Tests as boundary signifiers: level 6 tests and the primary secondary divide

COLDWELL, Michael <http://orcid.org/0000-0002-7385-3077> and WILLIS, Benjamin <http://orcid.org/0000-0002-4346-0459>

Tests as Boundary Signifiers: Level 6 tests and the primary secondary divide - The Curriculum Journal

Mike Coldwell: Head of Centre for Development and Research in Education in the Sheffield Institute of Education, Sheffield Hallam University, Arundel 10101, Howard Street, Sheffield S1 1WB, England; e-mail: M.R.Coldwell@shu.ac.uk. His interests center on educational policy evaluation, and the school workforce.

Ben Willis: Research Fellow at the Centre for Development and Research in Education in the Sheffield Institute of Education, Sheffield Hallam University, Arundel 10101, Howard Street, Sheffield S1 1WB, England; e-mail: b.willis@shu.ac.uk. His interests center on children's wellbeing and the wider dimensions of children's education.

Abstract

This paper addresses the question: how do teachers and school leaders respond to high stakes testing of pupils transitioning from primary to secondary school? It explores how a new test, the Level 6 test, operated with regard to primary/secondary school relationships in England. It draws on an analysis of qualitative interviews with teachers and school leaders in 20 primary schools that took part in the test, 40 school leaders that chose not to and 20 secondary school leaders. Theoretical work on social boundaries is utilised to develop an argument that this test and its results acted as a boundary signifier, crystallizing many of the tensions between primary/secondary schools. These tensions included the role of accountability regimes in requiring schools to demonstrate progress; narrowing of the curriculum and teaching to the test; and the extent to which test results can provide a true representation of pupil attainment. We conclude by suggesting the potential of the boundary signifier concept in relation to other tests at the primary/secondary boundary and other key transition points in education, and consider whether such tests can act as an ideal boundary object (Star 1989), serving to help overcome, rather than cement, barriers between schools.

Key words: boundary objects, testing, attainment, primary/secondary education, accountability, transition, curriculum.

Introduction

This paper utilises an analysis of interviews with teachers and leaders in primary and secondary schools to consider the role that high stakes tests play in the relationship between primary and secondary schools. We present evidence that the test examined in this paper - the Level 6 test in England - brought out tensions between primary and secondary schools, including the role of accountability regimes in requiring schools to demonstrate progress; narrowing of the curriculum and teaching to the test; and the extent to which test results can provide a true representation of pupil attainment. Drawing on the research literature on boundary theory, we argue that the differing views we uncovered from primary and secondary teachers in England towards this new test and its outcomes suggest that such high stakes tests can be seen as a particular kind of boundary object. We refer to this as a *boundary signifier*: a boundary object that draws out tensions and differences across boundaries, in this case tensions between primary and secondary schools that are made particularly acute within the accountability regime in operation in England.

To establish the policy and theoretical context for this analysis, we first discuss the place of Level 6 tests within England's strong accountability regime, and then provide an argument for using boundary theory to frame an analysis of primary and secondary teacher responses to these tests.

Level 6 tests and England's strong accountability regime

In England, the end of Key Stage 2 (KS2)[1], when pupils are about to move from primary to secondary school, is a crucial transition point in their schooling. Until 2014, this meant that in the Spring term of Year 6, when pupils are aged 10 or 11, those who were judged by teachers as being able to access the National Curriculum at level 3 or above took part in National Curriculum Tests (NCTs) - formerly known, and still widely referred to, as Standard Attainment Tests (SATs) - in English and Mathematics. In 2010 the English Education Ministry, the DfE, launched a call for evidence in relation to testing and accountability at KS2 framed by a rationale that external accountability is a driver for improvement in education as well as school autonomy.

In 2011, Lord Bew reported on a review of KS2 testing more broadly, and one significant finding was that the then current system was seen by the government to place a ceiling on the attainment of what were described as the most able pupils, potentially impeding their progress (Bew 2011). This was because the standard NCTs gave judgments of levels between 3 and 5. The Bew review highlighted that in 2010, for example 51 per cent of pupils taking the tests achieved level 5 in Reading and 30 per cent in Mathematics suggesting that the tests did not discriminate sufficiently for pupils with the potential to demonstrate higher attainment. Following a pilot in 2011, the DfE made Level 6 tests available to schools on an optional basis in 2012 to enable them to stretch and demonstrate the abilities of pupils with the potential to attain higher grades than the standard tests allowed.

This focus on allowing schools to demonstrate pupil attainment - which can be seen as an extension of an already existing policy focussing on a narrow range of subjects in the primary age phase - is strongly related to the wider English state education policy context. Since the 1988 Education Reform Act (ERA), the English school system can be loosely described as what Supovitz (2009) calls a 'test-based accountability system', a system that is 'based on the belief that attaching incentives (either positive or negative) to standardised achievement tests will improve student performance' (Supovitz 2009: 213) as part of a wider 'regulatory system' (Jones 2003) including a sanctions-based inspection system. Alexander (2011: 266) makes the point that in the English context *accountability* itself is as inherently important to politicians as student performance: 'tests are the instrument of choice in policy-makers' efforts to do the two things which they believe they must always be seen to do: raise educational standards and call teachers and schools to account'. In England, the accountability mechanism - via league tables and comparisons of school performance - has meant that the organisation has become the focus of reform and it is predominantly at the level of the organisation - the school - to which Supovitz's 'incentives' apply.

Whether the introduction of school-level league tables, and a school inspection regime increasingly concerned with comparative organisational performance, has actually improved outcomes for pupils has been debated in the media, policy and academic spheres more or less since the ERA was enacted. In curriculum terms, the introduction of the national curriculum at this time was intended to be 'balanced and broadly based' (Department for Education and Science, 1988), and the associated assessment tasks originally mooted were intended to focus on all aspects of the national curriculum. However, there was evidence from the outset that the assessment of the curriculum at the end of Key Stage 2 as it was implemented, looking only at English, Mathematics and - at that time - Science, narrowed the focus. As Boyle and Bragg (2006: 570) put it, 'the inevitable happened; a subject-centred curriculum saw the elevation of a number of subjects into a "core"'. This contributed to a process by which, as Wyse and Torrance (2009: 216) argue, 'a potentially positive educational

[1] There are four 'Key Stages' in the English Schooling system KS1 (Years 1 and 2, for pupils aged 5-7) and KS2 (Years 3 to 6, ages 7-11) in primary schools, and KS3 (Years 7 to 9, ages 11-14) and KS4 (Years 10 and 11, ages 14-16) in secondary schools.

deployment of "standard assessment tasks" was transformed by political pressure into a national testing system.' As the proportion of children achieving higher levels in tests in both primary and secondary schools increased year on year, the debate soon turned to whether in fact 'standards' - measured by 'the overall quality of education in our schools' as opposed to simply test scores (Mansell 2007: 25) - were improving. Whilst the debate has most recently decisively turned to international comparisons using the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMMS) scores in particular, from the late 90s onwards concerns were raised that, if one used other measures including alternative 'lower stakes' tests, standards were not rising at this rate (e.g. Tymms and Merrell 2007).
The reasons behind the disparities between higher stakes test results and other measures of standards relate to what Torrance (2011:464) identifies as a 'schism between the educational arguments for changes in assessment to enhance learning, and the policy demands for school improvement and accountability'. In practice, the incentives for schools related to maximising pupils performance in tests *funnel* teaching efforts to improve test performance. The research literature indicates how this works in at least two ways.

