Rethinking assessment and inequality: the production of disparities in attainment in early years education

Alice Bradbury

Department of Educational Foundations and Policy Studies, Institute of Education, University of London, London, UK

Available online: 22 Jun 2011
Rethinking assessment and inequality: the production of disparities in attainment in early years education

Alice Bradbury*

Department of Educational Foundations and Policy Studies, Institute of Education, University of London, London, UK

(Received 23 November 2010; final version received 28 February 2011)

Despite decades of research and debate, the issue of unequal outcomes continues to be a concern in educational systems worldwide. In England, published data relating to pupils’ attainment across ethnic groups and by class indicators has been used to demonstrate continued inequalities in schools. This article attempts to deconstruct the relationship between assessment results and inequality by questioning the assumption that results only record inequality, rather than being implicated in its production. Interview data related to the case of a statutory teacher assessment system in early years education are used to show how assessment results may be influenced by pressure from external advisors, who only recognise certain patterns of results as intelligible. These recognisable patterns, it is argued, relate to wider discourses of class, race and the ‘inner city’, through which the pupils in these schools are constituted as inevitably low attaining. In addition, monitoring systems based on ‘value added’ methodologies provide an incentive to deflate assessment results in this first year of school. The article concludes that we need to rethink exactly what apparent disparities in assessment results actually represent, particularly given the increasing use of teacher assessment in the school system in England.

Keywords: assessment; inequality; race; class; education policy; early years education

Introduction

Despite decades of research and debate, the issue of unequal outcomes continues to be a concern in educational systems worldwide. In England, research on educational inequalities in the education system has largely focused on pupils in the secondary school system (age 11–16), partly due to the use of GCSE exam results as evidence of significant disparities in attainment by class indicators, ethnic group and gender. In this paper, the focus is shifted to differences in attainment in assessments conducted at the very beginning of children’s school careers, in Reception classes at age five. Using data from ethnographic studies of two Reception classes in London, I discuss the impact of a statutory assessment system in terms of the production of inequalities, rather than the apparent recording of these. I argue that this policy provides an example of how differences in attainment can be created, maintained and become embedded in the education system.

*Now at Department of Education, Roehampton University, London, UK. Email: alice-bradbury@roehampton.ac.uk
The introduction of a standardised statutory assessment at age five in England revealed patterns of inequality similar to those found in assessments for older children (Gillborn 2006b; Gillborn 2008). The assessment, originally called the Foundation Stage Profile and renamed the Early Years Foundation Stage Profile (EYFSP) in 2008, was introduced in 2003 as part of the formalisation of the early years curriculum. The new assessment replaced local and non-statutory systems of assessing children in Reception classes, the first compulsory year of school. The EYFSP is a detailed assessment of all areas of children’s learning, conducted entirely through teacher assessment. Results from England since the introduction of the EYFSP show that high proportions of boys, children on free school meals (FSM) and from minority communities are designated as failing to reach ‘a good level of development’ (the government’s term for their benchmark EYFSP score); this is an assessment which labels significant proportions of children as already ‘behind’ in their first year of school. For example, only 26% of boys on FSM reached ‘a good level of development’ in the 2008–2009 academic year (DCSF 2010).

This set of results (see Table 1), which are broadly similar to the previous years (DCSF 2008; 2009; 2010), provides a complex picture of widely different levels of attainment. While all of the highest five attaining groups (shown in blue online, darker shading in print) defined by FSM status, gender and ‘ethnic group’ were non-FSM girls, all of the five lowest groups (shown in yellow online, lighter shading in print) were boys on FSM. The White British category provides some of the highest and the lowest proportions of children reaching the benchmark. Within each gender and FSM category, there are distinct and reasonably consistent differences between ethnic groups, with the smallest proportions of children from the Pakistani, Bangladeshi and Black groups reaching ‘a good level of development’.

However, as I argue through this article, the relationship between assessment results and educational inequalities needs to be scrutinised – to what extent can percentages such as those mentioned above be taken as representing or recording

