The impact of accountability measures on children and young people: emerging findings

Merryn Hutchings, Emeritus Professor, London Metropolitan University[1]

This research was commissioned by the National Union of Teachers (NUT). However, the analysis presented here is the author's and does not necessarily reflect the views of the NUT.

1. Introduction

This paper presents the emerging findings of ongoing research which aims to investigate the impact on children and young people of the various measures used to hold schools accountable. These include Ofsted inspections, floor standards, and the whole range of measures published in the school performance tables (attainment, pupil progress, attainment gaps, etc.). The full report will be published in summer 2015.

The study, commissioned by the NUT, draws together findings of relevant research with new data from:

- an on-line survey of teachers, completed by almost 8000 NUT members carried out between 21 November and 14 December 2014[2] ;
- case study visits to seven schools across the country, including primary, secondary and special schools, some rated 'Good' by Ofsted, and others as 'Requires Improvement'; in each school several members of staff and one or two groups of pupils were interviewed. Interviews were carried out in February and March 2015.[3]
- a survey of parents' views; this is not yet complete and is not included here.

This account of emerging findings first reviews evidence relating to the ways in which accountability measures are intended to benefit children and young people, and then identifies ways in which they have a negative impact.

[1] Research team: Professor Merryn Hutchings and Dr Naveed Kazmi

[2] Respondents were evenly split between primary (including early years) and secondary (including sixth form). They included a range of roles (e.g. headteachers, leadership posts, classroom teachers, supply teachers) and type of school (academies, maintained schools, special schools).

[3] In this paper,' interviewees' refers to teachers interviewed in the case studies.

2. Ways in which accountability measures are intended to benefit children and young people

Summary: accountability measures and their intended outcomes

This section argues that:

- There is evidence that high stakes testing results in an improvement in test scores because teachers focus their teaching on the test; however, higher test scores do not necessarily represent an increase in pupils' level of understanding and knowledge.
- There is no evidence that accountability measures can reduce the attainment gap between disadvantaged pupils and their peers.
- Accountability measures have achieved government aims of bringing about an increased focus on English, mathematics and (in secondary schools) academic subjects; however, this has been achieved at the cost of narrowing the curriculum that pupils experience, which teachers argued was detrimental to pupils.
- There is little evidence that providing market information to parents benefits pupils.

2.1. Introduction: intended benefits of accountability measures

The government has over time identified a number of different intended outcomes of accountability measure which are intended to benefit pupils by:

- improving attainment and progress;
- narrowing attainment gaps and thus increasing social mobility;
- ensuring that the qualifications that they study for are demanding, rigorous and a route to employment, and that all those leaving primary school are literate;
- providing information for parents to enable them to select effective schools for their children.

This section therefore considers evidence both from previous research and from the current study about the extent to which accountability measures are benefiting pupils in these ways.

2.2. Raising attainment and increasing pupil progress

While there is international evidence that external accountability has a significant positive impact on pupils' attainment in tests (e.g. Carnoy and Loeb 2002; Hanushek and Raymond 2005)[4], other research (e.g. Wiliam 2010) demonstrates that this does not necessarily indicate any greater understanding or knowledge, but simply that pupils have been taught how to succeed in that particular test. For example, Amrein and Berliner (2002), in a study of the impact of the introduction of high stakes testing in 18 US states, showed that while there was clear evidence that linking high stakes consequences to test outcomes had increased scores on those tests, use of a range of other tests showed no evidence of increased student learning.

[4] There is also evidence that it is possible for attainment to be high without having any high stakes accountability measures; Finland is an obvious example (Sahlberg 2011).

In our survey and case studies, teachers distinguished between test outcomes and pupils' overall level of knowledge and understanding; they argued that high test scores can be brought about by preparing pupils for a specific test, but that the scores they achieve do not necessarily imply having the level of skills and understanding that is needed as a foundation for future learning. Thus secondary teachers argued that the Key Stage 2 SATs scores that children arrive with in Year 7 overstate the level they have reached, and junior school staff said the same about infant schools.

Inevitably, high stakes testing results in teachers having to focus on the specific content that it is anticipated will be tested and on preparation for tests. This means that the amount of time spent teaching other aspects of the curriculum is reduced; this is discussed in Section 2.3.

The current pattern of statutory tests and examinations is intended both to measure the effectiveness of schools and to give useful formative feedback to learners (purposes which are not necessarily compatible). In our survey, only six per cent of teachers agreed 'a lot', and a further 40 per cent agreed 'a little', that "*Testing pupils helps them focus on what they do not understand/ know"*. There was a similar pattern in the responses to "*In this school testing and targets have helped raise attainment";* six per cent agreed 'a lot' and 50 per cent agreed 'a little'.

