

THE WILDERNESS YEARS: AN ANALYSIS OF GOVES'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

*Andrew Chandler-Grevatt**

ABSTRACT

During Michael Gove's educational reforms between 2010–2014, he imposed several policy changes that changed the nature of assessment in terms of grading, terminal examinations and classroom expectations. Despite his vision of England rising up the international league tables, there has been little change in England's position and even signs of stagnation of attainment at upper secondary. This paper uses the Teacher Assessment Literacy in Practice (TALiP) framework to understand why the reforms associated with assessment have had little impact on attainment and reveals the devastating effect of such wholesale change to school assessment systems, without time or support to change, leaving teachers in a decade of assessment wilderness.

INTRODUCTION

Wilderness: If politicians or other well-known people spend time in the wilderness, *they are not in an influential position or very active in their profession* for that time.

A significant feature of the 2010 education reforms in England were changes to assessment. The reforms reduced the number of state-imposed examinations throughout the school system, overhauled what was to be examined in GCSE and A-level qualifications through a completely new national curriculum, and even removed the long established grading systems (national curriculum levels and GCSE grades) and replaced the GCSE letter grades (A*–G) with a numbered grading system (9-1). One politician drove these changes, Michael Gove in his role as Secretary of Education (2010–2014), and is understood to have taken a more active role in the changes than any other Education Secretary.

* Andrew Chandler-Grevatt EdD, MSc, BSc (Hons), PGCE, CSciTeach

THE WILDERNESS YEARS: AN ANALYSIS OF GOVE'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

Classroom assessment is a complex practice and is far more than just formal examinations: it can be seen as an ongoing process; a professional skill; and a set of skills that employ strategies to enhance teaching and learning (Black & Wiliam, 1998; Heritage & Harrison, 2020). Teacher assessment literacy in its simplest form is what teachers understand, know and can do regarding assessment (Stiggins, 1991). It can be used to support trainee teachers and established teachers to analyse their current understanding of assessment (Koh, 2011). In this paper, I employ a robust empirically based framework established by Xu & Brown (2016), Teacher Assessment Literacy in Practice (TALiP), to analyse the known impact of the neo-conservative policies on classroom assessment, teacher assessment identity, and assessment outcomes for this generation of school leavers.

GOVE'S ASSESSMENT POLICY

The interactions between policy, assessment and classroom practices are complex (e.g. Black & Wiliam, 2018; Stiggins, 2001). Following the Department for Education (2010) White Paper *Importance of Teaching*, Gove's rationale for world-leading national curriculum reform was set out by Oates (2011), unusually with a forward from the Secretary of State, Gove himself. Oates (2011) states his proposals are founded upon the highest performing jurisdictions, using international assessments as an evidence base on which to draw conclusions. In terms of assessment, the focus of the present paper, Oates' (2011, p. 126) main critiques of the, then, current national curriculum, were threefold: assessment was overbearing, with adverse impact on teaching and learning (evidenced by England's position in the international comparison tables); specific issues with drilling for tests; and lack of robust information for policymakers on national standards (particularly concerns over grade inflation).

In response to each of these criticisms, Gove made several reforms on England's assessment systems. To counter what was considered the overbearing assessment, GCSEs and A-levels moved from the established modular system to a terminal examination system. Meaning that there were no external assessment for learners from age 11 until age 16.

However, the claims made by the Coalition Government about the state of education in England were questioned. In a review by Oxford University curriculum and assessment academics, they showed that there was no evidence for England's decline in international tests, there is no evidence for GCSE grade inflation, raised doubts about the advantages of moving to linear examinations and the government claims about failures in teacher assessment (Baird et al., 2013). The case for the changes being evidence based was not as clear as Oates (2011) and Gove presented.

The translation of policy into practice is always problematic, but it helps if the policy is rolled out in a logical and supported way. Assessment and curriculum are closely entwined, so both need to be considered simultaneously. This was not the case in this reform, the structure of assessments were changed before the new curriculum was announced, and teachers started teaching GCSE and A-level course without any idea of the style and content of the examinations themselves (Mansell, 2012), almost as if assessment had no bearing on what and how teachers teach.

ASSESSMENT LITERACY

The term ‘Assessment Literacy’ appears to have been coined by Stiggins (1991) in his critique of teachers’ lack of knowledge of assessment and associated processes. Since then there has been a proliferation of studies into the development of, the features of, and the application of teacher assessment literacy. Although concerns about its validity as a concept have been muted (Popham, 2009), the consensus is that it is a useful concept through which to improve teachers’ knowledge and understanding of assesses in their practice (Hill *et al.*, 2010).