Firstly, targets set in tests encourage schools to concentrate their efforts on particular groups of pupils. In the English system, schools' positions in GCSE league tables were for many years based on the proportion of pupils gaining five or more subjects graded at C or above. Therefore schools, quite rationally, focussed their efforts on pupils who were 'on the borderline' between grade C and D providing to those selected for additional support via a process of what Gillborn and Youdell (2000) refer to as 'educational triage' to be provided with what Ball et al (2012: 520) describe as a 'diverse range of often very imaginative techniques' - from mentoring and school trips to revision classes to having photographs of targeted students in the staffroom - to ensure these pupils reached grade C. Similarly, in primary schools, Troman's (2008) detailed qualitative study of six schools in different English regions identified approaches including 'teacher-led Saturday morning and holiday revision sessions for Year 6 pupils in the run-up to the SATs tests' (p. 623), and - in the case of one school - use of 'literacy and numeracy "experts" from the local education authority to support Year 6 teachers in planning for SATs tests and "coaching" of children targeted as needing to improve their performance.' (p. 624).

Secondly, the particular testing regime in place will affect curriculum content, pedagogic approaches and increase the time spent teaching to the test (Stecher and Barron 1999; Berliner 2011; Ball et al 2012). Au (2007)'s metasynthesis of research into high stakes testing indicates that accountability pressures that focus on tests can shift curriculum time onto tested subjects, and in fact alter the curriculum to become more fragmented into 'test-sized pieces' (Au 2007: 262), as well as altering pedagogy to become more teacher-centred with a focus on 'direct transmission of test-related facts' (ibid, 263). Boyle and Bragg (2006: 597) demonstrate that, in English primary schools, curriculum time spent on Mathematics and English increased over time from the late 90s indicating 'a primary curriculum dominated by teaching time allocated to English and Mathematics' which they associate with a range of policy instruments including testing focussed on these two subjects. In further analysis of the same dataset, these researchers identify changing patterns in curriculum organisation towards subject-based teaching in primary schools which appears in Year 6 to be associated with test preparation (Boyle and Bragg, 2008). Beyond the focus on particular subjects, Troman, Jeffrey and Raggl (2007) provide evidence of how this can lead to pressures to reduce and restrict creativity across the curriculum, especially in the final year of Key Stage 2 due to the pressures associated with the SATs tests. Torrance (2011: 9) demonstrates that this narrowing of the curriculum has been recognised for many years, including by the English school inspectorate, Ofsted, and politicians and other public bodies including, particularly pertinently in relation to this study, the writers of the Bew Review which identified concerns that the system is too "high stakes", which can lead to unintended consequences such as over-rehearsal and "teaching to the test".' (Bew 2011: 9). .

However despite this raft of evidence that high stakes testing, including in relation to English primary education, has the detrimental effects outlined in the paragraph above, there is little published research on the potential for tension to emerge between primary and secondary schools in relation to their responses to such tests and their results, that is inherent in a system where both types of schools need to demonstrate progress. We are not concerned primarily here with pupils' transition across the divide, written about widely in the education field (e.g. Galton, Gray and Rudduck's (1999) work); rather, we are interested in teacher, curriculum and organisational relations. For this, we turn to another body of literature, focussing on boundaries.

Boundary Theory and the primary/secondary divide

Engestrom et al's (1995) concept of a 'boundary' is described by Akkerman and Bakker (2011: 133) as a 'sociocultural difference leading to discontinuity of action or interaction', focussing largely on the boundary as being between two social spaces or sites. In contrast, Tilly (1999) writes of the boundary as a dividing point that can be used as a method of maintaining power and resource, with 'boundary maintenance' activities important in shoring up the higher status of those 'within' the boundary. Tilly therefore conceptualises a boundary being 'around' something to be protected, not simply as a border or divide between two spaces. Aikenhead's (2001) writing on borders uses the metaphor from political geography seeing borders as sites of crossing from one space to another (and in one sense this is also how Engestrom et al and other theorists constitute 'boundary crossing').

All three elements - boundaries between; boundaries around; boundary crossing - are important in the context of this paper. The differences across the divide between primary and secondary are many and various - differences in location and space; differences in scale and size; differences in pedagogical practices; differences in the workforce (again size, but also qualifications, gender, and identity as - for example - subject teachers on the one hand and teachers of children on the other). There are differences in status and pay, with secondary schools seen as having higher status than primaries. Coldron et al (2014), drawing on Bourdieu's (1977, 1990) concepts of capitals and Habitus and Schatzki's (2002) conceptualisation of practice, present an analysis of how the differing sizes, contexts and institutional practices of primary and secondary schools emphasise the differences between the two and lead to systematic and durable inequalities between them. These differences become sharper during the crucial time of crossing from primary to secondary schooling for pupils, when primary and secondary schools have to work together, and can become extremely pointed in relation to assessment of pupil attainment and progress. Following Tilly (1999, 2004), Coldron et al (2014) go on to argue that taken together this set of differences have amounted, over time, into primary and secondary schools becoming different 'social categories', as a consequence of which 'community building and boundary maintenance occur[s]' (p7).

Coldron and colleagues' analysis provides evidence that the primary/secondary divide can be conceptualised as a social boundary: a structural instantiation of socio-cultural difference. Such boundaries are - in Tilly's view - in need of maintenance and vigilance, and it is here that the role of tests comes into view. There are a number measures of attainment that are important in the English context, in particular KS2 tests which provided, until recently, a measure of the National Curriculum Level achieved in certain subjects. And, of central concern to this paper, these included the newer measure of performance on Level 6 tests. Such tests and - crucially - the results attained by pupils have particular status for both primary and secondary schools. For primary schools, they are a means to demonstrate progress made by pupils, extremely important evidence of the quality of the school

within the English accountability regime as discussed earlier. It is vital to primary schools that they can show as great a degree of progress for all pupils as possible, so for many schools the Level 6 tests were an opportunity to remove an artificial ceiling in relation to National Curriculum levels, since the standard KS2 tests set level 5 as an upper limit.