2.3. Social inequalities and attainment gaps

2.3.1. Attainment gaps and accountability measures

The policy of successive governments has emphasised the importance of increasing social mobility by reducing the gap between the attainment of disadvantaged pupils and their peers, and ensuring that both groups progress at the same rate. Information about gaps and pupil progress is published in the performance tables. Ofsted also has a particular focus on this. Schools can have above average attainment, but be judged less effective because of attainment gaps relating to specific pupil groups (disadvantaged, SEND, EAL). Despite the government focus on reducing gaps, including Pupil Premium payments, the attainment gap at GCSE level between pupils eligible for Free School Meals and those who are not has remained at about 27 percentage points throughout the last decade. There is no evidence that holding schools accountable will reduce attainment gaps, particularly in a context in which the economic gap between the richest and the poorest in society is increasing. Research has shown that the school effect contributes only 7-8 per cent of the variance in attainment between pupils (Wiliam 2010, drawing on OECD analysis); home background is very much the larger influence, and thus attainment gaps are very difficult to reduce. Some research (discussed in Section 3.1.2) has suggested that accountability measures have the opposite effect, tending to widen gaps because those with lower attainment become discouraged following poor test results, and lose motivation.

Strand (2014), analysing data about attainment gaps in relation to Ofsted judgements, concluded that current accountability mechanisms, such as performance league tables and Ofsted inspections, fail to adequately take into account factors associated with pupil background or the socio-economic makeup of the school, and are therefore biased against schools serving more disadvantaged intakes.

Our analysis showed that Ofsted grades are strongly related to the proportion of disadvantaged pupils in a school. More than half the schools in the lowest quintile for percentage of disadvantaged pupils have been judged to be 'Outstanding' by Ofsted, whereas this is the case for less than 15 per cent of those in the highest quintile of disadvantage. At the other end of the scale, less than one per cent of those schools in the lowest quintile are rated 'Inadequate' in comparison with 13 per cent of

the schools in the highest quintile of disadvantage. One interpretation of this is that schools serving affluent communities do a better job than those serving disadvantaged communities (and this is the view Ofsted take in the 2013 report *Unseen Children)*. However, an alternative interpretation is that Ofsted judgements do not adequately reflect the challenge faced by schools in disadvantaged communities, as Strand suggests. This was certainly the view expressed by the teachers we interviewed in a case study school in an area of high social and economic deprivation, with "*a history of negative Ofsted categories*". One implication of this pattern of Ofsted judgement is that disadvantaged pupils are more likely than their peers to be taught in schools judged 'Requires Improvement' or 'Inadequate', in which staff are likely to be more stressed and the pressures to be greater.

2.3.2. Pupil Progress

Measuring gaps between groups by reviewing the percentage of pupils reaching the expected level has the disadvantage that it ignores the fact that some groups of pupils are already 'behind' when they enter school. There has therefore been a shift in emphasis to considering progress made while at a school. However this still ignores the vast differences in children's experiences outside school. Our interviewees highlighted the variation in the home environment and parental support for children's learning, which means that disadvantaged pupils are unlikely to progress at the same rate as their more affluent peers. While Ofsted are aware that "differences in educational attainment between individuals will always exist" and that "family backgrounds have a strong influence on attainment", they assert that "factors such as material poverty … are not by themselves insurmountable barriers to success," and "the very best early years providers, schools and colleges make an enormous difference to the life-chances of children and young people" (Ofsted 2013, p18). Thus their argument is that all schools should be able to achieve as well as the best. This assumes, of course, that the social and economic conditions of all pupils eligible for Free School Meals are the same, and ignores the potential differences, for example, between being poor in a large economically thriving city and being poor in an area where there are no jobs. Material poverty is not all the same, and schools and individual pupils face different challenges.

2.3.3. Measuring 'the gap'

A further concern is that when only a small number of children in a school are disadvantaged, the specific characteristics of the individuals and their circumstances assumes greater importance, and may easily be very different from the national average pattern. Thus it is particularly unfair to penalise schools for large attainment gaps when pupil numbers are low. Interviewees argued that the DfE and Ofsted focus on the mean attainment of groups of pupils is problematic. Some schools (particularly primary schools and those with low pupil numbers) have very low numbers of disadvantaged or SEND pupils. It was argued that it is statistically unsound to measure attainment gaps and compare them with the national average in such cases. However, interviewees reported that the Ofsted inspectors in their schools had focused only on the group level data, and were not prepared to listen to information about individual circumstances.