Xu and Brown (2016) extensively reviewed studies on assessment literacy from a thirty year period and offered a reconceptualisation of teacher assessment literacy in the form of the Teacher Assessment Literacy in Practice (TALiP) pyramid (Figure 1).

Their focus for their study was for the development of pre-service teachers and considered not just the knowledge base required to be an assessment literate

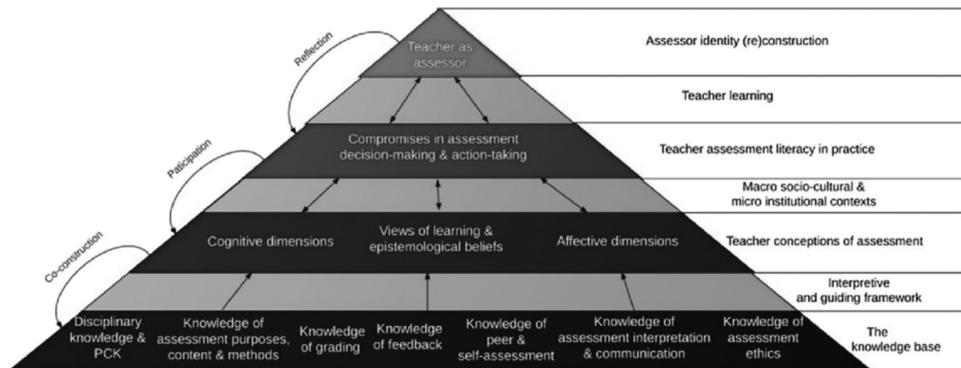


Figure 1: A conceptual Framework of teacher assessment literacy in practice (Xu & Brown, 2018)

THE WILDERNESS YEARS: AN ANALYSIS OF GOVES'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

teacher, but the interacting dimensions between that knowledge base, the sociocultural interactions, and teacher identity as assessor. As the authors explain:

TALiP is a dynamic, complex entity combining teachers' assessment knowledge, their conceptions of assessment, and their responses to the external contexts embedded with actual constraints and affordances in the environment (Xu & Brown, 2018, p.157).

The national curriculum changes from 2010 in England provide an interesting environment in which to analyse potential and actual impact on teacher assessment literacy in a changing policy landscape.

GOVE'S POLICY IMPACT ON STUDENT OUTCOMES

Before embarking on the analysis of changes in teacher assessment literacy, it is useful to ascertain what impact the education policy changes have had between on student outcomes 2014 and 2020.

As Gove envisioned, the education reforms would have England's school leavers being world-leading in qualifications, marching up the international comparison tables. However, indications so far would not suggest anything remotely close to this.

In terms of international comparisons, which are exalted by politicians, but are treated with caution by academics, the picture is mixed. Since the start of the policy changes in 2014, both Programme for International Student Assessment (PISA) and Trends in International Mathematical and Science Study (TIMSS) have produced a cycle of results.

The 2018 PISA tests 15 year olds were analysed by Sizmur *et al.*, (2019) and summarised that:

The mean scores in reading and science in England have not changed significantly over successive PISA cycles, but in mathematics, England's overall mean score showed a statistically significant increase compared with PISA 2015.

TIMSS results were published in 2019, giving an insight to international comparisons of Year 5 pupils and Year 9 pupils in Maths and Science. In their analysis Richardson *et al* (2020 p. 234) conclude that:

Overall, the 2019 TIMSS results saw an improvement in year 5 pupils' performance in mathematics, stability in year 9 mathematics and year 5 science, and a decline in year 9 performance in science.

The pupils involved in these tests had only had partial experience of the new curriculum in England from 2014. However, this remains a mixed picture and hardly the climb Gove anticipated from his policies.

In their recent analysis, Rogers & Spours (2000) call this the ‘great stagnation of upper secondary education.’ Highlighting this plateauing of attainment which is disproportionately affecting the middle to low attainers. This is supported by the Education Policy Institute (EPI) (Hutchinson et al, 2020) report that shows that since the coalition government’s policies were introduced that the attainment gap between disadvantaged pupils and their peers has stopped closing, and this started before the COVID-19 pandemic.

With this in mind, the following analysis using the TALiP framework, may go some way to explaining Gove’s policy failures.