The same pressures apply for secondary schools: from their perspective there is a strong incentive to depress or downplay the progress made by primary age pupils; or - at the very least - a concern to ensure that primary schools are not overstating the current attainment level of pupils on entry to secondary school, since secondary schools themselves need to demonstrate progress from the start of primary school to the end of secondary schooling. Work by the Datalab (2015) in England has produced an analysis to show that there is evidence of suppression of KS 1 teacher assessment data in junior schools, which enables progress from KS1 to KS2 to be shown, so it appears this is a real concern within the English system. Therefore, these sets of tests have a particular status in relation to the primary secondary divide.

Boundary theory provides a specific concept that is of use in considering the role of tests in relation to this divide, developed by Star and colleagues in the late 80s: the *boundary object*. Star (1989) considers boundary objects as *artefacts* that smooth boundary crossing by providing a bridging mechanism: boundary objects have 'different meanings in different social worlds, but their structure is common enough to more than one world to make them recognisable, a means of translation' (Star and Griesemer 1989: 393). They allow for 'interpretive flexibility' (Star 2010: 602) and so can facilitate action and learning. It is important to recognise that this perspective treats boundaries not as markers of difference to be maintained as with most other forms of boundary in the social science literature (Lamont and Molnar 2002) rather as a 'shared space, where exactly that sense of here and there are confounded' (Star 2010: 603). Thus the boundary is an 'interface, facilitating knowledge production' (Lamont and Molnar 2002: 180). Drawing on this conception, Akkerman and Bakker (2011) discuss 'ideal' boundary objects as artefacts that 'allow different groups to work together, based on a back-and-forth movement' (p141). If such artefacts do not work in this way, they can 'fail as boundary objects' (ibid).

Nolen et al (2011: 94) note that course grades can be seen as boundary objects as part of the group of 'standardised forms' - 'objects devised as methods of common communication across dispersed work groups' and therefore argue that as 'a primary means of reconciling the differing views of various groups in the educational system, assessment tools, practices and artefacts and their function are worthy of further study'. Cobb et al (2009: 179) discuss assessment tools as boundary objects, noting that 'members of the school leadership groups looked at teachers through the state standards and test scores and made judgments about individual teachers' competence as they did so.' This perspective opens up the space for the current paper, allowing us to explore how and to what extent Level 6 tests operate as boundary objects across the primary-secondary border.

The Aims of the Paper

The two sections above provide the context for this paper's main aim which is to address the question: how do teachers and school leaders in primary and secondary schools respond to high stakes testing of pupils transitioning from primary to secondary school? It does so by examining the responses of primary and secondary school teachers and leaders to Level 6 tests, in the context of the strong English accountability regime; drawing on boundary theory to tease out some of the latent tensions underlying the two phases. This further enables us to explore the potential for using such theory for future work in relation to cross-phase boundaries in education.

The study and methods

The Bew review of primary school testing (Bew, 2011) raised concerns regarding the additional tests, particularly the extent to which primary schools would be able to offer enough of the KS 3 curriculum to allow pupils to attain level 6. Such concerns led to the DfE commissioning a research study into the KS 2 Level 6 test. This study (Coldwell, Willis and McCaig, 2013) aimed to understand how schools supported pupils for entry to the new Level 6 tests, and investigate why they chose to enter the tests or not. The study used three main methods: school case study visits, telephone interviews with non-participating senior leaders and secondary school senior leaders.

The overall research approach employed was approved by the ethics committee of Sheffield Hallam University. All adult participants received an information sheet on the project and provided written consent to be included in the study. In relation to pupil data, schools were asked to send each pupil's primary care giver an information sheet and consent form requesting permission for their child to be involved. Prior to pupil interviews, researchers verbally discussed the key features of the information sheet in an age appropriate manner with the pupils emphasising the voluntary nature of participation and providing opportunity for questions to be asked; before taking verbal consent.

The largest body of data was a set of 20 case studies of schools taking part in the tests. The sample of case study schools was drawn from data supplied by the Standards and Testing Agency (STA) - an executive agency of the DfE responsible for setting the tests to assess children in education from early years to the end of KS3 - based on a list of all level 6 registered schools (i.e. those schools entering pupils for the L6 test). To understand the range of issues and circumstances which affect a school's decision to enter pupils for the Level 6 test and, importantly, the relationship between their decision and the pupil outcomes the sample was designed to reflect the diversity of English schools and their characteristics. To do so, a 'sampling matrix' was used that selected schools broadly representative in relation to the percentage of the Y6 cohort entered for the Level 6 tests as well as a number of overall school population characteristics. Sample schools were allocated to one of 12 'cells' in the matrix across two dimensions, cohort size and proportion of cohort achieving level 5 in the previous year. Other population characteristics including Free School Meals (FSM) entitlement, English as an additional language (EAL), Ofsted rating and geographical spread were monitored to create a balanced spread of schools. The matrix allowed replacement schools to be randomly selected within the same cell should certain schools refuse to participate, enabling the construction of a broadly balanced sample as evidenced in Link 1.

For both case studies and non-participating schools a thematic analysis was used as described below, broadly utilising a Framework Analysis (Smith and Davies, 2010) approach - involving gaining an initial overview of the data, building an initial framework drawing on research questions, then detailed coding or charting data according to themes from the framework and finally interpreting the data within the framework.

Each case study visit included semi-structured interviews with school senior leaders and Y6 teachers. In addition participating schools were asked to identify up to eight pupils covering a range of characteristics broadly reflective of a typical Y5 class (in relation to gender, ethnicity, FSM eligibility, EAL and attainment levels) to take part in a focus group. Interview schedules included a set of broad questions that were designed to contribute towards answering the main research questions of the project: key foci included the rationale for opting the school in to taking the Level 6 test, selection processes, school practices in terms of curriculum delivery and specific Level 6 test preparation,

extent of support for teachers and the impact of the test on the pupils (Link 2 - school case study interview and focus group schedules).