2.4. Curriculum

The government uses accountability measures to reinforce its policies in relation to the curriculum. For example, the Key Stage 2 SATs and the focus on five A*-C GCSEs including English and maths both encourage schools to focus more strongly on English and maths; the EBacc and Progress 8 to ensure that more pupils study academic subjects to age 16; and the phonics test is intended to ensure a particular approach to the teaching of reading;

In our survey, 97 per cent agreed that there is "*an increased focus on maths and English teaching"*. The inevitable consequence of requiring a greater focus on certain subjects is that others are allocated less teaching time and are seen as less important. Thus despite government references to a broad curriculum, accountability measures tend to narrow the range of what is taught. Donaldson, in his recent review of the curriculum in Wales (2015, p10) asserts that "At its most extreme, the mission of primary schools can almost be reduced to the teaching of literacy and numeracy and of secondary schools to preparation for qualifications." Harlen and Deakin Crick's systematic review identified the emphasis on subjects tested as being at "the expense of personal and social development" (2002, p6). Recent reports on science education in primary schools (CBI 2015) and the creative arts (Neelands et al 2015) have drawn attention to the reduction in time spent on these subjects.

In primary schools, many teachers reported that the amount of time spent on maths and English increases in Year 6 in order to prepare for the SATs, and that other curriculum areas (such as music, art, design technology, topics) are consequently taught less, or not at all. An interviewee explained, "*at the top of Key Stage 2, definitely in Year 6 and to some extent in Year 5, the curriculum's narrowed to reading, writing and maths because that's what we're held accountable for and we've got to get those children to a certain level."* In secondary schools the amount of time spent on maths and English has also increased.

Both primary and secondary teachers pointed out that the lower-attaining pupils are often removed from other lessons to do extra maths and English, and that they therefore spend more of their time on these subjects, and experience a narrower curriculum than their peers: "*some of those children are really struggling and they're the children that are taken out in the intervention groups in the afternoons and then they miss out on the art and the PE and the history and the geography and the ICT*" (primary interviewee). This was also described in by a Key Stage 1 teacher in the survey:

> *These children are pulled out of broad curriculum subjects to try to close the gap. Their experience at school must be horrible - in assembly they've got to do phonics intervention, then a phonics lesson, a literacy lesson, a maths lesson, lunch, reading, extra reading intervention and then speech intervention. What else are they learning about the world? They are 6 years old, and all their school experience tells them is that they are failure (already) and have to be pulled out constantly to work on things their peers can already do, and miss out on the fun bits of learning.*

These pupils who miss out on much of the curriculum in order to concentrate on literacy and maths are often the disadvantaged pupils who are less likely to have access to wider learning and cultural opportunities outside school. Moreover, as interviewees explained, children can feel resentful about missing a lesson they enjoy.

While some of the pupils interviewed in the case studies accepted the dominance of English and maths because, they said, these are the most important subjects, many others questioned that analysis, arguing that what they learned in maths and English would not all be useful to them in the future. Some argued strongly that they should be learning more things that were practically useful, and several primary pupil groups argued for more science.

In secondary schools, a major impact of accountability changes is that pupils are being encouraged or, as a number of comments on the survey and interviewees claimed, "*forced*", to take academic rather than vocational courses, and this was resulting in loss of self-confidence and motivation, and often poor behaviour. In the survey 93 per cent of respondents said there is "an *increased focus on*

academic subjects" in their schools, and 86% of secondary teachers agreed that "*Pupils are encouraged to take subjects that will count in the league tables irrespective of their own interests/ aptitudes."*

The drive to focus on academic subjects was a considerable concern identified in the comments of many of the special school teachers who completed the survey, as well as in the case study special school. One teacher commented:

> *At post 16, I'm still expected to assess their maths and English despite the fact that they are 16 plus and still can't talk, toilet themselves or feed themselves. The life skills that I try to promote and independence skills, don't show up on any official chart, but this is where I try to concentrate. The curriculum is totally unrealistic for most of my school.*

The introduction of the phonics test was designed to influence how children are taught to read. Staff in the case study schools said that the phonics test had not improved pupils' reading, or informed what teachers do: one headteacher explained:

> *We did pretty poorly the first year that the phonics checker came out and then we practised for it the following year and our results were marvellous, but of course that took time from other elements of the curriculum. [And has that improved their reading?] We've always been a very strong reading school. …. It doesn't give our teachers any additional knowledge and it doesn't inform our planning.*