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Non-FSM boys</th>
<th>Non-FSM girls</th>
<th>FSM boys</th>
<th>FSM girls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>White British</td>
<td>47.8</td>
<td>67.2</td>
<td>24.9</td>
<td>42.2</td>
<td>53.9</td>
</tr>
<tr>
<td>Any other White background</td>
<td>39.0</td>
<td>54.5</td>
<td>28.6</td>
<td>41.2</td>
<td>45.2</td>
</tr>
<tr>
<td>White and Black Caribbean</td>
<td>41.8</td>
<td>62.1</td>
<td>30.5</td>
<td>47.5</td>
<td>47.8</td>
</tr>
<tr>
<td>White and Black African</td>
<td>44.1</td>
<td>65.3</td>
<td>28.8</td>
<td>47.9</td>
<td>51.2</td>
</tr>
<tr>
<td>White and Asian</td>
<td>52.4</td>
<td>70.8</td>
<td>26.6</td>
<td>48.0</td>
<td>57.7</td>
</tr>
<tr>
<td>Any other mixed background</td>
<td>47.5</td>
<td>64.7</td>
<td>31.1</td>
<td>46.8</td>
<td>52.1</td>
</tr>
<tr>
<td>Indian</td>
<td>49.3</td>
<td>66.1</td>
<td>30.9</td>
<td>52.6</td>
<td>56.5</td>
</tr>
<tr>
<td>Pakistani</td>
<td>33.5</td>
<td>47.6</td>
<td>26.6</td>
<td>40.6</td>
<td>39.2</td>
</tr>
<tr>
<td>Bangladeshi</td>
<td>30.0</td>
<td>48.8</td>
<td>25.3</td>
<td>42.2</td>
<td>38.0</td>
</tr>
<tr>
<td>Any other Asian background</td>
<td>42.8</td>
<td>58.2</td>
<td>27.8</td>
<td>47.1</td>
<td>49.0</td>
</tr>
<tr>
<td>Black Caribbean</td>
<td>36.2</td>
<td>56.6</td>
<td>27.4</td>
<td>45.9</td>
<td>43.3</td>
</tr>
<tr>
<td>Black African</td>
<td>38.6</td>
<td>57.2</td>
<td>28.4</td>
<td>47.9</td>
<td>44.5</td>
</tr>
<tr>
<td>Any other Black background</td>
<td>36.3</td>
<td>56.1</td>
<td>25.0</td>
<td>45.1</td>
<td>42.4</td>
</tr>
<tr>
<td>Chinese</td>
<td>45.6</td>
<td>60.2</td>
<td>36.4</td>
<td>42.9</td>
<td>52.0</td>
</tr>
<tr>
<td>Any other ethnic group</td>
<td>36.3</td>
<td>49.7</td>
<td>26.7</td>
<td>42.1</td>
<td>40.5</td>
</tr>
<tr>
<td>Total</td>
<td>45.9</td>
<td>64.7</td>
<td>26.3</td>
<td>43.2</td>
<td>51.7</td>
</tr>
</tbody>
</table>
educational inequalities, in terms of experiences in schools and levels of attainment? I use data from interviews with Reception teachers to show how assessment systems themselves also produce inequality, and therefore contribute to the embedding of differences within the school system.

**Background: assessment policy and inequality**

Previous research from other stages of the education system in England has commented on the entwinement of assessment and associated school practices with the maintenance of inequalities by race and class. Most notably, Gillborn and Youdell’s (2000) study of two secondary schools’ practices described how pressure for greater numbers of higher GCSE grades resulted in systems of ‘educational triage’, whereby some students were designated ‘hopeless cases’, unlikely to gain a high enough grade to affect statistics based on A–C grades. This group of students were more likely to be pupils on free school meals or from minority communities. Gillborn and Youdell’s findings on practices related to ‘tiered’ exam entry (see also Gillborn 2006b), whereby some secondary school GCSE students are entered for exams in which the highest mark they can possibly achieve is a C grade, also demonstrated how assessment practices may contribute to the production of inequalities in and of themselves. For these pupils, who were mainly Black students, their exam results (and therefore their attainment in comparison with other groups) are limited by teachers’ choices. Thus, the exam results are not merely records of different attainment, but the product of discriminatory practices.

Other research on assessment policies has focused on the dramatic impact of statutory testing on school practices in general; Stobart (2008) argues that systems such as using ‘booster’ classes to ensure primary pupils gain the required level in their Key Stage 2 tests are now commonplace in schools. In this project, I aimed to explore the classroom practices associated with statutory assessment in the early years; I found, however, that specific assessment practices themselves were also implicated in the production of disparities in attainment.

Although there have been studies focused on race and class inequalities in primary schools in England (Connolly 1998; Wright 1992), there has been little research into the impact of the significant reforms to primary and early years education of the 2000s. The greater formalisation of the early years curriculum in England, which is distinctive in terms of early childhood education policy internationally, has rarely been discussed in relation to educational inequalities. Furthermore, despite its increased use in other statutory tests (Key Stage 1 and Key Stage 3), there is as yet no substantial research on the impact of teacher assessment on disparities in attainment. The shift towards teacher assessment as a tool in statutory assessments has been welcomed by teaching unions and critics of the practices associated with tests (NUT 2010); however, there is no reason to assume that teacher assessment will produce more equitable outcomes.