The dataset of interview transcripts for each school were cross-referenced with actual test outcomes and details regarding selection practices for pupils in the schools (for further details, see Coldwell, Willis and McCaig, 2013: 101-106). This permitted tentative initial judgements about how consistently schools were selecting pupils to enter into the Level 6 test and to assess how relatively successful their pupils were having undertaken the test. All teacher and pupil focus group interviews were digitally recorded, partially transcribed, anonymised and written up as detailed school-level summary reports. An initial thematic analysis of each of the school level summary reports allowed comparison within and between case studies, highlighting emerging issues and any differences captured in the data. Following this initial stage of analysis, *visible case maps* were created for each school: A3 sized flow diagrams based primarily on a conceptualisation of the original research questions but further informed by the themes identified following the initial thematic analysis. The diagram was based on three overarching columns Influencing/contextual factors; Actions and Outcomes each were composed of a series of sub-categories, each with open boxes [See Link 3- a unpopulated version of a visible case map]. This template was then populated with key data extracted from the more detailed summary reports. Methodologically having a top level data source was particularly helpful for ensuring that the analysis considered each case study as well as offering a quick means of comparing level 6 outcomes against school context. The 20 A3 analysis templates formed the basis for a team analysis meeting (involving an internal advisory group) which identified the key themes emerging from the data which in turn helped inform the report structure.

The second dataset was a group of 40 telephone interviews with primary school leaders that decided not to take part in the tests. A sampling method was designed, based on data from the Annual School Census (ASC) and National Pupil Database (NPD) supplied by DfE, with L6 schools removed. 200 maintained primary schools were selected and through a randomised sampling process similar to that previously outlined for selecting case study schools a sample of 40 non-L6 schools were achieved. Again, the sample was stratified by population characteristics and geographical spread. The focus of these interviews was to help ascertain levels of awareness of the L6 test, the rationale for not taking part, perceived impacts of not becoming engaged at school and pupil level and to explore what strategies were employed to 'stretch' high attaining pupils [See Link 4 - non-participating school telephone interview schedules]. The non-participating Level 6 test primary school telephone interview data was partially transcribed and entered into an excel spreadsheet as a case by theme matrix, allowing both qualitative and quantitative analysis.

As this set of data was analysed, it became apparent that one reason for not engaging in the tests was that some primary schools felt the results would not be taken seriously by secondary schools. Therefore, an additional set of interviews were undertaken with 20 senior leaders of secondary schools, utilising the same sampling frame as in the main study to obtain a range of settings, with the inclusion of some secondary schools from selective areas. These interviews asked about secondary leader awareness of the L6 tests, their views on their appropriateness, and whether and how they would make use of any data from them (See Link 5 - secondary school interview schedule).

As with all qualitative research, the findings are not statistically generalizable to the wider population of schools. Nevertheless, the systematic approach to sampling, data collection and analysis outlined above provided a strong evidence base for the findings section that follows, since the data gathered provided a balanced spread of approaches, contexts and perspectives that aligned with the wider relevant school populations.

Findings: test result security and primary-secondary boundary issues

As we noted earlier in the paper, Star and other writers that use the concept of boundary object identify such objects as artefacts that provide bridges to understanding across social divides (Star and Griesemer 1989) ideally allowing groups to work more effectively together across such divides and build knowledge and learning (Akkerman and Bakker, 2011). In this section, we present an examination of how teachers and senior leaders in primary and secondary schools interpreted the Level 6 tests and their results, and whether these interpretations operated to provide such a bridge. In fact, the analysis exposed that the tests rarely enabled closer working but instead revealed differences across the primary/secondary border, many of which are related to accountability pressures. Such differences relate to concerns from secondary schools about the *security* of the test results in representing the attainment, capabilities and potential of pupils. We identify a set of these concerns below related to:

- Narrowing the curriculum and teaching to the test
- The role of the test in positioning the school with stakeholders
- Concerns around interpretation of 'level 6'
- Test preparation practices
- Selection practices

Narrowing the curriculum and teaching to the test

A recurrent issue raised by secondary school headteachers, in line with Torrance's (2011) work and others, was that the Level 6 tests drive practice in primary schools to achieve a result at level 6 that resulted from teaching to the test rather than deriving from pupils having a deep level of learning revealed by the test. This is exemplified by the comment below:

> *It's appropriate to teach KS3 at level 6, that's not the problem, but it shouldn't be all about teaching to the test, that is the risk. Some pupils we have had, that had level 6 think they are going backwards when they come to us*. (Assistant head teacher, secondary school 8)

Further, this focus on teaching to test taking was related to primary school accountability measures - what another secondary leader refers to here as targets:

> *If it is rigorous and externally assessed then those students who have a natural aptitude to achieve level 6 should be given the opportunity to do that…. but if a teacher has a target* [to achieve level 6 then] *we would be in an even worse position for inheriting even more inflated SATs* [Standardised Assessment Tests] *scores than we already do*. (Vice principal, secondary school 4)

Interviews with head teachers in secondary schools indicated that the tests were experienced as having causal power: they were seen - in the context of the wider accountability system - to strongly drive the actions of teachers and senior leaders in primary and secondary schools. This perception reflected a concern expressed by primary teachers themselves who experienced this drive as a set of tensions around their approach to the test, a finding in line with previous research, for example Troman and colleagues' work on what they refer to as "cultures of performativity" in primary schools in England (Troman, Jeffrey and Raggl, 2007; Troman, 2008) . Some schools saw it as something to be shown to parents, teachers and others, and this involved pride - 'pride that the highest achieving are recognised as such' (Head teacher, primary school 18). For this group, the test results provided a true representation of individual achievement as expressed by this school leader:

> *I recognised the children's ability and I wanted to not have them limited by the fact that I could only level them to level 5...There's never been any pressure on me to get pupils to level 6 because it would be kudos for the school…we want the children to do the best that they can and to show what they can do.* (Senior teacher, primary school 11)

Such schools often emphasised there was a moral imperative for engaging with the Level 6 tests, with those pupils adjudged to be capable of exceeding an artificial ceiling of level 5 permitted the opportunity to do so. However, others were concerned that this might involve pushing pupils beyond what was appropriate for their age. Some schools discussed the dilemma around providing a broad, balanced curriculum versus a narrow experience for the pupil:

> *Shall we look at KS3 curriculum* [i.e. a narrow focus] *or shall we try and give these children a broad, as rich as we can* [curriculum] *in that level 5 and we decided to go down that route….going sideways rather than up…we didn't want to push, push and just move them up.* (Deputy head teacher, primary school 10)

The theme of narrowness in the curriculum - in line with Boyle and Bragg's (2006) work in England and Au's (2007) review of the effect of high stakes testing on school curricula - was picked up by a number of the secondary school leaders interviewed. The assistant head teacher of secondary school 1 was particularly uneasy about the potential for the L6 Tests in Mathematics and Reading to further tip the balance of emphasis at primary level towards the core subjects of Mathematics and English; thereby reducing coverage of other subjects such as Science, History, Geography and Languages, meaning that *'some…students come in with very poor knowledge…even the high end of children haven't done a lot of Science'*, emphasising the need to *'broaden the whole curriculum so they've [pupils] got a better understanding of a variety of subjects'*.