2.5. Providing market information to parents

Accountability measures were partly designed to improve information to parents so that they could make informed choices of schools which would benefit their children. It was assumed that this would have the effect of expanding popular (and successful) schools and forcing unpopular schools to close, and would therefore drive up standards. However, international research (Waslander et al 2010) has shown that markets have had very little effect; among the reasons for this are that parents consider school reputation and the characteristics of the pupils more important than performance data, and that they do not respond strongly to underperforming schools (e.g. by removing their children). In England, less than half of all parents say they used school performance data or Ofsted reports in choosing their children's schools (Francis and Hutchings 2012).

Three of the case study schools in this research had been judged by Ofsted to 'Require Improvement'. Coverage of these inspection outcomes in local newspapers was generally supportive of the schools and critical of Ofsted. Thus there appears to be some scepticism about the validity of Ofsted judgements, which reduce their value as market information. One eleven-year-old in a case study school judged to 'Require Improvement' said "*I told my mum about it, and she was like, I don't think that was fair, if [the Ofsted judgement was correct] you wouldn't be in this school right now."* While such scepticism exists, there is also undoubtedly a tipping point at which school reputation suffers, with consequent negative impact on the morale of teachers and pupils. Whether this results from mainly Ofsted judgements, league tables or simply local people's own observations is unclear; probably all three contribute.

3. Negative impacts of accountability measures on children and young people

Summary: negative impacts of accountability measures on children and young people

This section outlines a wide range of ways in which accountability measures have a negative impact on pupils:

- increased levels of anxiety, stress, mental health problems, disaffection, poor behaviour,
- less time to focus on pupils' social and emotional development
- reduction in the quality of teacher-pupil relationships
- pupils being asked to learn things for which they are not ready, experiencing less variety in lessons and a reduction in imaginative and creative approaches to teaching and learning;,
- a focus on borderline pupils at the expense of others;
- gaining an instrumental view of schooling;
- militating against inclusion.

While some of these affect all pupils, others disproportionately affect disadvantaged and SEN pupils.

3.1. Negative responses to tests and academic pressure

3.1.1. Test-related anxiety, stress and mental health problems

There is now substantial evidence that anxiety, stress and a variety of mental health problems have increased among young people, and the pressure to achieve in school work and tests/ examinations is among the causes (e.g. *The Times,* 12 March, 2015). The Word Health Organisation (2012) found that 11-year-old and 16 year-old pupils in England feel more pressured by their school work than is the case in the vast majority of other European countries.

Our survey showed that:

- over 90 per cent of teachers agreed that "*Many pupils become very anxious/stressed in the time leading up to SATs/public examinations;"*
- 76 per cent of primary teachers and 94 per cent of secondary agreed that "*Some pupils in this school have developed stress-related conditions around the time of SATs/public exams."*

In interviews and comments on the survey, teachers identified the group most affected by test anxiety and stress-related conditions as including being high-attaining and conscientious pupils, often girls, but said that some low-attaining pupils and some with special needs also suffered extreme stress. One secondary teacher wrote: "*Many girls self-harm, have panic attacks and emotional problems because they cannot be 'perfect.'*" (See similar evidence in Harlen and Deakin Crick 2002). But teachers identified a wide range of pupils who suffered from depression, self-harm, thoughts of suicide, and eating disorders. While acknowledging that there are other causes of stress among young people, teachers claimed that stress about exams or tests was often the immediate trigger. For example a primary teacher reported in the survey, "*Last year I had a year six pupil turn to physical self-harming which she attributed to the pressure she felt to achieve a level similar to that of her peers, and to hit a level four in her SATs (she is severely dyslexic and an incredibly hard worker).*"

The school case studies showed that the main cause of the stress/anxiety was simply the fact of having to do test or exams in which there is a real possibility of failure. This is exacerbated by the fact that school practices make the importance of tests and exams very clear to their students. The extent to which schools emphasise tests and exams varies, but even where staff said that they tried to "*protect*" their students from the pressure, pupils talked about ways in which teachers reinforced the importance of tests and exams, for example, by regular mentions of the SATs (a Year 5 pupil reported, "*our teacher, she's like, if you don't listen in class you're not going to do very well in the SATs, you're going to fail or you're not going to get good marks*"). Such comments, together with actions such as organising booster groups, made pupils feel under pressure. In some schools, teachers talked about "*pushing*" students, and the negative effect of this on pupil teacher relationships.