ANALYSIS AND DISCUSSION

In this section I use the Xu & Brown (2016) TALiP framework to analyse and discuss the impact of education reforms instigated by Gove on teacher assessment literacy, drawing on academic literature and relevant grey literature. In the following analysis I consider the impact of the policy on each of these aspects of ‘The Knowledge Base’ and through the lenses of teacher conceptions of assessment and teacher assessment literacy in practice, the implications for teacher as assessor.

UNSETTLING THE FOUNDATIONS OF TALiP

Considering the seven foundations of assessment literacy in the TALiP framework (Figure 1), I contend that at least five of this have been changed fundamentally by Gove’s educational reforms, and the remaining two (Knowledge of Feedback and Knowledge of Peer & Self-assessment) have required a shift in professional knowledge of teachers enacting these policies. In this analysis I will focus on three foundational areas of TALiP: Disciplinary knowledge and PCK; Knowledge of assessment purposes, content and methods; and Knowledge of grading, and the repercussions on teacher assessment literacy with indications to why Gove’s educational policies have failed.

DISCIPLINARY KNOWLEDGE AND PCK

Gove not only changed assessment regime, but foundations of *what* is actually taught in classrooms. One of the most contested policy changes that of the national curriculum style and content (e.g. Beck, 2012). Persuaded by the cultural literacy

THE WILDERNESS YEARS: AN ANALYSIS OF GOVE'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

arguments of Hirsch (2007), (see Gordon, 2018) and the cognitive science informed theories of learning from Willingham (2009), the education policies morphed from Labour's more cross curricular and vocational curriculum to a knowledge based, academic discipline curriculum (Department for Education, 2010).

Pedagogical Content Knowledge (PCK) describes the knowledge and skills associated with teaching a discipline (Shulman, 1986) and has been used to understand and support subject knowledge and professional development of teachers (e.g. Park & Oliver, 2008). PCK is not just what teachers teach, but *how* they teach it. Disciplinary knowledge is an essential foundation in assessment literacy, as it informs how they teach. As Xu & Brown (2016 p.156) justify:

Since educational assessment is about measuring the curriculum content taught in schools/universities, knowledge of disciplines and how to teach that content cannot be excluded from the assessment knowledge base.

In light of the radical changes of what is being taught by teachers, we can explore some of the reported impact on teacher assessment literacy using the TALiP framework.

TEACHER CONCEPTIONS OF ASSESSMENT

What teachers teach is 'filtered and interpreted by teacher conceptions of assessment' (Xu & Brown, 2016 p. 156). It can be argued that Gove's changes to the curriculum deskilled even the most experienced teachers. Whether teachers were in favour or opposed to the change, the curriculum shift challenged the very nature of what they teach and the status of their discipline within the curriculum.

For Gove, his proposals for the curriculum change got off to a bad start. The introduction of synthetic phonics in primary teaching has been highly controversial and problematic in its implementation and outcomes (Carter, 2020). The core subjects of English, Mathematics and Science at secondary level were overhauled becoming narrower, more prescriptive and in the case of English questionable inclusion or exclusion of texts (Isaacs, 2014) and the proposed removal of English Literature, until the English teaching profession protested (Marshall, 2017).

History teachers felt devalued as Gove himself was rewriting their curriculum (Watson, 2019), without any professional or academic understanding of history education, curriculum or assessment. This was challenged by the profession and some eventual compromises were made (Harris & Burn, 2016).

Geography teachers felt the changes threatened their subject as a discipline (Lambert, 2013). The art subjects were not only threatened by the imposition of what they considered a poor model for their curriculum (Steers, 2014), but have

continued to feel the side-lining of their subject. For example, due to the introduction of English Baccalaureate (EBacc), Music education has felt a significant decline in its status and uptake (Bath *et al.*). From a wider sociological perspective, Young (2011) predicted the move to subjects in this way would lead to new inequalities in education.

For many teachers therefore, this curriculum change challenged their very epistemology of their discipline as a curriculum subject, overturned long held beliefs and had direct on their everyday practice and practices and created tensions between the Department for Education and professional teaching bodies. On all three aspects of the TALiP framework (Figure 1), most teachers were having tensions in the cognitive, epistemological and emotional domains.

COMPROMISES IN ASSESSMENT

Teachers were affected by the 2014 national curriculum for England, the change would have had an effect on their assessment practice and the compromises they make. Xu & Brown (2018 p. 157) explain:

Teachers' assessment decision making is a process by which teachers balance the demands of external factors and constraints with their own beliefs and values... TALiP is constantly negotiating between teachers' conceptions of assessment and the macro socio-cultural, micro institutional contexts and expected knowledge base, it reflects a temporary equilibrium reached among tensions.