**How the EYFSP works**

The EYFSP assessment relies entirely on teachers’ judgements, based on their observations of the pupils throughout the school year. The assessment comprises 13 scales, based on the 6 areas of learning that make up the Early Years Foundation Stage curriculum; these scales are:
Each of these scales is made up of nine numbered statements, which the teacher decides to award or not award, depending on their knowledge of the child. These 117 statements, commonly known as EYFSP ‘points’, define what is valued in Reception pupils; they extend to all elements of classroom life. The teacher is expected to collect evidence in the form of observation notes, photographs and children’s work to demonstrate these points if they decide to award them. Some examples of the points and the kinds of observations that would function as evidence of them (taken from the guidance booklet on the EYFSP) are provided below: (see Table 2)

During the summer term, the class teacher (sometimes in consultation with other involved adults) decides which of the EYFSP points to award for each child, giving each a total score out of 117. The completion of these assessments for class of between 20 and 30 children is therefore a lengthy and complicated process.

The increased formalisation of early years assessment is part of wider process of transformation in early childhood education in England and elsewhere (Cannella and Viruru, 2004; Dahlberg and Moss 2005; Yelland 2010). This transformation has included the introduction of formal curricula, expected levels of attainment and accountability measures to provision for young children in school and pre-school settings. In England, although principles of ‘learning through play’ are preserved in the Early Years Foundation Stage (EYFS) curriculum, the more detailed planning requirements, record keeping and monitoring involved have been criticised for being overly formal and the EYFS has been derided in the press as the ‘nappy curriculum’ (BBC News 2010; Beckford 2008; Woolcock 2010).

My focus in terms of analysing the EYFSP policy has been on how it represents the spread of accountability measures into Reception classes and the impact of this move in terms of equality. Although the EYFSP results are not published in league tables5 and so are not formally part of the market-driven provision of information to parents, they are used by Ofsted and local authority inspectors to assess a school’s early years provision. Furthermore, school-level results may be given to prospective parents. Thus, EYFSP results are part of a framework of assessment based on accountability and monitoring, which is a key element of the neoliberal agenda in education (Apple 2001). Research on the impact of monitoring systems such as inspections or league tables based on test results has found that they have a real impact on teachers and their practices (Jeffrey and Woods 1998; Seddon 1997; Troman 2000). In Ball’s work, discussion of this impact focuses on the ‘terrors of performativity’ (2003). He describes ‘performativity’ as:
Table 2. FSP points and examples of observation evidence (from QCA 2008).

<table>
<thead>
<tr>
<th>EYFS Profile Point</th>
<th>Example of observation that would demonstrate the point</th>
</tr>
</thead>
</table>
| The child is confident to explore new experiences and talk about them with adults and peers.  
(Personal, Social and Emotional Development – Dispositions and Attitudes Point 7) | Jack and Ellie are playing outside with the cars. Ellie decides to get the blocks to build a garage. |
| The child attempts to read more complex words.  
(Communication, Language and Literacy – Linking Sounds and Letters Point 8) | Walking round the sea life centre Jack noticed the sign ‘crabs’ and says ‘We haven’t seen the crabs, yet.’ |
| The child consistently recognises numerals in a range of contexts.  
(Problem-solving, Reasoning and Numeracy – Numbers as Labels and for Counting Point 5) | Keith walks past the year 5 classroom. He notices the year group sign and says ‘Oh look, a five. I’m five.’ |
| The child knows about technology and its use in his or her life and local environment. The child exploits the technological opportunities around him- or herself to enhance his or her learning.  
(Knowledge and Understanding of the World – Point 7) | Tamsin picks up a tin of beans in the ‘shop’. She holds a wooden block against the bar code and makes a beeping noise. ‘That’s 20p,’ she says. |
| The child reacts to stories, music and rhythm, copying gestures and movements.  
(Physical Development – Point 1) | Sean hears a plane flying overhead and looks up to watch it. He puts out his arms and moves around, making engine noises. |
a technology, a culture and a mode of regulation that employs judgements, comparisons and displays as means of incentive, control, attrition and change – based on rewards and sanctions (both material and symbolic). (Ball 2003, 216)

As I discuss below, the EYFSP functions as a performative technology: it involves judgements of teachers and their ability to assess, and comparisons between schools and within them. Furthermore, there are sanctions if the right practices are not performed. This conceptualisation is important in the analysis of how this assessment relates to the production of inequalities because it emphasises the ways in which the EYFSP makes demands on teachers and schools; these cannot be considered separately from the practices the assessment engenders and the inequalities it produces.