Another factor that increases stress for some pupils is the way pupils talk among themselves about levels and test outcomes. Primary pupils said that classmates sometimes boasted about the levels that they had reached ("*I'm a 5b"*) or jeered at those who were less successful. A teacher reported a conversation where a child who would not be taking the SATs was put down by a classmate ("*you're not even taking the SATs"*). It is unfortunate that levels of attainment and test results have provided fresh ammunition for children to use to put one another down.

3.1.2. Disaffection

Disaffection is a second way in which tests and the drive to raise standards impact on some pupils. This pattern of low achievers becoming "overwhelmed by assessments and demotivated by constant evidence of their low achievement", which then further increased the gap between low and high achieving students, was highlighted by Harlen and Deakin Crick (2002). In our case studies and survey comments disaffection and demotivation was described in all age groups and types of school. Interviewees described pupils who were aware that they were doing less well than others in the class, or who found it difficult to understand what they were being taught. The consequent loss of self-esteem and motivation resulted in disaffection, which sometimes manifested itself as disruptive behaviour.

In our survey, 96 per cent of teachers agreed that "*When pupils know they are doing less well than others in class and in tests, their confidence and motivation suffers"* (with 70 per cent agreeing 'a lot'). There were no differences in these figures across school phases or different Ofsted categories. Pupils interviewed also talked about the negative impact of poor marks on self-confidence and motivation; a Year 6 pupil explained, "*it makes people that aren't as good and don't have enough confidence in themselves less confident, have, like, less confidence."*

It is clearly a concern that any children becomes disaffected, but is a particular concern with the youngest children. Heyman et al (1992) found that 5 and 6 year olds who failed in a task were more likely to make global negative self-judgements ('I am no good'), whereas older children were more likely to compartmentalise, and say they were no good at that specific area of task (e.g. 'no good at maths').

The drive for every pupil to take rigorous academic subjects and the devaluing of vocational subjects has also contributed to disaffection in secondary schools. Teachers argued that some pupils are now studying and being examined on courses that are inappropriate for the level they have reached, and that this has a negative impact which manifests itself in a variety of ways (poor behaviour, low self-esteem, etc.) A secondary teacher interviewee said:

My year group went through the options process last year and … we didn't choose the subjects for them but we had to force them down a certain pathway more than we would have done in the past, so we had many more students having to pick history, geography, to do a computer science or a language, many more students having to be encouraged along that pathway which isn't suitable for all of them, and for some students, the things that we might have offered in the past which would have really suited them, now we can't offer, and they're some of the ones that are actually now causing us the most problems because they're not engaged.

3.1.3. Summary: Negative responses to tests and academic pressure

Disaffection and anxiety are not distinct categories; some of the accounts we collected described pupils who had become extremely anxious about tests, and the longer term outcome was that despite having high attainment, they become disaffected.

These pupil responses to testing and academic pressure are of even greater concern because teachers argued that there is now less time to focus on pupils' social and emotional development; 84 per cent agreed that "*The focus on academic targets means that social and emotional aspects of education tend to be neglected.*"

It is somewhat ironic that the recent *Times* manifesto on young people's mental health (12 March 2015) recommends that Ofsted should inspect emotional support and mental health provision in schools, rather than tackling the causes of the problems, by reducing the emphasis on high stakes testing and the way in which Ofsted reinforces the importance of this.

3.2. Impacts on quality of pupil-teacher relationships

Donaldson (2015 p10) reviewing the curriculum and assessment in Wales, argued that one of the impacts of the high level of prescription and "increasingly powerful accountability mechanisms" is that the key task for many teachers has become "to implement external expectations faithfully, with a consequent diminution of … responsiveness to the needs of children and young people."

Our survey showed that many teachers felt that the quality of their relationships with pupils had been reduced by:

- pressure to cover the syllabus and maintain focus in lessons (and thus less time to deal with pupil distress, or to allow pupils to talk about their own experiences and the things that interest them);
- lack of time as a consequence of teachers' workload: 96 per cent of survey respondents agreed (and 76 per cent agreed 'a lot') that "*I do not have enough time to focus on the needs of individual pupils*". Some commented in the survey that they did not know their pupils as well as they did in the past;
- teachers' stress levels : 93 per cent of survey respondents agreed that "*My stress levels sometimes impact on the way I interact with pupils*". Many interviewees also commented that this was the case.