Any single policy change imposes the need to rework professional knowledge and understanding, and its application to practice. This takes time, applying the TALiP framework, teachers needed to react to these changes (cognitively, epistemologically and emotionally), make compromises through decision making and action taking and use this learning to reconstruct their identity as an assessor.

Assessor identity: from defending to resenting

Between 2010 when curriculum changes were announced and 2014 when they were enforced, was a time of uncertainty. Many teachers were defending their discipline. From 2014 until the first GCSEs were sat in 2017, although the curriculum was established, teachers were still learning what it meant in relation to the other changes. Considering this aspect of curriculum change in isolation, this (re)construction of teacher identity as assessor could arguably be a complete 'construction' of identity due to the severity of the changes made. The changes in

THE WILDERNESS YEARS: AN ANALYSIS OF GOVES'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

disciplinary knowledge have particularly caused cognitive and epistemological tensions that needed reconciling in many teachers and associated with that will be affective tensions such as questioning their status, role as a teacher, role as assessor and the personal attacks from Gove himself on the status of the teaching profession and education academics (Lupton & Thomson, 2013). This was a particularly unusual aspect of policy change. The personal involvement of Secretary of State, the confrontational style to the very people who had to instigate that change (including a letter to *The Independent* newspaper from one hundred academics opposing the curriculum reform) and rather than appeasing the concerned professional workforce, Gove chose to attack them as the “enemies of promise”. This was followed by all four teaching unions calling a vote of no confidence in Gove. Not only did this curriculum reform deskill teachers, but the minister was attacking the profession and many of their epistemological beliefs. This emotional impact still resonates and many teachers hold resentment as part of their TALiP.

KNOWLEDGE OF ASSESSMENT PURPOSES, CONTENT AND METHODS

Having discussed the unsettling effects of changing the national curriculum content on teacher assessment literacy, this next foundational aspect of TALiP considers *how* that content is assessed. Xu and Brown (2018 p. 56) established that:

Teachers need to know how and why they assess (i.e., formative and summative), how different assessment methods can be related to the learning goals and specific content being learned, and what a variety of relevant assessment strategies are.

Again, these changes to formal assessments were wholesale, not just adjustments. Long established approaches to assessment at Key Stage 3 in the form of SATs and National Curriculum levels were abolished. I will discuss the impact of changes to grading in the next section. For this section, the focus is on the impact on teacher assessment literacy between 2010 and 2017 when the new style GCSE examinations were sat by students using the new grades 1–9, the impact of the removal of coursework, and the impact of terminal examinations dominating the assessment model. The stakes in high stakes assessment had become higher, on a much smaller evidence base (Torrance, 2018 p. 5).

By 2014, building on the White Paper (DfE, 2010), several changes started to be implemented, including the EBacc, and the teaching of GCSE subjects and A-level subjects. However, there was a period of time when the teaching of the

new curriculum had started but the new assessment purposes and methods had not. Teachers were given the frustrating situation of teaching their modified curriculum without any understanding of how it might be assessed (Mansell, 2014).

The removal of coursework may have left teachers with more time to teach content, but for practical based subjects, the disappearance has changed the nature of the subject itself. For example, in the sciences, practical coursework has become more and more controlled over the years to the point now, that it has been removed at GCSE and has become an add on to A-level Sciences (Childs & Baird, 2020). Practical work is a traditional part of science education and a much needed skill for future scientists.

TEACHER CONCEPTIONS OF ASSESSMENT

The changes to the examinations were wholesale and fast. Cognitively, teachers were kept in the dark for sometime about the exact nature of examinations: what they would assess and how they would assess it. It was particularly emotionally unsettling for teachers teaching a course before they understood the assessment, particularly when Ofqual and the Examination Boards were unable to give timely guidance. Epistemologically, many teachers were unsure or opposed to the new terminal assessments often due to accessibility for many learners. This was compounded by the strong accountability measures (also new) of the EBacc and Progress 8.

ASSESSOR IDENTITY: FROM DISEMPOWERED TO PLAYING THE GAME

Most GCSE and A-level subjects have sat their first round of new style examinations and it is only now that they can start reflecting. The TALiP framework illustrates how teachers' identify as assessor can change depending on current influences. In the current situation, I suspect many teachers identity is that of uncertainty and disempowerment: uncertain about the examinations, the types of questions, how their student may respond and unable to support their students with answers to the fundamental question of what the test will be like. It is only once those first examinations have been sat, the papers are revealed, and the results come out that teachers can feel more confident and more empowered to teach effectively with improved assessment literacy.