The EYFSP and racial inequalities

Gillborn’s discussion of the (then) FSP and the significance of the shift from local assessment to a national system provides a starting point for my research into early years inequalities. In a 2005 article and his subsequent book, Gillborn argues that his work with Heidi Mirza using results from a number of local education authorities’ ‘baseline’ tests (informal assessments conducted when a child first enters school at the beginning of the Reception year) meant that it became ‘received wisdom’ that Black children often started school with equally high if not higher attainment than their peers. When the FSP was introduced, Black children were one of the lowest scoring groups, however. Gillborn argues that this shift, and the absence of any outcry over the reversal of previous patterns, reveals much about the racist workings of the education system. The case of the FSP, he comments ‘highlights a fundamental question … does assessment do more than merely record inequity, or does assessment produce inequity?’ (Gillborn 2006b, 335). This question reveals one of the fundamental issues for any researcher interested in issues of educational inequality: although assessment scores provide evidence of inequality which justifies a focus on these issues, these scores cannot be taken as factual information, given the complexity of the practices surrounding them. As discussed, research on tiered exam entry, for example, indicates that some groups of pupils are never given the opportunity to reach higher grades (Gillborn 2006b; Gillborn and Youdell 2000). In these cases, and the ‘educational triage’ described above, assessment produces inequality as much as it measures it. This paper focuses on how, at a classroom level, the practices associated with assessing children using the Early Years Foundation Stage Profile work to produce educational inequality among children as young as four and five.

Theoretical tools

As with Gillborn’s analysis of FSP scores (2006b; 2008), this paper employs some tools from Critical Race Theory to analyse the relationship between inequality and assessment. However, I also use concepts from critical policy sociology informed by a Foucauldian conception of discourse (Foucault 1980) and poststructural ideas about the intelligibility and recognisability of the subject (Butler 1990; 2004) to examine the discursive frameworks through which these Reception pupils and their ‘learning’ are constituted.
Critical Race Theory (CRT) is a wide and varied body of literature, originating in the United States, which focuses on the production and maintenance of racial inequalities in society; in recent years, it has been increasingly employed in relation to education (Gillborn 2006a; Ladson-Billings 2004; Tate 1997). One of the main tenets of CRT is that racism is endemic in society and inequalities are created unwittingly but systematically. Much of the literature influenced by CRT concentrates on the outcomes of practices, not the expressed intentions. In this way there are some similarities with the focus on ‘institutional racism’ in Britain in the early 2000s following the report into the racist murder of Stephen Lawrence; this term broadened out conceptions of racism to include ‘processes, attitudes and behaviour which amount to discrimination through unwitting prejudice, ignorance, thoughtlessness and racist stereotyping which disadvantage minority ethnic people’ (Macpherson 1999, 28). Similarly, my intention here is not to prove deliberately racist intentions; as Apple comments, ‘racial dynamics can operate in subtle and powerful ways even when not overtly in the minds of the actors involved’ (1999, 10). Instead, my focus is on how assessment systems may work in ways which serve to reproduce the very inequalities they appear to merely record. Furthermore, although Gillborn’s work on race and the FSP provides a starting point for this research, my focus is not limited to issues of racial equality; as indicated by the EYFSP results shown above, inequalities in the early years are multi-faceted. ‘Class’ effects, crudely indicated by free school meals data, are also significant in the EYFSP and are marked by gender, and the data demands that I consider these aspects of inequality together.

In my analysis of the data I also employ concepts from critical policy sociology, such as the idea of performativity outlined above, and a concern with the language of policy. Policies are powerful in that they ‘construct the inevitable and the necessary’ (Ball 2008, 5); embedded within the EYFSP is the idea that young children can and should be assessed, and that this information can be used to judge teachers, schools and pupils. I use a Foucauldian conception of discourse to discuss how particular ways of thinking and talking about the children in the study schools and their backgrounds has an impact on how they can be understood as pupils. I argue that discourses operate as ‘regimes of truth’ to constitute these children as successful or not (Foucault 1980). Following Butler’s discussion of intelligibility (1990; 2004) and Youdell’s application of these ideas to school settings (2006a; 2006b), I explore how pupils can be discursively constituted both as a group and individually as distant from idealised learner identities, and the impact of this constitution on assessment practices. This range of tools allows for a detailed analysis of the complex workings of these school settings within a wider context of educational inequality.

The study
The data used in this paper are taken from ethnographic studies of two Reception classrooms in primary schools in inner London during the academic year 2008–2009. The two schools, which I call Gatehouse and St Mary’s, are in the same local authority and both have school populations with high proportions of children from minority communities, with English as an additional language and in receipt of free school meals. During the course of the year, I interviewed two Reception class teachers – Jim at Gatehouse and Paul at St Mary’s (unusually both men) – on a regular
basis. At Gatehouse I also conducted one-off interviews with other teachers, including Lynn, the Early Years Foundation Stage coordinator. I also observed in Jim’s and Paul’s classrooms as a non-participant for one day each week through the year, a strategy informed by critical school ethnography (Youdell 2006a). Relevant documents, including the final EYFSP results for each class, were also collected. For ethical reasons of anonymity, information regarding the teachers, the children and the schools is kept deliberately ambiguous and pseudonyms are used throughout.