The implications of such thinking were frequently expressed as concerns about 'teaching to the test', which was sometimes recognised by primary schools:

> *I do feel sorry for secondary schools on that basis* [the growing prominence of performance tables].*Children aren't given a broad base. In certain schools where children are deemed more able, the pressure is placed on them to produce these high results, all they do is teach to the tests. That is replicated in secondary schools where everything is based around GCSEs to the detriment of everything else.* (Head teacher, non-participating primary School 4)

These concerns were such that some schools - including the school quoted immediately above - opted out of engaging in the tests because they argued that taking part in the test would necessitate a narrowing of the curriculum.

The role of the test in positioning the school with stakeholders

For some primary schools, the focus of accountability was on the school demonstrating high standards or progress to others in the community, including teachers and parents (in line with Troman's 2008 study) - *'To inform stakeholders e.g. parents within our ranks we have pupils working at the absolute highest level and that's important for everybody to hear. It's important for staff to hear too*'. (Deputy head teacher, primary school 12). This slipped easily into the tests contributing to a positioning of the school to these stakeholders, internal and external, such as other schools and governors.

For example, in primary school 8, located within a selective authority and with a significant proportion of pupils transferring onwards to independent schools, the deputy head teacher noted that

school governors were eager for the school to *'extend the able, gifted and talented*' having previously conveyed frustration at the absence of a 'higher paper'. Although they felt that the school was already appropriately *'stretching*' pupils the deputy head teacher acknowledged the Level 6 test presented a concrete means of demonstrating this to governors: '*So this way at least I can say we've entered this many kids into the Level 6 test as more evidence that we are teaching at this level*'. (Deputy head teacher, primary school 8).

For this and other schools, involvement in the Level 6 test was strongly related to external pressures to demonstrate progress especially from Ofsted, the English inspection body, which provides a school level judgment:

> …*being told we are satisfactory* [i.e. not good or outstanding according to Ofsted inspection categories] *is not good for morale, so showing we have children not just above, but way above national expectations is a boost for all of us*. (Deputy head teacher, primary school 20)

In some cases, this amounted to admissions of 'games playing' linked to these specific accountability measures, with schools aiming to 'balance out' lower scores on some measures with others: *'…an ulterior motive really. Because it gives us more ability to show progress because of the cut off at the top'.* (Deputy head teacher, primary school 1). One head teacher revealed that their school was '*much of a muchness with many local schools*' but through the Level 6 test they could evidence a number of 'high achievers' that boosted their average points score, a measure of school effectiveness, which meant being able to 'further distinguish' their school and become more marketable. Teachers and senior leaders were, in the main, not happy with this perceived need to fit school policy to external judgments, a finding in line with wider research in this area such as Mausethagen's (2013a) work in Norway and a raft of further studies identified by Mausethagen's (2013b) review including Webb et al (2004) in England, Wong (2008) in China and Ng (2006) in the United States. Yet most school leaders seemed resigned to it:

> *Another part of my thinking is, if it's there and the kids are capable, you have to be pragmatic as well. I'm not one to cut my nose off to spite my face. That's not helpful for me; it's not helpful for anybody. You have to go with the system, whether you like it or not*. (Head teacher, primary school 5)

Concerns around interpretation of 'level 6'

Lack of adequate guidance was raised by a number of primary schools. Primary teachers felt confident in teaching up to level 5, but had seldom previously taught to level 6, and the curriculum support, materials and training they received focussed on levels 2 to 5. Almost all of the primary interviewees involved in teaching to level 6 felt that this meant they were unable to effectively support children working at this higher level. The deputy head teacher in one school laid out the issues starkly:

> *We have a test at level 6 but we have nothing in the curriculum...there is no framework at the moment to tell us what that looks like. If there weren't any pilot materials I would have had no idea what our children would have had to learn to pass level 6... It took me three to four years to learn how to teach and a good deal of that was learning familiarity with the primary curriculum it's not something even an experienced teacher could just pick up a textbook and go oh that's alright...it's not like that, it's way beyond that! (*Deputy head teacher, primary school 12)

Specifically, there were concerns about understanding what level 6 meant at primary age compared with secondary age. So even where there was training, primary teachers didn't feel supported in understanding what this meant:

> *They told us the type of questions the children could expect and the kind of things children do at level 6 when they go to secondary school but I haven't had anything written about the type of questions to expect... Because really I've had one morning of training*. (Deputy head teacher, primary school 3)

In a separate example, the lack of accessible training or guidance beyond example papers led a senior teacher to seek advice from a local secondary to help '*find out exactly what the year 7 curriculum was and go on for from there, because it was such a new area for me…I was a bit fumbling in the dark*'. (Senior teacher, primary school 11)

The consequence of the lack of clarity about what exactly should be being taught and ultimately tested through the Level 6 test meant that schools interpreted its parameters in different ways. For example, primary school 13 (a small rural school, in a selective area with very varied levels of attainment) took the view that the Level 6 test assessed content beyond what would ordinarily be covered within the KS2 curriculum and therefore sought out new materials to help deal with this:

> *You are constantly looking for resources...for extension tasks for my highest ability children so that awareness is there all the time that you've got to push these children and challenge them and to teach the kind of things that they're going to come across....* (Deputy head teacher, primary school 13)

In contrast the head teacher at primary school 18 made the assumption that the Level 6 test was sufficiently nuanced to be able to assess *'KS3 Mathematical understanding within the context of a KS2 curriculum*'.

These two examples illustrate that schools varied in their understanding of how this new test fitted within the current curriculum framework: on the one hand primary school 18 assumed the KS2 curriculum would be enough; on the other primary school 13 actively sought out KS3 content and incorporated that into their teaching in order to permit access to the Level 6 test. This variation related to school approaches to test preparation and the selection of pupils as we explore in the next two subsections.

Test preparation practices

There was wide variation regarding how much dedicated time schools spent preparing pupils for the Level 6 tests.