It is only now that teachers can start the processes of reflection, participation and co-construction from the TALiP framework (Figure 1). Only after the first sittings are teachers empowered to reflect on what was in the exam, how the students

THE WILDERNESS YEARS: AN ANALYSIS OF GOVE'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

responded and how to improve their teaching in response. I have equated 'washback' with these processes, which is the positive or negative effects of assessments on teaching (Cheng & Curtis, 2004). For example, a negative effect is 'teaching to the test' or worse the test becoming the teaching as in the case of phonics screening checks (Carter, 2020). Positive washback can be seen as modifications in teaching to promote improved learning, supporting students with exam technique, and understanding the different grade boundaries. In high accountability assessment and administrative regimes, washback can be skewed to more negative responses. The challenge for examination boards and education administrations is to write texts that are worth teaching to and avoid teachers 'playing the game.'

KNOWLEDGE OF GRADING

Gove's two main policies that affected grading were the removal of National Curriculum Levels for KS2 and KS3 (with no replacement) and the replacement of the GCSE A*-G system with grades 1-9 which were to be awarded by ranking the national cohort of students. None of these changes were a modification of existing policy, they were a seismic shift, a complete ground zero from which teachers, subject leaders and schools had to make sense of and build into their practice.

Immediately teachers were facing three significant changes to their assessment literacy. Firstly, their knowledge base had been removed of established grading systems, secondly, there was nothing to replace national curriculum levels at Key Stage 3 and thirdly there was little information on what the new GCSE grades meant until the year of the first examinations.

National curriculum levels had had several iterations in their history since 1988, and the most current form of levels was the application of them to a system called Assessing Pupil Progress. There were all criterion based, with level descriptors for various domains of each subject. It had merits for whole school assessment practices (Ofsted, 2011) and it suited some subjects better than others. Alongside the rise in Assessment for Learning in England (Black & Wiliam, 2003), national curriculum levels descriptors became a way of communicating progress in some cases lesson by lesson, or individual pieces of work and in some cases to a meaningless sub-level (a division of levels). This proliferation of using levels in this way raised a lot of criticism (Reay & Wiliam, 1999). In the background there has been a demonisation of criterion based assessment in England (Christodoulou, 2017) despite it being a valid and useful form of assessment in other jurisdictions (Panadero & Jonsson, 2013).

The void of Key Stage 2 and Key Stage 3 levels led to teachers and education publishers to invent their own assessment strategies to fill the vacuum. In 2014, when it was clear that the government were not going to replace levels there were

a number of consultations held by various unions and subject associations. For example for Primary education, the National Association of Headteachers commissioned a report on assessment (NAHT, 2014) and the government did release a report on Primary Assessment and Accountability (Department for Education, 2014), that controversially insisted on baseline assessment for Reception pupils (Robert-Holmes & Bradbury, 2016) but then leaves primary schools to do their own summative assessments throughout the Key Stage 1 & 2 until they do English and Maths SATs at the end of Key Stage 2. However, this void has been filled with a resurgence in using comparative judgement to assess primary writing (Wheadon et al, 2020).

In secondary schools, the lack of meaningful information from the Key Stage 2 SATs in English and Maths has led to many schools giving Year 7 pupils a secondary baseline test at the start of their secondary school career (e.g. AQA, 2021). In the chasm, some schools continued using levels for a number of years, more adventurous schools attempted their own quasi-level assessment model (Lilly *et al.*, 2014) and then there were a multitude of progress trackers at Key Stage 3.

We had moved from one model, with limitations, of which schools had a shared understanding. This could have been modified and its appropriate use have been supported with professional development of teachers. Instead, schools were left to invent their own tracking approach, often without suitable assessment literacy or buy in a package, which varied considerably in quality and assessment integrity. This led to a fragmented informal assessment system between schools, further losing the ability to communicate progress of individuals or groups between them.

A further consequence was that instead of filling the vacuum with another unknown, schools have attempted to extrapolate GCSE grading down to Key Stage 3. This has a variety of incarnations, but they all have significant flaws. This means that Year 7 pupils are being graded on final GCSE grades (that until recently had not yet been officially awarded). There was at least one diamond amongst the coal, the maligned and disgruntled music teachers got a team of assessment experts together to produce a bespoke assessment and progression framework using a criterion approach (Fautley & Daubney, 2014; 2019).

