

‘[It] isn’t designed to be assessed how we assess’: rethinking assessment for qualification in the context of the implementation of the Curriculum for Wales

Elizabeth Titley^{a,b,*}, Andrew James Davies^a and Stephen Atherton^a

^a*Aberystwyth University, UK;* ^b*Birmingham City University, UK*

This paper reports teacher and learner perspectives on how assessment and reform influences pedagogical practices and behaviours. The research was conducted in a context of policy reform, at a time when Wales’ revised General Certificate of Secondary Education (GCSE) specifications had been implemented, and learners were preparing for their assessments; but, also during the period of debate on the development of Wales’ new curriculum, which has taken a distinct and contrasting position on assessment to the assumptions underlying the reform of Welsh GCSEs implemented from 2015. These data, therefore, offer unique insights into the affordances and limitations of two sharply contrasting systems at a time of considerable change, offering reflections on the current curriculum and its attendant assessment practices, and also a prospective analysis of how the principles embedded in the new curriculum could challenge these existing assumptions and conventions. Findings suggest that teachers and learners currently inhabit an assessment-driven system, which encourages performative practices in pedagogy and is governed by external accountability; and that these practices are at odds with the principles of assessment articulated in *Successful Futures*. Consequently, teachers in this study expressed uncertainty about how assessment for certification purposes at GCSE could be compatible with the principles of the Curriculum for Wales.

Keywords: assessment; accountability; reform; performativity; Curriculum for Wales

Introduction

It has recently been argued that the history of Welsh education since devolution in 1999 can be characterised by three distinct policy ‘movements’ (Egan, 2017; Davies *et al*, 2018; Connolly *et al*, 2018), each of which have taken specific positions on the linked

*Corresponding author. Department of Childhood, Youth and Community, Room SCT230 City South Campus, Birmingham City University, Westbourne Road, B15 3TN, UK. Email: elizabeth.titley@bcu.ac.uk

© 2020 The Authors. *The Curriculum Journal* published by John Wiley & Sons Ltd on behalf of British Educational Research Association

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

issues of agency, assessment and accountability. Davies *et al.* (2018) provide an overview of these three movements, noting that the early years of the National Assembly for Wales witnessed an ‘experimental’ approach in the newly created policy ‘laboratory’ of Welsh devolution (Reynolds, 2008; Moon, 2012); this period being marked by a loosening of external accountability, with the abolishment of SATs and school league tables (Davies *et al.*, 2018; Connolly *et al.*, 2018). Power (2016) further suggests that the Welsh approach since devolution has been based on a ‘high trust’ rather than ‘mistrust’ approach to the teaching profession (Davies *et al.*, 2018). Yet, from 2010 onwards a second phase of policymaking was initiated: poor PISA results in 2009 prompted a reappraisal, by the then Minister for Education Leighton Andrews, of the approach taken during the first decade of devolution. In his much-cited speech of February 2011, Andrews described education in Wales as a ‘complacent system’ and signalled a ‘new approach to accountability’ (Andrews, 2011; Davies *et al.*, 2018). Davies *et al.* (2018) have argued that this period leveraged greater external accountability and challenge into the Welsh system, citing key developments such as the introduction of national testing for 7 to 11-year old pupils, the consolidation and increased centralisation of school improvement services, and the introduction of performance-related ‘banding’ (subsequently ‘categorisation’) of schools from 2011 onwards.

This period also saw a review and radical reform of qualifications, with Wales moving away from the ‘three-country’ regulatory settlement on General Certificate of Secondary Education (GCSE)s which had existed between Wales, England and Northern Ireland since 1988 (Andrews, 2014, p. 169; Barrance & Elwood, 2018, p. 252; Barrance, 2019). At this time of heightened external accountability, in Welsh education, the reform to GCSE specifications which followed this review also arguably injected greater externality into assessment procedures for some subjects. This was achieved through limiting controlled assessment, and aligning pupil outcomes with skills measured by PISA (Estyn, 2018), which has itself been described as a key measure of system-wide external accountability (Egan, 2017, p. 4). Since 2016, the third policy ‘movement’ identified by Davies *et al.* (2018) and Egan (2017) has signalled a shift back towards ‘trust’ (Power, 2016), and a proposed relaxation of some of these measures of external accountability (National Assembly for Wales, 2018). There has also been a commitment that assessment policies and practices should be ‘uncoupled’ from high-stakes accountability and external monitoring (Welsh Government, 2019, pp. 38–39), in the context of the implementation of the new curriculum, which has been under development since a landmark review of the Welsh curriculum conducted by Professor Graham Donaldson in 2015. The extent to which assessment for certification can realistically be ‘uncoupled’ from curriculum development a delivery remains a major concern.