**Producing intelligible and acceptable results**

The main argument of this paper relates to the relationship between assessment practices, the publication of assessment results and how we understand educational inequality. Using these two Reception classrooms as case studies of teacher assessment practices in a context of established disparities in attainment, I examine how the production of results in a performative culture may also be the production of inequalities. Although the EYFSP had an impact on classroom practices all through the academic year, I concentrate in this paper on the production of the final results in the summer term. I use the term ‘production’ deliberately to emphasise how this is a process focused not on assessing children in detail, but on creating a set of scores that would be acceptable.

The teachers began the process by using their ‘teacher knowledge’ to award points to individual children, as advised by the EYFSP guidance (QCA 2008). This sometimes involved using the evidence from the children’s EYFSP folders, but frequently also included relying on their own ‘knowledge’, or as Jim (class teacher at Gatehouse) put it, making decisions ‘just from your own brain’. This process was complex in itself, given the teachers’ certainty in the veracity and objectivity of their ‘knowledge’, and the way in which children were constituted as different types of learner through discourses of class, race, gender, language and religion in the classroom. However, this system of awarding points on an individual basis was only the first stage of the production of EYFSP results.

The second stage in this production, and the main focus of this article, was an adjustment of results informed by the need to provide scores to the school management and the local authority (LA) that would be accepted as accurate. Because the EYFSP system relies on teacher assessment, a monitoring process known as ‘moderation’ is used by the LA to ‘check’ teachers’ judgements and decide if they are correct. Thus, the idea that teachers are able to accurately assess the children in their class based on their ‘knowledge’ is undermined by the external LA advisors operating as arbiters of accuracy. This has a distinct impact on how teachers assess, because it encourages them to produce results which will be acceptable to the LA. What is significant in terms of the maintenance of inequalities is that, in these schools, the only acceptable results were those below the national average.

**Responses to LA pressure**

The moderation process formally involved providing six EYFSP folders to the LA, who would then re-assess the children by talking to them and using the evidence provided. However, there were also more informal moderation processes, through which the LA advisor would visit the school and ‘look over’ and comment on the
Both levels of intervention caused the teachers great anxiety and frustration; at St Mary’s, Paul commented:

[With] the profiles you can make up any number and they’re not going to bloody know. They do come and moderate every two or three years, whatever it is, and you show them three samples’ … and they check that you’re making the correct assumptions. Well my experience is that most people just go and say right, what do we have to write to make them that level? They know the child is at a different level, but then they just go ahead and make up the points. (Paul, St Mary’s)

Paul suggests that the moderation process simply encourages teachers to fit children into ‘levels’ and give them points based on these; the implication is that however much teachers build up a ‘knowledge’ of the pupils, under the pressure of moderation they will simply produce the figures that are required. Lynn, the EYFS coordinator at Gatehouse, made similar comments:

I mean, look, you’re human, obviously if someone keeps saying to you, ‘Oh that’s not right, no, that’s not right, that’s not right, I don’t want you to mark them into a curve, I don’t want you to fit them into a pigeon hole, I want you to mark them how you think you mark them’ and then you get told off for that, then you think – and you do get to the point where you think I’m just going to fill this out – what do you think I should write, so that you will stop bothering me please. (Lynn, Gatehouse)

Here, Lynn describes the LA’s advice as contradictory and critical, asking the teachers not to mark into a curve and then telling them they are wrong when they don’t. Her frustration with the EYFSP system and the LA is evident, and this frustration leads, she suggests, to teachers filling in the EYFSP in ways that will discourage the LA from criticising them. This shows the regulatory function of the LA, who are able to define what is accurate in ways which result in Lynn feeling that she just wants them to stop ‘bothering’ her. Also at Gatehouse, the class teacher Jim commented ‘they want us to fit into a bell curve, that’s all it is. You know, a few low, a few middle, and a few high’, suggesting he felt a need to produce results in a normal distribution in order to make them acceptable to the LA. These comments suggest these teachers feel the ‘terrors of performativity’, the pressure to produce the right performance in an accountability culture.