ASSESSOR IDENTITY: FROM DE-GRADED TO RESIGNATION

The impact on these changes on teacher conceptions of assessment was immediately catastrophic: all prior knowledge of the assessment systems, grading systems and exam systems was irrelevant. Teachers who were dissatisfied with national curriculum levels, would have been pleased with their removal, but there was no replacement. How could progress be assessed or measured at Key Stage 2 and Key Stage 3? Teachers who were comfortable and confident with using levels (some

THE WILDERNESS YEARS: AN ANALYSIS OF GOVES'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

had their whole career using them) would have been left reeling from the loss of what they perceived as a useful and workable system. There was no useful interpretive and guiding framework.

Using the TALiP framework, the impact of these changes on the cognitive, belief and affective domains meant that teacher conception of assessment was not just reduced, but eliminated. Together with the changes to the curriculum, teacher assessment identity had moved into the wilderness years between 2014–2020.

In a study of how these policies affected teachers' conceptions of assessment, Braun & Maguire (2018) give some insights into how primary teachers perceived the enactment of these policies. The pressure and uncertainty of policy change and the fact that it is at odds with teacher epistemologies, for example shifting the focus of teaching from individuals to targeted groups and second guessing policy, caused teachers to experience 'disaffected consent' (Gilbert, 2015), 'doing without believing.' (Braun & Maguire, 2018). Although this has been unresearched at Key Stage 3 and Key Stage 4, it is likely in that 2014–2020 period teachers at all key stages and in most subjects were feeling lost, disaffected and then resigned to taking on untried, unfit for purpose and assessment strategies.

Grading, culturally and pragmatically, has been the main mode of communication for teachers about their students' attainment, progress and potential. This shared language was removed, making teachers feel de-professionalised in assessing, predicting and communication of these important aspects with their peers, their students and the parents of those students. Teachers were on a steep learning curve with little support from the government. The (re)construction of 'teacher as assessor' could be argued as a catastrophic change in identity, from being informed and knowledgeable, to being in the dark and incapable of decision: totally degraded, followed by a resigned compliance.

IMPACT OF COVID-19

The global pandemic has forced significant changes in education, particularly the cancellation of examinations in GCSE and A-level. In the examinations of Summer 2020, students were unable to sit their examinations due to national lockdown measures. Instead, teachers were asked to provide a grade for their students in each subject (Ofqual, 2020). This brought about significant uneasiness in the profession, with schools asking examination boards and the government for support and guidance (Jadhav, 2020). The solution was far from satisfactory, with an algorithm that randomly assigned grades to students, causing huge dismay for students, teachers and parents (Paulden, 2020).

In England, teachers regularly make predictions for GCSE grades and A-level Grades, but when assigning actual attainment at the end of these qualifications,

they lack confidence. This could be accounted for by lack of training in assessment literacy: an understanding of validity and reliability skills in moderation; and appropriate sources of evidence. However, I contend that is mainly due to the reliance on examinations as the only form of assessment. Usually in the form of mock examinations that use previous exam papers.

The forced cancellation of examinations has opened a flaw in this approach to assessment of qualifications and exposed teachers need for improved assessment literacy, so that they, with professional confidence, supported with appropriate evidence, make professional and moderated judgements about their students' attainment at any time.

CONCLUSIONS

One politician was able to exert his influential position to make sweeping changes to educational policy, in doing so, he has professionalised, disempowered and degraded the teaching profession, casting the profession into an assessment wilderness for a decade.

Gove intended to make ambitious changes to education in England, positioning us in the top-performing countries. So far, there is little evidence that any of the policy changes have had any impact on improving standards in education. In this paper, I have explored the impact of his changes to assessment policy to explain why these policies have failed, using the TALiP framework.

In the rationale for change (Oates, 2011) what was seen as overbearing external assessment has been replaced with so few and such high stakes assessment, that the assessment system has become too fragile (Torrance, 2018). With regard to the concerns of the negative impact of assessment on teaching and learning, it is clear that the way these reforms have been managed has exacerbated that and led the education system into a period of darkness and disorientation with regard to teaching, learning and assessment. Even though there have been some positive developments such as subject organisations developing their own assessment systems, this is fragmented and out of necessity rather than strategy.

Our children and their education deserved better than this, and we need to ensure that governments do not allow a single person to exert such power. Education, teaching and learning and assessment are too complex to allow someone with more confidence than competence to make such destructive changes. We need to ask why this can happen, and prevent it happening again.