The development of the new Curriculum for Wales is broadly line with a wider shift to what Shapira and Priestley (2018, p. 75) term a new ‘genericism’ in curriculum design, where curricula are increasingly being designed to be outcome- and competency-focussed (see Donaldson, 2015, pp. 5–6). Donaldson argues that, going forward, articulations of progression in the Welsh curriculum should encompass a broader range of outcomes than just ‘narrower measures of attainment’, noting that outcomes should be described from the learner’s point of view, with assessment conducted using a ‘a wide repertoire of assessment techniques matched to the different

outcomes' (Donaldson, 2015, pp. 56, 79). He further recommends a move towards long-term reference points, developed as achievement outcomes for each of the progression steps in the new curriculum: progression is to be measured at three yearly intervals, allowing teachers' scope to plan learning and assessments. The introduction of progression steps will remove reference to the Foundation Phase and Key Stages as part of the new curriculum (Welsh Government, 2019). Progression steps will be introduced at ages 5, 8, 11, 14 and 16, and while statutory guidance has not yet been published, research conducted as part of the CAMAU project (Hayward *et al.* 2017) notes that in each Area of Learning and Experience achievement outcomes will be developed as broad expectations, rather than a prescriptive list of knowledge and skills. The CAMAU project also suggest that progression of learning should be used in a more formative manner and incorporate assessment for learning to support pupils, and their teachers, to make pedagogical decisions on how they can be supported to improve. Achievement outcomes will, therefore, encompass a series of statements describing achievements by pupils at each progression step, and teachers will decide which of the statements captures pupil performance most appropriately (Welsh Government, 2019). The Welsh Government (2019) note that this approach will empower schools to develop their own approaches to assessment and learning. The Welsh Government (2019) has further suggested that, under the new curriculum arrangements, Head teachers will be empowered to set these achievement outcomes against each of the new progression steps.

As noted above, Donaldson's review is also vocal about the impact that high-stakes and external assessment, focussed on small range of outcomes, can have on the breadth of the curriculum, most notably when it bears a close relationship with measures of external accountability (Donaldson, 2015, pp. 6, 75 and 79). The review also calls for teacher assessment to be the 'main vehicle' for assessment before qualification (Donaldson, 2015, p. 80), although it acknowledges the issues of validity, reliability and alignment with external benchmarks, which have been a cause of concern in Wales for some time (Andrews, 2011), and calls for support for teachers and robust moderation to address these.

The reform of GCSE specifications in Wales and the debate over alignment with the new curriculum

Given such a direction of travel with regard to assessment policy and practice, it is therefore inevitable, and generally acknowledged, that further reform of GCSE qualifications will be required to ensure alignment between the principles which have guided the development of the new curriculum, and those which will form the basis of certification and its associated assessment norms and practices at (what is currently) Key Stage 4 (Qualifications Wales, 2019).

By contrast, the reforms to GCSE in Wales, implemented in 2015 during the second phase of devolved policymaking centred on developing more 'rigorous' qualifications (Andrews, 2014, p. 170), leading to substantive changes in course and assessment structure (Barrance & Elwood, 2018); changes which would now appear to be significantly at odds with the principles of assessment espoused in the new curriculum. From 2015, tiering was retained only in exceptional cases, limits were

placed on controlled assessment, with a greater dependence on external examinations (Barrance & Elwood, 2018). For example, the English Language GCSE specification was adapted so that only 20% of the marks are awarded for controlled assessment (a reduction from 40% in previous specifications) with the remaining 80% awarded for externally set examinations (Welsh Government, 2014).

Public-facing external accountability with regard to GCSE grades, it has been suggested, also led to a disproportionate emphasis being placed on the outcomes of narrow groups of learners on the C/D borderline, focussing much of the attention of teachers on pupils who would raise school attainment figures with marginal improvements in their performance (Welsh Government, 2019). Such an emphasis can clearly produce problematic consequences, such as failing to encourage all learners to reach their full potential (Donaldson, 2015), or in promoting surface learning, and encouraging pedagogical practices which focus on test-specific measures (Koretz, cited in Ofqual, 2016). Much of the challenge for the implementation of the Curriculum for Wales will be the move from high-stakes assessment to a system that focusses more on the needs of individual pupils through formative assessment and assessment for learning (Donaldson, 2015; Welsh Government, 2019). And, while a wide range of evidence can be called upon to suggest that moves towards foregrounding formative assessment may well benefit learners (See Dann, 2002; Bennett, 2011; Black, 2015; Hopfenbeck & Stobart, 2015), there are clearly concerns regarding the validity of such assessments (Andrews, 2011; Stobart, 2012), as well as the ways in which significant changes such as these will be implemented in practice in schools (Priestley & Minty, 2013). This process could well be further problematized through the staggered implementation of the new curriculum beginning in 2022 with learners in Year 7 who will consequently complete Year 11 studies in 2027 (Qualifications Wales, 2019). This will lead to dual systems of accountability and measurement working in tandem in schools until the new curriculum is fully implemented.