However, although the teachers at Gatehouse talked about producing a normal distribution, there was a further ambiguous element to producing acceptable results which caused them great frustration:

I was moderated, which was a total farce. They told me [each year] that I’d been marking too high, then too low and then too high, and that basically that I need to make sure I mark the children a certain way. And I was like well, you know, I can’t mark them a certain way, they come out with what they get. (Jim, Gatehouse)

In Jim’s explanation (like Lynn’s above), there is a specific way in which he is supposed to score the children, which he refuses to engage in, instead arguing that ‘true’ results are produced. This tension between ‘true’ results and acceptable results, he went on to argue, relates to the context of the school. In this quote, he refers to the school being part of an Education Action Zone (EAZ); these were areas designated by a New Labour policy as needing additional funding due to poor education results.
Jim: [Our scores] were ‘too high’, they were way above the national average. And because the school is in an EEZ, or education action...
AB: An EAZ?
Jim: Yes, we shouldn’t be that high, because we’re achieving really high so by the time they’ve got to Year 6 they’ve gone down, and what are we doing wrong?
AB: And that doesn’t look good?
Jim: No, exactly. So the second year we were told to make sure, that between 4–5 points is a good score, so we did that.
(Interview with Jim, Gatehouse)

Jim explains that he was advised that 4–5 points per scale (out of nine) was a ‘good score’, in order to ensure that his results were not ‘too high’. There are two connected reasons why these results are not acceptable: because the school was previously in an EAZ, and because high EYFSP results will damage the school if and when ‘value added’ measures are used. In the following sections, I examine these two pressures on the production of EYFSP results.

The ‘difficult intake’

Education Action Zones were designated in the early 2000s, and were usually areas with low income levels and high proportions of social housing. In comments from Jim and Lynn, this term seemed to serve as a synonym for the problems of the ‘inner city’: when I asked Jim if he thought he would be told his EYFSP results were ‘too high’ again, he said:

I’m sure it’ll be exactly the same. It’s all because you know, we’re in, shouldn’t really, we’re in an EAZ, underprivileged children – [cynically] there should be no chance of them getting nines. (Jim, Gatehouse)

Lynn made a similar connection:

Another thing that was brought up last year was that, you know, our scores weren’t appropriate [...] This is the issue we have as a school because of it being a deprived neighbourhood and this is the kind of intake we have. (Lynn, Gatehouse)

This discourse of the ‘underprivileged’ children and EAZs worked powerfully at Gatehouse to distance the children from positions of educational success. Although individual children might get some high scores, there was a general assumption that these children would ‘underachieve’. A number of issues relating to race, class, parenting and housing were cited as part of a discourse of a ‘difficult intake’, through which these ‘inner city’ children were contrasted with ‘normal’ children elsewhere. Lynn commented ‘if you measure against the country we’re just always going to look hopeless’, and although Jim was critical of this assumption in the quote above, he also engaged with this discourse when discussing the children: he contrasted his class with ‘an English speaking, White middle-class primary school in the middle of England’.

Across both schools, this ‘difficult intake’ discourse constituted these children as inevitably low attaining because of their parents, their local environment, their language and their socio-economic status:
These goals are fair enough if you’re an English-speaking middle-class child whose parents work with you at home, but our children are not like that – they’re few and far between whose parents work with them – who will engage with them, who know what it means to develop a child’s mind. (Lynn, Gatehouse)

A lot of children have obesity issues, and health issues, and with a lot of them it’s that they live in these tall flats, they don’t get to go out, they don’t get to do all the running around. Their parents a lot of the time come from very hot countries and get very concerned about the cold, they don’t like to let their children go out. (Lynn, Gatehouse)

With the social backgrounds they have, they don’t see a lot of books, so they’ve got really into them here. (Paul, St Mary’s)

You’ve got all the issues to deal with – you’ve got English as an additional language, you’ve got sort of fear, and all those sorts of social things. (Paul, St Mary’s)

Here, Lynn and Paul’s comments reflect the tendency towards the ‘re-socialization’ of working-class parents (Gerwirtz 2001) in New Labour policy, where parents are encouraged to take on middle-class attitudes to schooling. In this policy context:

‘Good’ parents are viewed as those who, as well as choosing ‘good’ schools, instil confidence in their children, read to them, encourage them to talk and show an interest in what is around them, take them on educational outings and take an interest in their school work (Gerwirtz 2001, 369)

The urban location of the schools was another important part of the ‘difficult intake’ discourse, as shown in Lynn’s comments about ‘tall flats’. The association of the ‘inner city’ with social problems is connected to a policy context where ‘urban’ communities are repeatedly constituted negatively (Gulson 2008; Leonardo and Hunter 2009). As Lupton and Tunstall explain, through the current policy framework in the UK, ‘disadvantaged neighbourhoods … are discursively repositioned as irredeemably problematic’. (2008, 114)

The perceived lack of ‘good’ parents and the urban location of the school (and the racial and class associations of both of these) combined to render anything other than low attainment among the children at Gatehouse and St Mary’s completely unintelligible. Like the Australian school in Youdell’s (2004) study, these schools and communities themselves are constituted as ‘hopeless cases’. As we see in the data below, this had a real impact on how results are produced. However, when I asked Jim if he thought the idea of his results being ‘too high’ indicated low expectations, he said: ‘Apparently really low aspirations and expectations, and I don’t think that’s right’. Although Jim was more positive about the children in his class than Lynn, and he saw low expectations as problematic, he nonetheless did not appear to recognise the role of the teachers in reinforcing this connection.