The TALiP framework exposes the complexities of the impact of change to assessment and the considerations needed for making such changes. Teachers need to be part of the process of educational change on cognitive, epistemological and affective levels. Change needs to be managed, not imposed. Working with the

THE WILDERNESS YEARS: AN ANALYSIS OF GOVES'S EDUCATION REFORMS ON TEACHER ASSESSMENT LITERACY

profession, trusting the profession and developing the profession are essential aspects of change. In addition, assessment needs to be seen as important as teaching, learning and curriculum. It needs investment in time, finance and recognition in policy to make meaningful and positive impact on children's education.

REFERENCES

- AQA (2021) *Key Stage 3 Maths Baseline Tests*, <https://allaboutmaths.aqa.org.uk/KS3baselinetest> (Accessed 12th February 2021) (Accessed 12th February 2021).
- Baird, J. A., Ahmed, A., Hopfenbeck, T., Brown, C., & Elliott, V. (2013). Research evidence relating to proposals for reform of the GCSE.
- Bath, N., Daubney, A., Mackrill, D., & Spruce, G. (2020). The declining place of music education in schools in England. *Children & Society*, 34(5), 443–457.
- Beck, J., (2012). Reinstating knowledge: diagnoses and prescriptions for England's curriculum ills. *International Studies in Sociology of Education*, 22(1), pp.1–18.
- Black, P., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British educational research journal*, 29(5), 623–637.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.
- Braun, A. & Maguire, M. (2018) Doing without believing – enacting policy in the English primary school, *Critical Studies in Education*, 1–15.
- Carter, J. (2020). The assessment has become the curriculum: Teachers' views on the Phonics Screening Check in England. *British Educational Research Journal*, 46(3), 593–609.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. *Washback in language testing: Research contexts and methods*, 27, 3–17.
- Childs, A., & Baird, J. A. (2020). General Certificate of Secondary Education (GCSE) and the assessment of science practical work: an historical review of assessment policy. *The Curriculum Journal*, 31(3), 357–378.
- Christodoulou, D. (2017). *Making good progress?: The future of assessment for learning*. Oxford: Oxford University Press.
- Collins Dictionary (2020) <https://www.collinsdictionary.com/dictionary/english/in-the-wilderness> (Accessed 12th February 2021)
- Department for Education (2014) Reforming assessment and accountability for primary schools. Government response to consultation on primary school assessment and accountability. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/297595/

- Primary_Accountability_and_Assessment_Consultation_Response.pdf (Accessed 12th February 2021).
- Department for Education (2010). White Paper: Importance of Teaching <http://www.educationengland.org.uk/documents/pdfs/2010-white-paper-teaching.pdf> (Accessed 12th February 2021).
- Fautley, M., & Daubney, A. (2014). The national curriculum for music: An assessment and progression framework. *Incorporated Society of Musicians*. https://www.ism.org/images/images/ISM_The-National-Curriculum-for-Music-booklet_KS3_2019_digital.pdf (Accessed 12th February 2021).
- Gilbert, J. (2015). Disaffected consent: That post-democratic feeling. *Soundings*, Summer, 60, 29–41
- Gordon, J. (2018). Reading from nowhere: assessed literary response, Practical Criticism and situated cultural literacy. *English in Education*, 52(1), 20–35.
- Harris, R., & Burn, K. (2016). English history teachers' views on what substantive content young people should be taught. *Journal of Curriculum Studies*, 48(4), 518–546.
- Hill, M., Cowie, B., Gilmore, A., & Smith, L. F. (2010). Preparing assessment-capable teachers: What should preservice teachers know and be able to do?. *Assessment Matters*, 2, 43–64.
- Hirsch, E. D. (2007). *The knowledge deficit: Closing the shocking education gap for American children*. New York: Houghton Mifflin Harcourt.
- Huchinson, J., Reader, M. & Akhal, A. (2020) Education in England: Annual report 2020. Education Policy Institute. <https://epi.org.uk/publications-and-research/education-in-england-annual-report-2020/> (Accessed 12th February 2021).
- Isaacs, T. (2014). Curriculum and assessment reform gone wrong: The perfect storm of GCSE English. *Curriculum Journal*, 25(1), 130–147.
- Jadhav, C. (2020). Arrangements for Summer 2020. <https://ofqual.blog.gov.uk/2020/04/09/arrangements-for-summer-2020/> (Accessed 12th February 2021).
- Lambert, D. (2013). Collecting our thoughts: school geography in retrospect and prospect. *Geography*, 98(1), 10–17.
- Lilly, J., Peacock, A., Shoveller, S., & Struthers, D. R. (2014). *Beyond levels: Alternative assessment approaches developed by teaching schools*. The National College for Teaching and Leadership. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/349266/beyond-levels-alternative-assessment-approaches-developed-by-teaching-schools.pdf
- Lupton, R. & Thomson, S. (2015) *The Coalition's Record on Schools: Policy, Spending and Outcomes 2010–2015*. University of Manchester. <https://sticerd.lse.ac.uk/dps/case/spcc/wp13.pdf> (Accessed 12th February 2021).