While pupils were aware that their teachers were feeling stressed during Ofsted inspections, they were less aware of ongoing stress among teachers (though said they were sometimes "*grumpy*"). Most said their teachers had time to offer them the support they needed. This suggests a high level of teacher professionalism, but may also reflect pupils' experience in schools; teachers have been stressed and overworked for many years.

However, teachers were very clear that the quality of their relationships was less good than it had been. One primary teacher wrote on the survey, "*I have less time to get to know individual pupils and rarely have show and tell, which is a shame as I teach mixed years 1 and 2.*" Another argued that pressures on her time meant that she was now less likely to be aware o*f "things seriously wrong in pupils' lives"* and to refer them in relation to child protection.

3.3. Impacts on teaching and learning

3.3.1. Pupils being asked to learn things for which they are not ready

The coalition government has introduced a curriculum designed to be "challenging and ambitious", and which includes "more demanding content" at earlier ages (Gove, 2013). Teachers reported that this means they have to "*push*" some pupils to learn things that they are not ready to learn. Almost 90 per cent of teachers agreed in our survey that this happens.

In the early years, teachers described having to make children sit down and tackle academic work in a way that was inappropriate to their level of emotional maturity. This was leading to "*silly*" behaviour and lack of motivation, particularly among summer-born boys. The introduction of the phonics test contributes to this pressure. A Year 1 teacher wrote in the survey:

> *This term we have seen year 1 pupils become anxious about not keeping up with the rest of the class. They feel they do not have enough time to finish work. Due to raised expectations of National Curriculum teachers have felt the need to increase maths and spelling homework in year 1. Parents have commented that they are concerned by the expectations and that their child is not ready. Some year 1 children are not ready for a formal style of learning but teachers feel under pressure to make progress despite knowing that socially and physically the children need more time to learn through play.*

Accounts were given of the impact of encouraging pupils to take academic GCSEs for which they were not ready. As discussed in Section 2.3, many teachers in special schools commented on the inappropriateness of trying to teach their students aspects of the academic curriculum for which they were not ready or which were not appropriate to their circumstances.

3.3.2. Lack of variety in lessons

There is extensive international evidence, reviewed by Lobascher (2011), that high stakes testing and accountability measures discourage creative teaching. Many teachers in our survey and case studies described the tendency for all lessons to have similar structures. They said that this was a result of the pressure of academic targets, the perceived need to cover the curriculum, and perceptions of what Ofsted require, or used to require; many teachers still use the lesson structure they adopted some years ago when Ofsted focused on the three part-lesson, because this way of teaching has become "*drilled in*". A requirement for uniformity in lesson plans, marking, displays and even PowerPoints was reported by a significantly higher percentage of those working in vulnerable and challenging schools (those with low attainment or negative Ofsted judgements or with a higher number of disadvantaged pupils). Teachers in one of the case study 'Requires Improvement' schools commented that the staff there had previously prided themselves on the imaginative and creative lessons they offered, but that in preparation for their next inspection they had moved to more uniform (and dull) lesson structures.

In our survey, 93 per cent of teachers agreed that "*the focus on academic targets means there are fewer opportunities for creative, investigative and practical activities"* while only 16 per cent agreed that "*pupils have ample opportunities for investigation/ exploration/ play."* Responses indicated that

even those in early years settings felt the impact of academic targets, though to a lesser extent than their primary and secondary counterparts; 87 per cent of teachers in early years settings agreed with the first statement and 57 per cent with the second one. The lack of creative, investigative and practical activities was reported by a significantly higher percentage of those in vulnerable and challenging schools. The survey also showed that stories play much less role in schools than they used to; even in primary schools, two thirds of teachers agreed that "*pupils rarely have opportunities to read/listen to stories for pleasure in school".*

A large majority (83 per cent) of teachers agreed that "*pupils do not have not enough time to reflect."* A secondary interviewee explained that impending exams meant she felt under pressure to cover the syllabus rather than to allow time for reflection and consolidation of learning.

Pupils in the case study schools said they preferred lessons that were "*different"*; they talked positively about lessons where they made models, engaged in role play, etc. They argued that they learned more in such lessons because they were memorable.