The influence of ‘value added’

The second influence on producing results that Jim mentioned was the issue of ‘value added’ (VA) scores. These scores are included in league tables for primary and secondary schools and are calculated by comparing results from two tests in a child’s school career to measure ‘progress’. Primary VA scores are based on the number of children making the expected level of progress between the Key Stage 1 tests at age seven and the Key Stage 2 tests at age 11. Although the EYFSP scores are not used in any official measure of the ‘value’ that the school has ‘added’ to the
children, the principle of measuring progress over several years appears to be spreading into early years education. It is also possible that schools do their own value-added analysis using EYFSP results as a ‘baseline’ figure to predict Key Stage 2 test results. At both schools the teachers mentioned the need to keep EYFSP results lower in order not to make life difficult in later years, although Jim was more explicit about this. He criticised his school’s concern with VA measures, saying: ‘this is totally different from the rest of the school, this is a different curriculum, it’s not the same so I don’t know why they’re using the figures’. At St Mary’s, Paul commented that using these scores was ‘completely wrong’, ‘cause what they’re assessing there is not national curriculum level-type predictable’. He also showed his awareness of VA scores when he commented on the problems caused by Reception teachers deliberately raising their EYFSP results because ‘they don’t want to stand out’:

But what’s stupid about that, is that although this is not used to predict national curriculum levels, they’re trying to […] so I think people who do it are trying to cover, trying to keep the wolf away from the door-type thing, but actually it means with value added if you’ve got a bunch of children who are really developmentally very behind, and you’re telling everybody that they’re normal – [correcting self] that their kind of expectation would be a normal expectation, you’re really setting up for massive failure, for the kids. So personally, I think it’s stupid. (Paul, St Mary’s)

Paul suggests that other teachers ‘keep the wolf away from the door’ by changing results but that the use of VA scores mean this is a flawed strategy. Paul’s argument that this practice means that ‘developmentally very behind’ children will be assessed as ‘normal’ implies that this adjustment will inevitably involve assessing children as higher than is accurate: he seems to suggest that teachers will falsely raise scores, not just make them more consistent. This stands in contrast to the kind of practices Jim describes at Gatehouse, where the focus seemed to be more on producing results which are not ‘too high’. Paul says that a side effect of artificially raising your EYFSP scores will be that when the results of subsequent tests are published, the children will do unexpectedly badly; it will be ‘a massive failure, for the kids’. This comment obscures the fact that value-added scores have little meaning (and are never published) at an individual level: they only have significance in terms of judging a school.

The issue of ‘value added’ was important in both schools as a factor which determined what results were acceptable and even desirable. One of the functions of value-added scores and particularly, *contextual* value-added scores (CVA) is to recognise the achievements of schools with lower overall attainment through a focus on ‘progress’. With contextual scores, this ‘progress’ is calculated within a statistical framework which expects slower progress for children from some minorities, children on FSM and from areas of greater deprivation, and boys (Bradbury 2011); as such, CVA is often presented as a mechanism which is ‘fairer’ to schools in deprived areas with high proportions of FSM and minoritised pupils. These schools’ positions in local league tables, importantly, are likely to be far higher when ranked by CVA than raw scores. Therefore, value-added scores are particularly important to schools like Gatehouse and St Mary’s; they provide a chance to demonstrate the quality of the school in the light of ‘contextual’ issues.

It would seem from the comments made about value added that the importance of these scores means that they function as a powerful but complex regulatory mechanism. The school management and LA find desirable results that are low
enough to show that children make progress as they move through the school; there is a real incentive to give low scores. Combined with the pressure to deflate results that are deemed too high for a deprived area, value-added scores work powerfully to ensure that EYFSP results at these schools remain below national averages. However, these scores remain one of several different ways of judging a school, just one disciplinary technology in the accountability system. There is always a danger for low attaining schools that they will not be judged on VA, and so lowering scores may be self-defeating. The ambiguity of the relative importance of different measures was shown in 2008 when the government used raw GCSE percentage scores to define 638 schools as ‘failing’, whatever their CVA scores, in an apparent reversal of previous policy which had prioritised progress and improvement over percentage figures (BBC News 2008). Thus the ‘fairer’ measure, presented as a gift to low attaining schools, can also be taken away without warning. In the study schools, the concern over value-added in the Early Years Foundation Stage, where it is not even officially used, suggests the power of this disciplinary technology to regulate results in all areas. Confusion allows the regulation to spread beyond its original remit; as Ball explains, ‘Constant doubts about which judgements may be in play at any one point mean that any and all comparisons and requirements to perform have to be attended to’ (2003, 220).

