow, D.A. and Benford, R.D., 1992. Master frames and cycles of protest. *Frontiers in social movement theory*, 133, p.155.

Snow, D.A., Rochford Jr, E.B., Worden, S.K. and Benford, R.D., 1986. Frame alignment processes, micromobilization, and movement participation. *American sociological review*, pp.464-481.

Spillane, J. P., 2004. *Standards deviation: How schools misunderstand education policy*, Harvard University Press, Cambridge, MA.

- Spillane, J.P., 2004. Educational leadership. *Educational evaluation and policy analysis*, 26(2), pp.169-172.
- Spillane, J.P., Reiser, B.J. and Reimer, T., 2002. Policy implementation and cognition: Reframing and refocusing implementation research. *Review of educational research*, 72(3), pp.387-431.
- Srivastava, P. and Hopwood, N., 2009. A practical iterative framework for qualitative data analysis. *International journal of qualitative methods*, 8(1), pp.76-84.
- Stake, R.E., 1995. *The art of case study research*. sage.
- Sternberg, R. J., 1994. Thinking styles: Theory and assessment at the interface between intelligence and personality. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 169–187). Cambridge University Press.
- Stobart, G., 2008. *Testing times: The uses and abuses of assessment*. Routledge.
- Stoecker, R., 1991. Evaluating and rethinking the case study. *The sociological review*, 39(1), pp.88-112.
- Sutton M, Levinson BA eds., 2001. *Policy as practice: Toward a comparative sociocultural analysis of educational policy* (Vol. 1). Greenwood Publishing Group.
- Swann, M., Peacock, A., Hart, S. and Drummond, M.J., 2012. *EBOOK: Creating Learning without Limits*. McGraw-Hill Education (UK).
- Sweller, J., 1989. Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of educational psychology*, 81(4), p.457.
- Teddlie, C. and Sammons, P., 2010. Applications of mixed methods to the field of educational effectiveness research. In B. P. M. Creemers, L. Kyriakides, and P. Sammons (Eds.) *Methodological advances in educational effectiveness research*. London: RoutledgeTaylor & Francis
- Teddlie, C. and Yu, F., 2007. Mixed methods sampling: A typology with examples. *Journal of mixed methods research*, 1(1), pp.77-100.
- Terry, G. and Hayfield, N., 2020. 38. Reflexive thematic analysis. *Handbook of qualitative research in Education*. Edward Elgar Publishing Limited, pp.430-441.

The Plowden Report., 1967. Children and their Primary Schools. A Report of the Central Advisory Council for Education (England). London: Her Majesty's Stationery Office 1967 © Crown copyright material is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland. Accessed Online 27th January 2017, <http://www.educationengland.org.uk/documents/plowden/plowden1967-1.html>

Thulasidas, M., 2008. Perception, physics, and the role of light in philosophy. *The Philosopher*, 96(2). Retrieved from <http://www.thephilosopher.co.uk/2008/05/the-role-of-light-in-philosophy-2008.html>

Thrupp, M., Lingard, B., Maguire, M., Hursh, D.W. and Peters, M.A., 2018. *The search for better educational standards: A cautionary tale*. Springer International Publishing.

Tidd, M., 2015. Plus ça change... plus c'est la même chose. *Ramblings of a Teacher*. With not even a hint of cohesion between posts, 26th May 2015.

<https://michaelt1979.wordpress.com/2015/05/26/plus-ca-change-plus-cest-la-meme-chose/>

Timmermans, S., Epstein, S., 2010. A world of standards but not a standard world: Toward a sociology of standards and standardization. *Annu. Rev. Sociol.* 36(1), 69-89.

Timperley, H.S. and Parr, J.M., 2009. Chain of influence from policy to practice in the New Zealand literacy strategy. *Research Papers in Education*, 24(2), pp.135-154.

Tolo, A., Lillejord, S., Flórez Petour, M.T. and Hopfenbeck, T.N., 2020. Intelligent accountability in schools: A study of how school leaders work with the implementation of assessment for learning. *Journal of Educational Change*, 21, pp.59-82.

Tomlinson, C. A., Gould, H., Schroth, S., & Jarvis, J., 2006. *Multiple case studies of teachers and classrooms successful in supporting academic success of high potential low economic students of color* (Research Monograph No. 06220). Storrs: The National Research Center on the Gifted and Talented, University of Connecticut.

Tomlinson, C. A., Moon, T. R. and Mcmillan, E. J. H., 2013. *Differentiation and Classroom Assessment*, SAGE Handbook of Research on Classroom Assessment. doi: 10.4135/9781452218649.n23.

Torrance, H., 2007. Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. , *Assessment in Education: Principles, Policy & Practice*, 14:3, 281-294,

Trent, A. and Cho, J., 2020. Interpretation in qualitative research: What, why, how.

Troman, G., 1997. Self-management and school inspection: Complementary forms of surveillance and control in the primary school. *Oxford Review of Education*, 23(3), pp.345-364.