3.3.3. Focus on borderline pupils

The focus on the percentage of pupils achieving five A*-C GCSEs including English and maths has resulted in many secondary schools focusing their attention on pupils who are borderline C/D grade in certain subjects (Ball et al 2012). This can be visible in staff rooms with photos of the relevant pupils. This inevitably means that there is less focus on both the high-achieving and low-achieving pupils. Our survey showed that such a focus is also common practice in primary schools, with 70 per cent of primary teachers and 79 per cent of secondary agreeing that *"explicit focus on borderline students"* is a *key* strategy in their school, and altogether 94 per cent of teachers saying that this strategy was used. A primary teacher interviewee reported:

> *We have to get a certain number of children to make the required amount of progress … And also the percentage of Level 4 plus for reading, writing and maths. So if a child is good at reading and writing but their maths is going to knock them out of that percentage then they're boostered in maths.*

Similarly a secondary teacher noted on the survey, "*Because the focus is on C/D borderline pupils and A/A* pupils, those that have worked hard for a D or E grade, or who are aiming around B grades, feel unappreciated*." While the introduction of Progress 8 should reduce this focus, it has not yet done so.

3.4. The development of an instrumental view of education

One of the impacts of high stakes testing is that many pupils see education entirely in terms of tests and qualifications. This develops as pupils go through their school careers. Thus while 62 per cent of primary teachers agreed that "*Most pupils think it is only worth learning things that are tested*," the equivalent figures for secondary teachers and those in sixth forms were 90 per cent and 96 per cent.

Similarly a majority of teacher survey respondents agreed that "*Pupils are more concerned with test outcomes than with learning for interest*" (primary 77 per cent, secondary 95 per cent, sixth form 98 per cent), and that "*Many pupils see the main purpose of schooling as gaining qualifications rather than gaining a rounded education"* (primary 88 per cent, secondary 98 per cent).

Secondary students interviewed (especially sixth formers) talked about the waste of time of having to continue with subjects other than those they were taking in exams. All the pupils interviewed (including those in primary schools) asserted that SATs or GCSE results would influence and potentially limit their future options.

3.5. The negative impacts of accountability on inclusion

Galton and MacBeath's (2015) report on inclusion draws attention to the reluctance of some schools to take on pupils who are likely to lower test scores, because of the assumed Ofsted reaction. They describe some schools setting limits to the number of such pupils they admit.

This was an issue raised in the case studies. For example, one single form entry primary school had experienced a fall of ten points in the percentage of pupils achieving the expected level, and attributed this to three statemented pupils who failed to score. One of these could have achieved Level 4 but the pressure of the test made her distraught on the day. The head had unsuccessfully asked for these pupils to be disapplied. The "*drop*" in attainment had attracted attention from the local authority, and would presumably be viewed negatively by Ofsted.

While headteachers interviewed argued that they prided themselves on the work their schools did with disadvantaged and SEND pupils, they were also acutely aware of the impact that their intakes could have on test results, and thus on Ofsted judgements. Thus increasing inclusion was seen as a risky option.

The headteacher of Burlington Danes academy recently spoke out on the covert selection strategies that some secondary school heads use to ensure an intake of high attaining, and in some cases, affluent, pupils, and thus avoid the potential negative impact on attainment of disadvantaged and SEND pupils (*The Independent*, 24 March, 2015)

It is clearly a major concern if the way that Ofsted views the attainment of disadvantaged and SEND pupils is encouraging schools to become less inclusive, and Galton and McBeath's report makes a number of recommendations in relation to this.

4. In conclusion

This paper presents emerging findings from ongoing research; the full report on the research will be published in summer 2015. It discusses some of the evidence that we have collected so far about both the achievements of accountability measures in relation to their aims, and the vast range of negative impacts on pupils. Clearly many of these effects are unintended consequences of initiatives by policy-makers. As Donaldson (2015) put it, "The unintended effects of over-exuberant accountability can unintentionally compromise good intentions."

All the issues above have been highlighted in research and reports over the years; only a very small selection of the relevant literature has been mentioned in this paper. For example, as long ago as 1896, Emerson White discussed "the propriety of making the results of examinations the basis for … determining the comparative standing or success of schools". His conclusions are still relevant:

> They have perverted the best efforts of teachers, and narrowed and grooved their instruction; they have occasioned and made well-nigh imperative the use of mechanical and rote methods of teaching; they have occasioned cramming and the most vicious habits of study; they have caused much of the overpressure charged upon the schools, some of which is real; they have tempted both teachers and pupils to dishonesty; and, last but not least, they have permitted a mechanical method of school supervision. (quoted in Wiliam 2010, p7)

More recently, Jones and Egley (2004) conducted a survey of teachers in Florida to explore the impacts of their high stakes testing program, and reported negative effects on the curriculum, teaching and learning, and student motivation. In England, in 2008 the House of Commons Children

Schools and Families Committee concluded that "a variety of classroom practices aimed at improving test results has distorted the education of some children, which may leave them unprepared for higher education and employment." They went on to identify the narrowing of the curriculum, and argued that "a focus on test results compromises teachers' creativity in the classroom and children's access to a balanced curriculum" (2008, p3). They pointed to shallow learning, pupil stress and demotivation, and a disproportionate focus of resources on the borderline of targets. Their recommendations included reform of the current system of national tests to separate out the various purposes of assessment.