At one extreme, a minority of schools claimed to have done no specific test preparation at all. For example, interviewees at primary school 2 emphasised that routine differentiation within class meant they did not need to alter their approach from previous years:

> *I did nothing different to the usual challenging, extending activities that I would normally do as part of differentiating. I did no additional coaching for the test. I said I think you're able - have a go.* (Year 6 teacher, primary school 2)

Most schools described as a minimum working through sample and past level 6 papers with the pupils due to be entered to take the test in order to orientate them to the type of questions they could expect to receive. However, some schools suggested they could not dedicate additional time to preparation for the Level 6 tests given the accountability focus on borderline pupils' results in the standard NCTs:

> *No* [booster classes specifically for level 6] *because our resources were to boost those 3's and 4's to push them up because that is our focus. You know we want to make it accessible to all children and that is the majority so they are the ones that need it. Level 6 is a sort of bonus in our eyes...we're stretched as it is* [in terms of extra support available]! (Year 6 teacher, primary school 8)

Nevertheless other schools provided much more extensive preparation for pupils selected to sit the Level 6 test. For example at primary school 3, as part of a Gifted and Talented group run across schools in a local learning community, additional booster sessions were offered to all children selected to undertake the Level 6 test following local authority (LA) training about how to teach Mathematics at level 6.

> *I was volunteered to deliver a series of weekly booster lessons for the children that were going to do the L6 Maths test...We had about 20 children* [not every school sent pupils] *who came here every week* [for 10 weeks] *for about an hour and it was just purely on L6 Maths and I went on some training with [*named education authority*] which was really really good, it was basically how to teach Maths at level 6 and I used what I'd learnt on that course and the materials to deliver this booster* (Deputy head teacher, primary school 3)

Elsewhere (*Author(s), Date, page range*) present evidence that indicates that pupils in schools at the extreme ends of the preparation continuum - those, like primary school 2, that undertook no test preparation and others, such as primary school 3, that provided extensive preparation - performed more poorly than those in schools that took a more balanced approach.

Selection practices

Selection of pupils to enter for the Level 6 test also varied markedly across the sample. Selection practices were influenced by a combination of objective measures of prior attainment, such as quantifiable tracking data, with teacher assessments of attainment coupled with other judgements about pupils' ability to cope with an additional test and the possibility of being deemed below level 6.

So some schools, such as primary school 16, adopted a very cautious approach, electing to only put forward pupils achieving a mark of at least 95% in the pilot Level 6 test paper. This position was a result of concerns that should pupils perform below expectation in the Level 6 test, that it could have an adverse effect on their wellbeing and confidence ahead of their transition to secondary school: '*The disappointment on the children would spoil them for secondary school. So therefore I wouldn't do it*'. (Year 5/6 teacher, primary school 16)

Contrastingly, others such as primary school 18 employed a more liberal approach to their selection and had a differing perspective on the implications of the test for pupil transition; arguing that it gave borderline pupils the opportunity to ensure secondary schools had the highest expectations of them.

> *There were some we had a punt on. On a good day they might just nudge over the line, you know - shall we deny them the opportunity was the question…We discussed this issue and we said what's the positive, what's the negative. What's the negative? They'll do a test they can't*

> *do. They are pretty robust children, they can handle that. What's the advantage? It really might mean that they're expected to achieve the highest possible grades.* (Head teacher, primary school 18)

This variation on selection approaches led to widely varying results at the school level. All of the pupils that undertook the Level 6 tests also sat the standard NCTs during the same testing period, and comparative analysis of results from the two tests showed that the percentage of those pupils achieving level 5 in the standard NCT in Mathematics, for example, and who were also entered into the Level 6 test varied from just 8% (2 out of 26 pupils) to 100% per cent (5 of 5 pupils plus one pupil that attained level 4), with a similar degree of variation in relation to the Reading test.

The results of the two schools we discuss above, primary school 16 and primary school 18, illustrate the point. Primary school 16, taking a cautious approach, entered three pupils for the Level 6 Mathematics test and all achieved a pass mark. Primary school 18, taking a more liberal approach, entered seven pupils for the Level 6 Mathematics test and none achieved a pass mark.

Implications of and responses to test result security concerns

The set of issues outlined above, relating principally to the Level 6 tests but also testing in the strong accountability system in place in England more broadly, can be read as indicating uncertainty over the security of test results. For example, given the variation in selection practices outlined immediately above, it is likely that pupils with the same prior attainment would have been selected to enter the Level 6 test if had they been taught in one school but not if they were in another; and differences in preparation practices appear to be related to differing outcomes for pupils.

One significant implication of this security concern was a clear lack of trust amongst secondary schools in the results attained by primary pupils, demonstrated by the fact that 18 of the 20 secondary schools we spoke with retested all primary pupils at the start of their secondary education; as one secondary school leader noted,

> *[We have to] test to even out the variability of how primaries work. There's too much ambiguity about, some have opted out of [the tests], some take them earlier in the year etc. It all has such an impact on our results further up the school.* (Assistant principal, secondary school 7)

This was directly related in some cases to concerns about accountability pressures forcing primary schools to 'inflate' their scores by teaching to the test:

> [We are]…*sceptical of the consistency and the quality of the assessment that takes place, the reliability of the data we receive. Across primaries and within primaries we are sceptical of the whole thing... The primary schools have a vested interest in inflating their scores...* (Assistant head teacher, secondary school 4)

The practice of retesting early at secondary school was anticipated by most primary schools and was a source of particular frustration for many of their teachers, since it both called into question their professional judgement and meant that pupils' prior attainment was being disregarded, leading in one case to a school leader directly contacting secondary schools to try to prevent this happening:

> *It's been good preparation for secondary school. I hope to God they don't have to re-sit Year 6 all over again in Year 7 essentially. That's another reason why I think it's important… I fully intend to email the results to neighbouring schools when their results come back because*

> *I'm not clear if it will be reported to them. So I think it's our duty to do it myself.* (Deputy head teacher, primary school 12)

For some primary schools, this concern was the main reason for not taking part in the tests:

> *To get them somewhere and then when they go to high school we know the schools round here will CAT* [Cognitive Abilities Tests] *test them straight away and group them from the CAT tests. So what is a level 6 going to show them? (*Head teacher, non-participating primary school 35*)*

From the perspective of the secondary school, however, it was rational to treat Level 6 test data with caution. Secondary schools often expressed concern about differing interpretations of levels across primary schools, and different levels of engagement with the Level 6 test, making any systematic use of level 6 results to inform decision making logistically problematic. So, for example, one school respondent (primary school 12) noted that only one of their fellow primary schools that acted as 'feeder schools' to a local secondary school had engaged in the Level 6 tests, so other pupils that may have been able to attain a level 6 were not given the opportunity. This, then, limited the ability of the secondary school to judge the attainment of pupils across feeder schools consistently.

Taken together with the accountability pressures to demonstrate pupil progress, these issues - and the differing responses of primary and secondary teachers to them - led in most cases to increased difficulty for primary and secondary schools working effectively together as this final quotation demonstrates:

> *I have talked to the headteacher of the secondary school, looking at whether they could support us for next year, but they are not so keen on L6. Because they have then got to convert them through. Although they said they would support us, I didn't get the impression they were keen.* (Head teacher, primary school 1)

In the discussion below we develop a boundary object analysis which we argue can help explain and help move forward from what appeared, from this study, to be an impasse.