GCSE assessment in Wales: pressure and performativity?

Elwood, Hopfenbeck and Baird (2017) note that high-stakes summative assessments can increase pressure on young people, obscuring the 'purpose' of schooling as learning, and reframing their educational experience as being geared towards 'performing' in examinations. Consequences of such an ethos can include greater assessment-related anxiety (Sommer & Arendasy, 2015), a tendency to adopt performance goal orientations (von der Embse *et al.*, 2018) and a trend towards surface-level learning rather than deep-level learning (Vandewalle *et al.*, 2019). As regards teachers, the literature on teacher orientations and experiences in such environments points towards performative practices being routinely employed, including teaching to the test by narrowing the curriculum and prioritising assessed content (Popham, 2000; Segal *et al.*, 2017), as well as the excessive use of past papers as an assessment preparation technique (Firestone *et al.*, 2004; Crocker, 2005). These practices, which can be attributed to the coupling of assessment outcomes and accountability for teachers and schools as a result of standards-based reforms, have resulted in systems where assessments exercise a disproportionate influence on what is taught in the classroom

(Brown, 2004). Furthermore, literature on teachers' practices has also examined how high-stakes assessment, and externally mandated systems of quality assessment, accountability and governance not only serve to encourage performative and strategic instrumentality in pedagogical practice, but also pose a very real threat to the 'identity' of teachers as agentic professionals and ethical subjects, empowered to make decisions based on their professional judgement (Ball, 2003; Holloway & Brass, 2018).

Materials and methodology

This paper reports on qualitative data gathered during the second phase of a wider, explanatory sequential mixed-methods study (Creswell & Plano Clark, 2011), the aims of which were to examine stakeholders' experiences and perspectives on the reformed GCSE and A-Level specifications in Wales, and to evaluate the impact of these reforms on classroom practices. The study also specifically sought perspectives on how the new curriculum under development, and its position on assessment, would likely influence further reform of GCSE qualifications, and consequently teachers' and pupils' assessment practices and learning behaviours—namely, the focus of this paper. Data were collected from four curriculum pioneer secondary schools in Wales, over a period of 12 months between 2017 and 2018. The schools which participated were, therefore, highly attuned to the complexities and challenges of developing the new curriculum, and embedding its principles in their arrangements for delivery, pedagogy and assessment. Schools involved in this phase of the research were selected using a purposive sampling method (Cohen *et al.*, 2018), with the sample generated to gather perspectives from a diverse range of such settings across all regions of Wales, encompassing both Welsh- and English-medium perspectives. The data presented in this paper are part of a wider research project that included questionnaires, interviews and focus groups with a range of key stakeholders, namely pupils, parents and teachers. Overall, 45 teachers, 244 pupils and 107 parents provided data as part of the study. To ensure the validity of research instruments a process of rigorous piloting was conducted. The pilot study involved a sample of 25 pupils and 5 teachers, who participated in pilot interviews/focus groups and questionnaires. The responses to the pilot study were evaluated for reliability and validity and research instruments were adapted prior to the completion of the main study.

The data presented in this paper were collected via 10 focus groups of GCSE and A-Level pupils (a total of 33 pupils), and 15 semi-structured interviews with classroom teachers. Pupils and teachers were also invited to participate based on a purposive sampling strategy, with the intention of obtaining a range of perspectives from both GCSE and A-Level pupils in different year groups, and from teachers of a range of subjects. Table 1 identifies the number of participants from each of the four data collection settings. The data presented in this paper are drawn from a relatively small subset of the main data set, and while we do not claim that the data can be generalizable, the data are offered as an exploration of situated insights that may have a wider application.

Focus groups were deemed appropriate for data collection with pupils as they allowed for the research agenda to be driven by issues that pupils perceived to be important to their experiences (Cohen *et al.*, 2018), and to allow for dynamic social interactions (Morgan, 1996) where perspectives could be co-constructed, challenged,

Table 1. Distribution of participants

	Number of teachers interviewed	Focus groups	
		Number of focus groups	Number of pupils
School 1	5	2	7
School 2	5	2	7
School 3	0	4	13
School 4	5	2	6

explored or refined through interaction. Focus groups were also considered to have the affordance of being a comfortable and informal context for pupils to articulate their views (Darbyshire *et al.*, 2005). By contrast, semi-structured interviews allowed teachers to explore, in a professional discursive space, issues relating to assessment that they considered to be important, and allowed for individualised, in-depth accounts of their own experiences (Newby, 2013). The qualitative data were analysed using a progressive coding procedure, starting with basic *in vivo* and descriptive codes, which identified commonalities in the original words and descriptions used by participants. The second stage of the coding procedure was to generate composite codes, where these commonalities were systematically combined and categorised. In these stages, pupil data and teacher data were analysed and coded separately, allowing the opportunity for distinct perspectives to be offered by these two groups and to encourage authenticity in the representation of differing perspectives. The final stage of the coding procedure was to develop more abstract codes designed to build theory and draw inferences (Braun & Clarke, 2006; Saldaña 2009). It was, at this stage, that the teacher data and pupil data were directly compared, to allow for a more complete and theoretical understanding of the experiences of both teachers and pupils in relation to the reform experience.