And yet, despite this mass of evidence, rather than reducing the accountability pressures on schools, politicians continue to increase them. In particular, the effects of testing have been exacerbated by Ofsted's increased focus on data and attainment gaps. Thus, for our case study teachers and headteachers, Ofsted posed the most worrying threat.

Contact: m.hutchings@londonmet.ac.uk

References

Amrein, A.and Berliner, D. (2002) High-Stakes Testing, Uncertainty, and Student Learning, *Education Policy Analysis Archives,* 10(18).

Ball, S., Maguire, M. and Braun, A. (2012) *How schools do policy: policy enactments in secondary schools*, Abingdon: Routledge.

Carnoy, M. and Loeb, S. (2002) Does external accountability affect student outcomes? A cross-state analysis, *Educational Evaluation and Policy Analysis,* 24(4), pp305-331.

CBI (2015) *Tomorrow's World: Inspiring Primary Scientists*, CBI.

Donaldson, G. (2015) *Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales* at https://hwbplus.wales.gov.uk/schools/6714052/Documents/Donaldson%20Report.pdf

Francis, B. and Hutchings, M. (2013) *Parent Power: Using money and information to boost children's chances of educational success,* Sutton Trust.

Galton, M. and McBeath, J. (2015) *Inclusion: Statements of Intent*, Cambridge University and NUT.

Gove, M. (2013) Curriculum, exam and accountability reform: oral statement to parliament, 7 February 2013, https://www.gov.uk/government/speeches/curriculum-exam-and-accountability-reform.

Hanushek, E. and Raymond, M. (2005) Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), pp297-327.

Harlen, W. and Deakin Crick, R. (2002) *A systematic review of the impact of summative assessment and tests on students' motivation for learning*, EPPI-Centre Review. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Heyman, G., Dweck, C., and Cain, K. (1992) Young children's vulnerability to self-blame and helplessness: relationship to beliefs about goodness, *Child Development*, 63, pp401-415.

House of Commons Children Schools and Families Committee (2008) *Testing and Assessment: Third Report of Session 2007-8, Volume 1*.

Independent (2015) Leading headteacher exposes 'underhand' tactics used by schools to get round curbs on selection, *The Independent*, 24 March.

Jones, B. and Egley, R. (2004) Voices from the Frontlines: Teachers' Perceptions of High-Stakes Testing, *Education Policy Analysis Archives*, 12(39).

Lobascher, S. (2011) What are the potential impacts of high-stakes testing on literacy education in Australia? *Literacy Learning: The Middle Years*, 19(2), pp9-19.

Neelands, J., Belfiore, E., Firth, C., Hart, N., Perrin, L., Brock S. Holdaway, D., and Woddis, J. (2015) *Enriching Britain: Culture Creativity and Growth,* Coventry: The University of Warwick.

OECD (2014) *Country Note: United Kingdom: Results from PISA 2012*, http://www.oecd.org/unitedkingdom/PISA-2012-results-UK.pdf

Ofsted (2013) *Unseen Children: access and achievement 20 years on: Evidence report*, Ofsted.

Sahlberg, P. (2011) *Finnish Lessons: What can the world learn from educational change in Finland?* New York: Teachers' College Press

Strand, S. (2014) *Moderators of the FSM achievement gap: being more able or poor in an affluent school*, paper presented at British Educational Research Association conference, University of Oxford on Wednesday, 24 September.

Times (2015) *True scale of child mental health crisis uncovered*, *The Times*, 12 March 2015, p1, p8.

Waslander, S., Pater, C. and van der Weide, M. (2010) *Markets in Education: an analytical review of empirical research on market mechanisms in education.* OECD Education Working Papers No 52, OECD Publishing.

Wiliam D. (2010) Standardized testing and school accountability, *Educational Psychologist* , 45 (2), pp. 107-122.

World Health Organisation (2012) *Social determinants of health and well-being among young people: health behaviour in school-aged children study: international report from the 2009/2010 survey*, http://www.euro.who.int/__data/assets/pdf_file/0003/163857/Social-determinants-of-health-and-well-being-among-young-people.pdf .