Tsatsaroni, A., Koutsiouri, S., 2022. Global Agenda on Knowledge and Governance and Language Literacy Practices in Secondary Education in Greece. In: *Educational Standardisation in a Complex World*, pp. 63-83. Emerald Publishing Limited (2022)

Vangrieken, K., Meredith, C., Packer, T. and Kyndt, E., 2017. Teacher communities as a context for professional development: A systematic review. *Teaching and teacher education*, 61, pp.47-59.

van Haren, R. and Harroun, J., 2019. CGScholar's Analytics: progress to mastery learning. *Ubiquitous Learning*, 12(2), p.1.

Vannoni, M., 2015. What are case studies good for? Nesting comparative case study research into the lakatosian research program. *Cross-Cultural Research*, 49(4), pp.331-357.

Viennet, R. and Pont, B., 2017. Education policy implementation: A literature review and proposed framework.

Vogl, S., Schmidt, E.M. and Zartler, U., 2019. Triangulating perspectives: ontology and epistemology in the analysis of qualitative multiple perspective interviews. *International Journal of Social Research Methodology*, 22(6), pp.611-624.

Wang, W., Day, C., 2002. Issues and Concerns about Classroom Observation: Teachers' Perspectives.

Waslander, S., Pater, C., and van der Weide., 2010. Markets in education: an analytical review of

Weaver, R.K., 2010. Paths and forks or chutes and ladders?: Negative feedbacks and policy regime change. *Journal of Public Policy*, pp.137-162.

Webb, N.M., 2004. Classroom assessment as a research context: variations on a theme of pedagogical decision making. Trabajo presentado en ICME.

Weick, K.E., Sutcliffe, K.M. and Obstfeld, D., 2005. Organizing and the process of sensemaking. *Organization science*, 16(4), pp.409-421.

Wenger, E., 1999. *Communities of practice: Learning, meaning, and identity*. Cambridge university press.

West, A., 2015. Education policy and governance in England under the Coalition Government (2010–15): Academies, the pupil premium, and free early education. *London Review of Education* Volume 13, Number 2, September 2015

Widodo, H.P., 2014. Methodological considerations in interview data transcription.

Wiener, M.J., 1981. *English Culture and the Decline of the Industrial Spirit, 1850-1980*. Cambridge University Press

Wiggins, G.P., & Mctighe, J., 2005. *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development. <https://doi.org/10.14483/calj.v19n1.11490>

Wilby, P., 2013. Margaret Thatcher's education legacy is still with us – driven on by Gove. *The Guardian* 15th April 2013. Accessed online on 22nd October 2017  
<https://www.theguardian.com/education/2013/apr/15/margaret-thatcher-education-legacy-gove>

Wiley, J., Goldman, S.R., Graesser, A.C., Sanchez, C.A., Ash, I.K. and Hemmerich, J.A., 2009. Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, 46(4), pp.1060-1106.

Wiliam, D., 2010. Teacher quality: why it matters, and how to get more of it. Available at: [http://web.me.com/dylanwiliam/Dylan\\_Wiliams\\_website/Papers\\_04-11.html](http://web.me.com/dylanwiliam/Dylan_Wiliams_website/Papers_04-11.html) (Accessed 15 November 2017).

Wiliam, D., 2011. Bryggan mellan undervisning och lärande” (The bridge between teaching and learning) in the September 16, 2011 issue of *Pedagogiska Magasinet*,

Wilkins, A., 2015. Professionalizing school governance: the disciplinary effects of school autonomy and inspection on the changing role of school governors. *Journal of education policy*, 30(2), pp.182-200.

Wilkinson, S.D. and Penney, D., 2021. Setting policy and student agency in physical education: Students as policy actors. *Sport, education and society*, 26(3), pp.267-280.

Williams, J. S., 2003. Why Great Teachers Stay. *Educational Leadership*, 60(8), 71-74.

Willingham, D., 2017. Here's a 21st Century Skill--and How to Teach It! 10/2/2017. Available at <http://www.danielwillingham.com/daniel-willingham-science-and-education-blog> (Accessed on 14 November 2017)

Willis, J., 2016. Life After Levels: A Small Primary's Approach. Times Educational Supplement Teacher Toolkit. Accessed Online 27th January 2017, <http://www.teachertoolkit.me/2016/06/28/life-after-levels/>

Willis, J., McGraw, K. and Graham, L., 2019. Conditions that mediate teacher agency during assessment reform. *English Teaching: Practice & Critique*.

Wilson, M., 2009. Assessment from the ground up. *Phi Delta Kappan*, 91(1), pp.68-71.

Yanow, D inside Yanow, D. and Schwartz-Shea, P., 2014. *Interpretation and method: Empirical research methods and the interpretive turn*. Routledge.

Yin, R., 1993. *Applications of case study research*. Beverly Hills, CA: Sage Publishing.

Yin, R. K., 1989. *Case study research: Design and methods*. London: Sage Publications.

Yin, R. K., 2014. *Case Study Research Design and Methods (5th ed.)*. Thousand Oaks, CA: Sage. 282 pages. (ISBN 978-1-4522-4256-9)

Young, C. Savola, K. and Phelps, E. (eds)., 1991. *Inventory of Longitudinal Studies in the Social Sciences*, Newbury Park, CA: Sage

Zhang, Y. and Wildemuth, B.M., 2009. Unstructured interviews. *Applications of social research methods to questions in information and library science*, pp.222-231.

Zuckerman, S.J., 2019. Making sense of place: A case study of a sensemaking in a rural school-community partnership.