Discussion: Level 6 tests as Boundary Objects - Boundary Signifiers

This paper set out to address the question of how teachers and school leaders in primary and secondary schools respond to high stakes testing of pupils transitioning from primary to secondary school, using boundary theory to explore responses of teachers to Level 6 tests. To do so we need first to consider to what extent these findings indicate that the Level 6 tests acted as boundary objects.

By discussing the meaning of Level 6 test outcomes, and other issues relating to these tests, primary and secondary teachers reveal tensions around the boundary between primary and secondary educations over issues including:

- the extent to which Level 6 test results provide a true representation of pupil attainment, given differences in preparation practices, selection and 'teaching to the test'
- whether National Curriculum 'levelling' at primary and secondary is the same
- the role of accountability regimes in requiring schools to demonstrate progress
- whether the upper primary curriculum leads to a narrowing of pupil learning

These disagreements indicate that the tests have one characteristic of boundary objects: they allow for 'interpretive flexibility' (Star 2010: 602) since primary and secondary teachers understand the test to produce results to which they attach quite different meanings. For some, mainly primary teachers, the tests reveal high pupil attainment, which they value, and present externally to parents, other schools and school inspectors. For others - secondary teachers and some of the primary leaders that either decided not to enter pupils for the tests - the tests represent shallow learning and a stripped back curriculum. The Level 6 tests also share another characteristic of boundary objects: they act to reify concerns - they 'congeal' the complexities around pupil progress, attainment, ability into 'thingness' (Wenger, 1998: 58). Akkerman and Bakker (2011) refer to this process as *crystallization*: the crystallization of a set of concerns around the primary secondary border, usually not directly expressed, but brought out into the open by the tests.

However, there is one important way in which the tests seem to 'fail' in the words of Alkkerman and Bakker (2011): they do not appear to smooth working across boundaries. Thus whilst they had the potential to act as a boundary object as conceived of by Star, they fall short. As Oswick and Robertson (2009: 179) note, boundary objects can 'be perceived as barricades and mazes, reinforcing existing power structures and occupational hierarchies, and creating barriers to change' and in the strong accountability regime of England, this is apparent in the results presented above. To take into account this deviation from the 'ideal' boundary object definition as developed by Star and colleagues, we suggest that the Level 6 test acts as a specific kind of object that we refer to as a boundary signifier. We use the term boundary signifier for objects that operate to reveal tensions around boundaries - in this case, drawing out issues in relation to the primary-secondary divide - but that do not usually help overcome these tensions. Thus the tests help identify issues around the boundary that may otherwise be hidden, but they do not serve to act as 'a means of translation' (Star and Griesemer 1989: 393) to aid smoother cross-boundary working. On the contrary, by laying bare boundary tensions they can actually have the effect of shoring up such boundaries.

Whilst Level 6 tests did not act as ideal boundary objects in the study reported here, it is possible that they might. If primary and secondary school representatives can work together to understand each other's responses to a particular test in open dialogue, then this has the potential to enable them to work together more effectively to ease transition, and cross-boundary working. The system as it is would still mean that with many primary schools and few secondary schools there would be boundary problems. However, these might be easier to overcome without the tensions that are created within the current accountability regime. Differences and difficulties would still exist; but there wouldn't be the same incentive to ignore caveats and therefore it is likely that conversations would be easier and less loaded with distrust.

Level 6 tests themselves are now a footnote in English testing history, abandoned along with the use of National Curriculum levels in primary schools (to be replaced by a different set of tests). However the analysis reported here is of wider interest. Firstly, the analysis presented here could equally be applied to the standard National Curriculum Tests that were in operation well before the introduction of level 6 tests and are likely to be of value in considering the newer tests in place at the end of Primary Education in England that have replaced them. The new tests, introduced in 2016, focus on English, Mathematics and spelling, punctuation and grammar, indicating that the concerns highlighted in this paper (and of course the substantial wider research evidence) around the narrowing of the primary curriculum, teaching to the test and so on are likely to persist. Secondly, within jurisdictions across the world, tests operate at many boundary points in education, including the pre-school to school boundary; the infant to junior school boundary; the secondary school to further education boundary; and the further education to higher education boundary. It is likely that tests act as

boundary signifiers in each of these cases, especially where there are tensions evident. A boundary analysis that considers the tests in each of these cases as potential boundary objects or signifiers can help shed light on the way those on either side of the boundary think about and act towards tests which may help researchers understand these tensions more clearly.

Conclusion

In this paper we contribute to the substantive literature on organisational responses to assessment, presenting a qualitative analysis of primary and secondary teachers' and leaders' differing experiences and understanding of upper primary tests. We do so by providing a new theorisation drawing on boundary theory, arguing that the Level 6 test acted as what we term a boundary signifier, an object that reveals aspects of the nature of a social boundary, without acting in all respects as a boundary object as defined by Star. We argue this is due in part to the strong accountability regime in place in England that creates mistrust on both sides of the primary/secondary divide. We further argue that such tests could fully act as boundary objects if primary and secondary schools were enabled to work together effectively. Finally, we suggest that this new boundary signifier concept may be of value in other analyses of the primary/secondary divide and other boundaries in education systems.

(word count 9425 excluding references and title)

Acknowledgements

Thanks are due to Dr Mark Boylan, Professor Cathy Burnett and Professor John Coldron for comments on earlier drafts of this paper.

References

Aikenhead, G. (2001). Students' ease in crossing cultural borders into school science. *Science Education,* 85, 2.180-188. doi: 10.1002/1098-237X(200103)85:2

Akkerman, S. & Bakker, A. (2011). Boundary Crossing and Boundary Objects. *Review of Educational Research,* 81, 2, 132-169. doi: 10.3102/0034654311404435

Alexander, R. (2011). Evidence, rhetoric and collateral damage: the problematic pursuit of 'world class' standards. *Cambridge Journal of Education,* 41, 3, 265-286. doi: 10.1080/0305764X.2011.607153

Au, W. (2007). High-stakes Testing and curricular control: a qualitative metasynthesis. *Educational Researcher,* 36, 5, 258-267. doi:10.3102/0013189X07306523

Ball, S, Maguire, M, Braun, A, Perryman, J & Hoskins. K (2012). Assessment technologies in schools: 'deliverology' and the 'play of dominations'. *Research Papers in Education,* 27, 5, 513-533.

Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum. *Cambridge Journal of Education,* 41, 3, 287-302. doi:10.1080/0305764X.2011.607151

Bew, P. (2011). *Independent Review of Key Stage 2 testing, assessment and accountability*. London: Department for Education.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge: Cambridge university press.

Bourdieu, P. (1990). *The Logic of Practice*: *translated by R. Nice*. Stanford CA: Stanford University Press.

Boyle, B & Bragg, J (2006) A curriculum without foundation, *British Educational Research Journa*l, 32:4, 569-582. doi: 10.1080/01411920600775225

Boyle, B., & Bragg, J. (2008). Making primary connections: The cross-curriculum story. *The Curriculum Journal*, *19*(1), 5-21. doi: 10.1080/09585170801903183

Cobb, P., Zhao, Q. and Dean, C. (2009). Conducting design experiments to support teachers' learning: A reflection from the field. Jo*urnal of the Learning Sciences*, 18, 2, 165–199. doi: 10.1080/10508400902797933

Coldron, J., Crawford, M., Jones, S., & Simkins, T. (2014). The positions of primary and secondary schools in the English school field: a case of durable inequality. *Journal of Education Policy*, 30, 5, 671-687. doi: *10.1080/02680939.2014.972989*

Coldwell, M., Willis, B. & McCaig, C., (2013). Investigation of Key Stage 2 Level 6 tests. London: Department for Education.

Datalab (2015). *Seven things you might not know about our schools*. Retrieved from education data lab website: http://www.educationdatalab.org.uk/wp-content/uploads/2016/02/EduDataLab-7things.pdf

Department of Education and Science (DES) (1988) Education Reform Act (London, HMSO).

Engeström, Y., Engeström, R. & Kärkkäinen, M., (1995). Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities. *Learning and instruction*, 5, 4, 319-336. doi: 10.1016/0959-4752(95)00021-6

Galton, M., Gray, J. & Rudduck, J., (1999). *The impact of school transitions and transfers on pupil attainment and progress*. Norwich: HMSO.

Gillborn, D., & Youdell, D. (2000). *Rationing education: Policy, practice, reform and equity.* Buckingham: Open University Press.

Jones, K. (2003). *Education in Britain: 1944 to the present.* Cambridge: Polity Press.

Lamont, M. & Molnar, V. (2002). The study of boundaries in the social sciences. *Annual Review of Sociology,* 28,167-195. doi: 10.1146/annurev.soc.28.110601.141107

Mansell, W. (2007). *Education by Numbers: the tyranny of testing*. London: Politico's.

Mausethagen, S., (2013a). Accountable for what and to whom? Changing representations and new legitimation discourses among teachers under increased external control. *Journal of Educational Change*, 14, 4, 423-444. doi: 10.1007/s10833-013-9212-y

Mausethagen, S., (2013b). A research review of the impact of accountability policies on teachers' workplace relations. *Educational research review*, 9, 16-33. doi: 0.1016/j.edurev.2012.12.001

Ng, J. C. (2006). Understanding the impact of accountability on preservice teachers' decisions about where to teach. *Urban Review*, 38, 5, 353–372. doi: 10.1007/s11256-006-0038-2

Nolen, S.B., Horn, I.S., Ward, C.J. & Childers, S.A., (2011). Novice teacher learning and motivation across contexts: Assessment tools as boundary objects. *Cognition and Instruction*, 29, 1, 88-122. doi: 10.1080/07370008.2010.533221

Oswick, C. and Robertson, M. (2009). Boundary objects reconsidered: from bridges and anchors to barricades and mazes. J*ournal of Change Management*, 9, 2, 179-193. doi: 10.1080/14697010902879137

Schatzki, T. (2002). *The site of the social: A philosophical exploration of the constitution of social life and change*. Pennsylvania: The Pennsylvania State University Press.

Smith, K. & Davies, J., (2010). Qualitative data analysis *In:* L. Dahlberg & C. McCaig, eds. *Practical Research and Evaluation* (145-159). London: Sage Publications.

Star, S. (1989). *Regions of the mind: brain research and the quest for scientific certainty*. Stanford CA: Stanford University Press.

Star, S. (2010). This is not a boundary: reflections on the origin of a concept. *Science, Technology and Human Values*, 35, 5, 601-617. doi: 10.1177/0162243910377624

Star, S. & Griesemer, J. (1989). Institutional ecology: 'translations' and boundary objects: amateurs and professionals on Berkeley's museum of vertebrae zoology. *Social Studies of Science*, 19: 387-420. doi: 10.1177/030631289019003001

Stecher, B. & Barron, S. (1999). *Quadrennial Milepost Accountability Testing in Kentucky*. CSE Technical Report.

Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change,* 10, 211–227. doi: 10.1007/s10833-009-9105-2

Tilly, C. (1999). *Durable Inequality*. Berkeley CA: University of California Press.

Tilly, C. (2004). Social boundary mechanisms. *Philosophy of the Social Sciences,* 34, 211-236. doi: 10.1177/0048393103262551

Torrance, H (2011) Using Assessment to Drive the Reform of Schooling: Time to Stop Pursuing the Chimera?, *British Journal of Educational Studies*, 59:4, 459-485. doi: 10.1080/00071005.2011.620944
Tymms, P. & Merrell C. (2007). *Standards and Quality in English Primary Schools over Time: the national evidence*. Cambridge: Cambridge Primary Review Research.

Troman, G. (2008). Primary teacher identity, commitment and career in performative school cultures. British Educational Research Journal, 34(5), 619-633. doi: 10.1080/01411920802223925

Troman, G., Jeffrey, B., & Raggl, A. (2007). Creativity and performativity policies in primary school cultures. Journal of education policy, 22(5), 549-572. doi: 10.1080/02680930701541741

Webb, R., Vulliamy, G., Hämäläinen, S., Sarja, A., Kimonen, E., & Nevalainen, R. (2004). A comparative analysis of primary teacher professionalism in England and Finland. *Comparative Education*, 40, 1, 83–107. doi: 10.1080/0305006042000184890

Wenger, E.(1998). Communities of practice: Learning as a social system. *Systems thinker*, 9, 5, 2-3.

Whetton, C. (2009). A brief history of a testing time: national curriculum assessment in England 1989–2008. *Educational Research*, 51, 2, 137–159. doi 10.1080/00131880902891222

Wong, J. L. N. (2008). How does the new emphasis on managerialism in education redefine teacher professionalism? A case study in Guangdong Province of China. *Educational Review,* 60, 3, 267-282. doi: 10.1080/00131910802195869

Wyse, D & Torrance, H (2009) The development and consequences of national curriculum assessment for primary education in England, Educational Research, 51:2, 213-228. doi: 10.1080/00131880902891479