As noted above, the data gathered during this study were collected at a very specific juncture of reform, when the post-2015 GCSE specifications had been implemented for the first time, and when the schools involved in the study were also deeply involved in planning their future provision in response to the principles of new curriculum. As such, the data offer a unique ‘bifurcated’ insight on two highly contrasting systems: reflections *in situ* on the current curriculum and its attendant assessment conventions, and the impact these have on pedagogy; and, also an *ex ante* analysis of how the principles of the new curriculum could pose challenges to assessment practice.

Results

Inhabiting an assessment-driven system

Throughout the research, pupils showed a keen awareness of the assessment-driven climate within which they were learning, and of the performative practices that this engendered. They spoke extensively of their experiences of a system, which many felt reduced ‘learning’ to a narrow suite of circumscribed assessed outcomes, particularly

during the later years of their education as they progressed towards assessment for certification:

Pupil L: *I think as well, the way you're taught when you get to like GCSE/A level, it's obviously just to pass those assessments, it's just to pass exams.*

Pupils also spoke critically of what they deemed to be the 'manufactured' nature of some exams. Many felt that the full range of their subject-specific knowledge or skills was not being appropriately assessed; rather, they felt that they were being assessed on their ability to replicate contrived answers and re-enact answering strategies they had rehearsed previously in class:

Pupil B: *I don't think the exam system in general portrays your knowledge of a subject, – it portrays your ability to replicate or to [...] execute what you have been told to execute.*

Pupil A: *[exams] can be very manufactured.*

Reflections from teachers also supported a view that the Welsh education system had moved towards engendering an assessment-driven ethos. Teachers identified increasing 'top-down' pressure to meet and exceed attainment targets, and accountability measures and inspection procedures as increasingly influential in prioritising assessment outcomes, above other indicators and outcomes of pupil learning:

Teacher B: *the focus again is on the sausage factory and the outcomes, not the learner.*

Teacher A: *the time you are taking to test, you used to take to teach. So, if you're testing and not teaching, they're learning less.*

Assessment, performativity and preparation

Furthermore, pupils also elaborated on what many saw as the excessive use of assessment-specific preparation methods used in schools, most notably teaching to the test and the frequent use of past papers. They spoke of, in some cases, years of strategic preparation for GCSE assessments, starting as early as Year 8 for some pupils:

Pupil G: *when it comes to the actual [science] assessment, I would say it would probably be more or less every lesson.*

Pupil I: *I feel like English give you a lot of time... and they started preparing us for maths I would say as soon as we went back [in the new academic year in September].*

Pupils in every focus group cited past papers as a primary method used by schools to prepare pupils for assessments. They spoke of an excessive cycle of 're-doing [and] re-doing' 'past paper after past paper after past paper', and had begun to identify that they felt the use of past papers were 'the only way [they] get through' their assessments:

Pupil B: *Past papers, revision sessions, revision lessons, erm, working during free lessons on revising subjects through erm past papers mainly.*

Pupil H: *if you have a past paper, they will say 'right do this by Monday' and then you mark it and you go over it, so whatever you do wrong, you go over it and you know how to improve, and then you just re-do it, re-do it until you get it.*

One teacher outlined a relentless regime of practicing using past papers and engaging in a strategic ‘guessing game’ as to the likely content of each paper, immediately prior to an exam:

Teacher B: *first week back [in September], we tested them, so that was a maths lesson gone... this week we are back after half term, erm, they had a paper Monday morning, they had maths Monday afternoon anyway. [...] Wednesday morning they had an exam and Thursday lunch time was an additional session instead of being in lunch and Thursday tutor was another session in additional maths. Friday morning they had an exam, Monday morning I'm hoping to get a paper emailed out; I've literally sent all the year 11 parents' email addresses to myself, get a paper sent to them over the weekend, probably tomorrow so they can look at it Sunday. Because I'm doing a – erm, predicted, I don't want to say predicted, erm, practice paper, you know, topics that haven't been in paper 1, we look at where they will arise in paper 2, so you know, that will go out let's – hopefully on Saturday for the three tiers [...] Then Monday morning at 8.30 we are starting early and we will focus on probable misconceptions and errors that the kids have for that, at 20 past 9 they have to go into the exam hall for 10 minutes of you know, getting settled down for starting [their formal exam] at 9.30.*