THE WILDERNESS YEARS: AN ANALYSIS OF GOVES'S EDUCATION
REFORMS ON TEACHER ASSESSMENT LITERACY

- Mansell, W. (2012). A race to the bottom?. *Phi Delta Kappan*, 94(4), 76–77.
- Marshall, B. (2017). The politics of testing. *English in Education*, 51(1), 27–43.
- NAHT (2014) Report of the NAHT Commission on (Primary) Assessment – February 2014 https://www.cfey.org/wp-content/uploads/naht_report_-_assessment.pdf (Accessed 12th February 2021).
- Oates, T. (2011), “Could do better: using international comparisons to refine the National Curriculum in England”, *The Curriculum Journal: Reviewing the National Curriculum 5–19 Two Decades On*, vol. 22, no. 2, pp.121–150.
- Ofqual (2020). Awarding Qualifications in Summer 2020. <https://www.gov.uk/government/publications/awarding-qualifications-in-summer-2020>. (Accessed 12th February 2021).
- Ofsted (2011) Report summary: The impact of the ‘Assessing pupils’ progress’ initiative. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/379103/Summary_20-_20The_20impact_20of_20the_20Assessing_20pupils_20progress_20initiative.pdf (Accessed 12th February 2021).
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational research review*, 9, 129–144.
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in science Education*, 38(3), 261–284.
- Paulden, T. (2020). A cutting re-mark. *Significance*, 17(5), 4–5.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental?. *Theory into practice*, 48(1), 4–11.
- Reay, D., & Wiliam, D. (1999). ‘I’ll be a nothing’: structure, agency and the construction of identity through assessment. *British educational research journal*, 25(3), 343–354.
- Richardson, M., Tina Isaacs, T, Barnes, I., Swensson, C., Wilkinson, D., & Golding, J. (2020) *Trends in International Mathematics and Science Study (TIMSS) 2019: National report for England*, Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/941351/TIMSS_2019_National_Report.pdf (Accessed 12th February 2021).
- Roberts-Holmes, G. P., & Bradbury, A. (2016). “They are children... not robots, not machines”: The introduction of Reception baseline assessment. UCL Institute of Education. <https://discovery.ucl.ac.uk/id/eprint/1476041/1/baseline-assessment-final-10404.pdf> (Accessed 12th February 2021).

- Rogers, L., & Spours, K. (2020). The great stagnation of upper secondary education in England: A historical and system perspective. *British Educational Research Journal*, 46(6), 1232–1255.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, 15(2), 4–14.
- Sizmur, J., Ager, R., Bradshaw, J., Classick, R., Galvis, M., Packer, J., Thomas, D. & Wheater, R. (2019). *Achievement of 15-year-olds in England: PISA 2018 results: Research report*, December 2019. National Foundation for Educational Research.
- Steers, J. (2014). Reforming the school curriculum and assessment in England to match the best in the World—A cautionary tale. *International Journal of Art & Design Education*, 33(1), 6–18.
- Stiggins, R. J. (1991). Assessment Literacy. *Phi Delta Kappan*, 72(7), 534–39.
- Koh, K. H. (2011). Improving Teachers' Assessment Literacy through Professional Development. *Teaching Education*, 22(3), 255–276.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5–15.
- Torrance, H. (2018). The return to final paper examining in English national curriculum assessment and school examinations: Issues of validity, accountability and politics. *British Journal of Educational Studies*, 66(1), 3–27.
- Watson, M. (2019). Michael Gove's war on professional historical expertise: conservative curriculum reform, extreme whig history and the place of imperial heroes in modern multicultural Britain. *British Politics*, 1–20.
- Wheadon, C., Barmby, P., Christodoulou, D., & Henderson, B. (2020). A comparative judgement approach to the large-scale assessment of primary writing in England. *Assessment in Education: Principles, Policy & Practice*, 27(1), 46–64.
- Willingham, D. T. (2009). *Why don't students like school?: A cognitive scientist answers questions about how the mind works and what it means for the classroom*. New York: John Wiley & Sons
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.
- Young, M. (2011). The return to subjects: A sociological perspective on the UK coalition government's approach to the 14–19 curriculum. *Curriculum Journal*, 22(2), 265–278.