As noted above, pupils were also very self-aware about such performative practices which involved teaching to the test, the prioritising and foregrounding of content likely to appear in the assessments, and preparation routines within which they were engaged with their teachers. In fact, some pupils were among the first cohort to take the reformed GCSEs, and many felt that their learned reliance on regimes of preparation using past papers had left them in a difficult situation, guessing the likely areas which would appear in the assessments but with limited precedents from which to take their cue:

Pupil X: *But there is no [past papers], and it's like the teachers are guessing, like 'it could come up, I think this might come up' but then it's hit or miss, well will that sort of question come up? Nobody knows. So it's hard to prepare.*

When reflecting on their own practices some teachers also felt that they had little choice but to prioritise content that they thought would be assessed because of external standards and accountability measures:

Teacher E: *think the curriculum is perhaps dictated a little bit – not a little bit let's be honest, it's dictated an awful lot by the measures in which we are judged.*

Teacher B: *Can we get through this in one lesson you know? [...] you don't actually go into any in-depth analysis, yes, you're scratching the surface, you know, and you lose the depth and the enrichment that you might like [...] there's no pedagogical understanding there or knowledge, it's literally procedural. Why do I do it? You don't need to know, we have 15 minutes, let's get through this now, you know, we have got to do another topic.*

Experiences of reforms

As noted above, some pupils also articulated their perception that they were disadvantaged by the changes, owing to a lack of assessment precedents to study. And, while Barrance and Elwood (2018, p. 255) note that examination results can often dip suddenly following assessment and qualification reform, and improve gradually thereafter, our study makes no claims about the validity of this perception as

expressed by participants. However, we do note that a lack of past papers for assessment did seem to make it difficult for some pupils and teachers to engage in the familiar preparation regimes outlined above:

Pupil K: *for the new specs, we had like one past paper which was like, it was problematic... the teachers were more stressed, [they] didn't know what to assess in class, like practice and things.*

Pupil AF: *there wasn't resources in abundance.*

Teachers also identified the relationship between classroom teaching and assessment content, with one openly identifying that the assessments drove the curriculum, particularly in times of educational change, and acknowledged that this experience was 'miserable for the pupils':

Teacher F: *When there is a new change, there's always a tendency for teachers to teach for exams, and that is the problem, so although it might get them good grades at the end, it's incredibly miserable for the pupils themselves [...] I think that is the danger, and especially for the new cohort that went through last year, it was very much taught for the exams rather than teaching them, [...] it was very much exam-based driven, the lessons, so we worked through a lot of past papers, erm, and I, as I said, I think that is mainly due to teachers not perhaps being as confident with the new syllabus, they wanted to make sure they were covering everything in the exam.*

The same teacher also commented that:

Teacher F: *in terms of results, I can see that it is working. In the classroom it feels like it's not, but in terms of results it is. So, that is all you can go on.*

Perceived challenges presented by the Curriculum for Wales

Teachers participating in the study were invited to reflect on the changes that the new curriculum would bring about. The responses given by teachers throughout this section are in response to a number of interview questions which aimed specifically to elicit teachers' perspectives on how the reformed specifications at GCSE and A-Level aligned to the principles laid out in *Successful Futures*. These interview questions included: 'In relation to assessment, what strengths are there in the vision of *Successful Futures*?', 'In relation to assessment, what weaknesses do you foresee in the vision of *Successful Futures*?', 'With regards to assessment for certification, how do the reformed qualifications align with the principles laid out in *Successful Futures*?' and 'Looking forward to the implementation of the new curriculum, how would you like to see assessment develop to align to the principles of *Successful Futures*?'.

A number of general risks and challenges were identified, which included the potential for tokenistic and surface-level implementation, and a lack of time and capacity to undertake the type of involved, collaborative work curriculum development involves:

Teacher B: *I am a little bit worried as to how much individual schools have the capacity to change and reform and that [...] possibly where opportunities might be lost for the sake of tokenism.*

Teacher J: *I think one of the challenges that we are facing as a school is not so much erm, teachers' attitude, but whether or not they get enough time allocated to be able to collaborate, to come up with new schemes of work and a new way of thinking.*

Teacher B: *So for example, I'm the head of maths, let's go down Donaldson, I'm the head of maths and numeracy, have I actually fundamentally changed anything by doing that? No, but I've hit that tick box for Donaldson.*

As regards assessment specifically, the prospect of a new curriculum with a contrasting ethos, and attendant assessment practices from those currently in use, did cause some concern. Teachers seemed uncertain as to how increased autonomy in relation to curriculum delivery and teacher assessment could be faithfully aggregated into system-wide measures which were consistent and mutually cognate: one teacher noted that Donaldson's (2015) call for the curriculum to assess a wide range of competences, using a variety of assessment methods, could prove difficult to transpose to the sphere of assessment for certification:

Teacher K: *[the design of the Curriculum for Wales] isn't designed to be assessed how we assess.*

Teacher H: *we've been presented with some of the options, which are in my mind [...] un-assessable at qualification level in terms of GCSE and A Levels, if – or whatever replaces them in the future.*

Although teachers acknowledged that the structure of qualifications would require reform to align with the new curriculum, they found it difficult to conceptualise an approach based on areas of learning as conducive to deep-level learning of subject areas and preparation for formal assessments:

Teacher C: *arguably you could say that the old curriculum prepared them for what is to come at GCSE, I'm not quite sure that [the new curriculum] will.*

Teacher D: *You bring them into this school, you do topic-related teaching, how are you going to get the facts across in science that they then need to build on to get to GCSE, to get to A Level and to get to University?*

Teacher L: *I've noticed where they want to have themed areas [...] being] merged with another topic area could be quite alarming if it means that the students aren't getting into as much depth in a particular area.*

Teachers found it difficult to envisage what a GCSE assessment system which was not standards- or exam-driven would look like in practice, and expressed scepticism that creativity and teacher agency would be sustained during the latter stages of compulsory education when the focus returned to the outcome of assessments:

Teacher F: *[the idea behind the Curriculum for Wales is] that it wouldn't be teaching for exams, and that end goal, but I don't understand how then, it would be assessed.*

Teacher G: *I'm worried [about] what happens at GCSE when you've got an examination to sit [...] how much freedom would we have? [...] because] your time isn't there and because they have an examination to sit, and your standards are going to be compared by the results of those examinations.*

Beyond this, one teacher suggested that, while the principles behind the new curriculum were a welcome addition to twenty-first century schooling in Wales, the lack

of measurability of the four curriculum purposes and the overarching ethos of the curriculum, would ultimately lead to a system built on tokenism and ‘lip service’:

Teacher B: *what I would hope is that those four curriculum purposes actually do create ambitious, ethically aware, mindful individuals, but at that point the mathematician in me says, Governments can't quantify that, and so it's going to come down to some sort of citizenship style, PSE, GCSE nonsense so we can tick box and assess and monitor and league table and again, at that point, we are back to what I was saying earlier on about tokenism [...] that we are playing lip service to the intent of what it was, simply because we can't necessarily monitor it.*

Discussion and conclusion

The data presented in this paper from the experiences of pupils and teachers corroborate received accounts of education in Wales being in transition between two contrasting systems (Egan, 2017; Davies *et al.*, 2018; Connolly *et al.*, 2018): from one which prizes high levels of accountability, externality and which encourages performativity to one which will ostensibly offer greater levels of curriculum freedom, teacher autonomy and agency, and less high-stakes assessment (Donaldson, 2015; Welsh Government, 2017). Pupils in this study echoed the views of participants in Elwood *et al.* (2017)'s study, articulating a clear sense that their educational experiences during their final years of compulsory schooling were too keenly focussed on performance rehearsal for examinations, which they did not always feel reflected the full range of their learning. This paper offers evidence that the accountability frameworks under which schools operated, and the reforms to GCSE and A-level assessment implemented from 2015 have led to an ‘assessment-driven curriculum’ (Brown, 2004), which prompted teachers to engage in performance-driven classroom practices designed to improve visible and externally mandated measures of progress and school standards—often against the professional and pedagogical judgment of teachers themselves. Teachers spoke of their powerlessness to act in ways that they saw as meeting what they regarded as the ‘true’ learning needs of their learners, echoing Ball’s (2003, p. 216) contention that systems which encourage high levels of performativity can undermine teachers ‘as ethical subjects’ who find ‘their values challenged or displaced’. The performative practices, which teachers perceive as central to the effective implementation of the reformed specifications, are clearly misaligned with the principles of assessment articulated in *Successful Futures*.

Somewhat understandably, teachers in this study found it difficult to conceptualise a system of assessment, in what is now Key Stage 4, which is not only simultaneously aligned with the Curriculum for Wales, attuned to contextual progression steps and distinctive curricula, but also able to be meaningfully consistent and cognate across Wales, for the purposes of GCSE certification. This perception of an ‘un-assessable’ (Teacher H) curriculum will need to be considered carefully and addressed as the next stages of implementation are planned. A negotiated position will need to be articulated precisely, which clarifies the proposed link between assessment and certification, beyond the initial commitment to uncoupling assessment and external accountability. And, while some teachers clearly do seem beleaguered by a system where assessment and qualification lead pedagogy, some seem also uncertain and cautiously sceptical about the prospect of system where pedagogy and curriculum

are wholly unburdened by the requirements of qualification, and where learning outcomes may not always necessarily lend themselves to cognate assessment and comparison across contexts.

In response to Welsh Government, Qualifications Wales have recently set out the timetable for development of new GCSE qualifications, to be delivered from September 2025, with the qualifications themselves, in all likelihood, being finalised at least 12 months prior to implementation (Qualifications Wales, 2019). Yet, in spite of what appears to be a consensus that greater alignment is needed to ensure that any revised qualifications do not undermine the integrity of the curriculum, there are clear risks that are posed by the timing of the changes. Priestley and Minty (2013, pp. 49–50) note that in Scotland some secondary schools operated a ‘wait and see’ approach to developing the senior phases of the Curriculum for Excellence, until after detailed qualification specifications had been published. One of the potential risks of the current situation in Wales is that some teachers and schools may be unwilling to invest themselves conceptually and practically in the ethos of the new curriculum until they understand fully the requirements of the revised qualifications, and their associated assessment modes. This study also suggests that there is also the potential for weak, tentative or surface-level implementation; or by contrast, we note that curriculum implementation in schools could come in two phases: one prior to the development of the qualifications where the principles of the curriculum may be fully embraced; and another following the publication of the qualification specifications, where teachers and schools may be tempted to review their assumptions and practices hitherto, to realign them strategically with assessment goals in the new qualification specifications. In any case, as we move, ostensibly at least, to a curriculum and set of assessment and certification practices that purport to take a more rounded view of learner achievement, schools and teachers in Wales will need to be reassured that they can safely move away from the types of strategic, instrumental and performative pedagogical behaviours that are described above, and which, we contend, can be understood as their response to a system which has historically encouraged such practices. Doing so will involve truly uncoupling assessment from the pressures of external accountability, in order that assessment can have at its heart the needs of Welsh learners and can be in line with the principles of the Curriculum for Wales.

Acknowledgements

The authors would like to thank all of the pupils and teachers who participated in this research.

Disclosure statement

No financial interest or benefit has arisen as a direct application of the research presented in this paper.

Funding

This research arises from a PhD study that was part-funded by the School of Education at Aberystwyth University.

Geolocation information

The data presented in this study was collected from four settings across North-, Mid- and South-Wales, United Kingdom.

Data availability statement

There is not a publically available data set to support this study.

References

- Andrews, L. (2011) *Teaching makes a difference* (Speech at the Reardon Smith Lecture Theatre, Cardiff, 2 February). Available online at: <http://www.iwa.wales/click/2011/02/teaching-makes-a-difference/>
- Andrews, L. (2014) *Ministering to education: A reformer reports* (Cardigan, Parthian Books).
- Ball, S. (2003) The teacher's soul and the terrors of performativity, *Journal of Education Policy*, 18(2), 215–228.
- Barrance, R. (2019) The fairness of internal assessment in the GCSE: The value of students' accounts, *Assessment in Education: Principles, Policy and Practice*, 26(5), 563–583.
- Barrance, R. & Elwood, J. (2018) National assessment policy reform 14–16 and its consequences for young people: Student views and experiences of GCSE reform in Northern Ireland and Wales, *Assessment in Education: Principles, Policy and Practice*, 25(3), 252–271.
- Bennett, R. E. (2011) Formative assessment: A critical review, *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.
- Black, P. (2015) Formative assessment—An optimistic but incomplete vision, *Assessment in Education: Principles, Policy & Practice*, 22(1), 161–177.
- Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology, *Qualitative research in psychology*, 3(2), 77–101.
- Brown, G. T. L. (2004) Teachers' conceptions of assessment: Implications for policy and professional development, *Assessment in Education Principles Policy and Practice*, 11(3), 301–318.
- Cohen, L., Manion, L. & Morrison, K. (2018) *Research methods in education* (8th edn) (Oxon, Routledge).
- Connolly, M., Milton, E., Davies, A. J. & Barrance, R. (2018) Turning Heads: the impact of political reform on the professional role, identity and recruitment of head teachers in Wales, *British Educational Research Journal*, 44(4), 608–625.
- Creswell, J. W. & Plano Clark, V. L. (2011) *Designing and conducting mixed methods research* (2nd edn) (Thousand Oaks, CA, Sage).
- Crocker, L. (2005) Teaching for the test: How and why test preparation is appropriate, in: R. Phelps (Ed) *Defending standardized testing* (Hove, Psychology Press), 159–174.
- Dann, R. (2002) *Promoting assessment as learning: Improving the learning process* (London, Routledge/Falmer).
- Darbyshire, P., MacDougall, C. & Schiller, W. (2005) Multiple methods in qualitative research with children: More insight or just more? *Qualitative Research*, 5(4), 417–436.
- Davies, A. J., Milton, E., Connolly, M. & Barrance, R. (2018) Headteacher recruitment, retention and professional development in wales: Challenges and opportunities, *Wales Journal of Education*, 20(2), 204–224.
- Donaldson, G. (2015) *Successful futures: Independent review of curriculum and assessment arrangements in Wales* (Cardiff, Welsh Government).

- Egan, D. (2017) *After PISA: A way forward for education in Wales?* (Merthyr Tydfil, Bevan Foundation).
- Elwood, J., Hopfenbeck, T. & Baird, J.-A. (2017) Predictability in high-stakes examinations: Students' perspectives on a perennial assessment dilemma, *Research Papers in Education*, 32(1), 1–17.
- Estyn. (2018) *New qualifications* (Cardiff, Estyn).
- Firestone, W. A., Monfils, L. F. & Schorr, R. Y. (2004) *The ambiguity of teaching to the test standards, assessment, and educational reform* (Mahwah, NJ, L. Erlbaum Associates Publishers).
- Hayward, E., Morrison-Love, D., Makara Fuller, K., Spencer, E., Hughes, S., Ryder, N., et al. (2017). *CAMAU project: Progression frameworks and progression steps*. Available online at: https://www.gla.ac.uk/media/Media_649431_smxx.pdf
- Holloway, J. & Brass, J. (2018) Making accountable teachers: Terrors and pleasures of performativity, *Journal of Education Policy*, 33(3), 361–382.
- Hopfenbeck, T. & Stobart, G. (2015) Large-scale implementation of assessment for learning, *Assessment in Education: Principles, Policy & Practice*, 22, 1–2.
- Popham, W. J. (2000) The mismeasurement of educational quality, *School Administrator*, 57(11), 12–15.
- Moon, D. (2012) Rhetoric and policy learning: On Rhodri Morgan's 'Clear Red Water' and 'Made in Wales' health policies, *Public Policy and Administration*, 28(3), 306–323.
- Morgan, D. L. (1996) Focus groups, *Annual Review of Sociology*, 22, 129–152.
- National Assembly for Wales. (2018) *Explanatory memorandum: The education (Amendments relating to teacher assessment information) (Wales) regulations 2018*. Available online at: <http://www.assembly.wales/laid%20documents/sub-ld11633-em/sub-ld11633-em-e.pdf>
- Newby, P. (2013) *Research methods for education* (Oxon, Routledge).
- Ofqual. (2016) *An Investigation into the 'sawtooth effect' in GCSE and AS/A-Level assessments*. Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/549686/an-investigation-into-the-sawtooth-effect-in-gcse-as-and-a-level-assessments.pdf
- Power, S. (2016) The politics of education and the misrecognition of Wales, *Oxford Review of Education*, 42(3), 285–298.
- Qualifications Wales. (2019) *The new curriculum—Correspondence with Welsh Government*. Available online at: <https://www.qualificationswales.org/english/qualifications/qualification-reform/curriculum-reform/the-new-curriculum—correspondence-with-welsh-government/>
- Priestley, M. & Minty, S. (2013) Curriculum for excellence: 'A brilliant idea, but...', *Scottish Educational Review*, 45(1), 39–52.
- Reynolds, D. (2008) New labour, education and Wales: The devolution decade, *Oxford Review of Education*, 34(6), 753–765.
- Saldaña, J. (2009) *The coding manual for qualitative researchers* (Thousand Oaks (CA), Sage).
- Segal, A., Snell, J. & Lefstein, A. (2017) Dialogic teaching to the high-stakes standardised test? *Research Papers in Education*, 32(5), 596–610.
- Shapira, M. & Priestley, M. (2018) Narrowing the curriculum? Contemporary trends in provision and attainment in the Scottish curriculum, *Scottish Educational Review*, 50(1), 75–107.
- Sommer, M. & Arendasy, M. E. (2015) Further evidence for the deficit account of the test anxiety–test performance relationship from a high-stakes admission testing setting, *Intelligence*, 53, 72–80.
- Stobart, G. (2012) Validity in formative assessment, in: J. Gardner (Ed) *Assessment and learning* (London, Sage), 233–243.
- Vandewalle, D., Nerstad, C. G. L. & Dysvik, A. (2019) Goal orientation: A review of the miles traveled and the miles to go, *Annual Review of Organizational Psychology and Organizational Behavior*, 6, 115–144.
- von der Embse, N., Jester, D., Roy, D. & Post, J. (2018) Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review, *Journal of Affective Disorders*, 227, 483–493.
- Welsh Government. (2017) *Education in Wales: Our national mission. Action plan 2017–21* (Cardiff, Welsh Government).