### Changing EYFSP results

Although these two influences on the production of EYFSP results – the risk of marking ‘too high’ for the ‘difficult intake’ and the need to mark low for the value-added scores – were discussed throughout the year at both schools, it was not until the summer term that I was able to observe how they impacted on the process of finalising the EYFSP results, and then only in detail at Gatehouse.

As Jim predicted, he was again subject to LA ‘intervention’, this time in relation to the results of the parallel Reception class, whose teacher (Liz) was new to the profession. He explained:

> So basically, [bit awkward] Liz had a couple of days off with a bit of illness, and hadn’t got her reports sorted. So [the head] gave me a day and got some cover in and I went off and looked through her reports [...] and had to change some of her results. And we had the early years specialist in ‘cause when I put them all into my own table as they were [...] [it] made the school half a percent above the kind of required mark of 78. And [guiltily] although I was not told to change any of the figures, the question arose as to how a school with a very difficult intake, could achieve half a percent, like average, how could it be average? And it also would then skew the Key Stage 1 Sats results which had come out just below average. Which would show, if you were to chart it on graph [cynically], would show that the children haven’t achieved, well have gone backwards from the end of Foundation to the end of Year 2. So therefore [laughing], stuff was changed. (Jim, Gatehouse)

Jim explains here how he was either told, or felt under obligation to lower some of the children’s marks because the overall average points score of the two classes was too close to the government’s benchmark of 78. As he described in relation to previous years, his marks were ‘too high’ for the kind of school that Gatehouse is; they are unintelligible within the discourses of race, class and deprivation that constitute the ‘inner city’. A further consideration was that the school’s low Key Stage 1 results (though obviously for a different group of children) would look even
worse if the EYFSP results were good, or even just average: this is the power of value added to function as a performative technology, often in illogical ways. Jim’s concern is not to boost his results to make the school look better, but to lower them in order to fit the LA’s perception of the school as ‘challenging’, and to make sure the school’s value added does not look like the children are going ‘backwards’ (even though this is nonsensical as it is a different group of children). We see here how these two elements – value-added scores and the difficult intake discourse – work together to ensure that the school has a vested interest in keeping marks low and has a ready-made justification. It is perhaps acceptable for the Reception children to do badly because they are still so influenced by their homes and backgrounds, so it can be argued that the school has not had a chance to ‘add value’ yet. This deflation of the EYFSP scores will only make the school look better in the long run, and it makes sense as the only intelligible set of results given the dominant discourses around poverty, race and the local area.

It is in the detail of how the results were changed, however, that the real impact of this process is apparent. Jim explained:

Basically, I had all of the children, I put them all in a list, from top to bottom, and I picked out the ones that looked in the wrong place. And then had a closer look. I noticed that Liz had marked quite a lot of children as achieving I think Point 7 in the third section of maths, which is about shape, space and measure, which is something to do with ‘uses mathematical language such as larger, greater, smaller’ and she ticked quite a lot of children as achieving that. But the children who had achieved that hadn’t achieved Point 6 and 5 and 4 and so on. Now this has been a big push for her this year, mathematical vocabulary […] But I think, and I’m quite confident in the fact that I took it away from the majority of her class – I don’t think I gave it to very many children in my class. They’re perfectly capable of being able to copy the words in lessons and being able to use the words in lessons, but they don’t use it independently, and because they don’t use it independently … well it’s a bit cheeky ‘cause I’d be upset if another teacher came and crossed off loads of my results, but … I would have left it personally but because of the conversation I’d had about the average marks, I had to kind of revisit… […] I knocked it off for most of the lower ability children.

The removal of marks, which appears logical and even sensible in Jim’s explanation, plays out in complex ways, as shown by the removal of marks from the ‘low ability’ children. This reveals the relative difference in value of the children in the class: these children, already assessed as ‘low’, are expendable within the system – it does not seem to matter so much if they lose a few marks. Jim does not take a point away from everyone, thereby keeping the same pattern of attainment but lowering overall scores; instead his tactic increases the spread of results in the class, distancing the ‘low ability’ even further from the ‘high ability’. This is, perhaps, informed by the need to maintain a normal distribution of scores: removing these marks will just produce a longer tail end to the ‘bell curve’ without flattening it out. But the effects are not just statistical; these children, already given low scores, are further disadvantaged. They will arrive in Year 1 with an EYFSP assessment which decrees that they cannot use mathematical language, despite their teacher observing that they can; they are victims of the external definition of what can be ‘correct’. Furthermore, this will not seem incongruous, given the strength of the ‘difficult intake’ discourse. This process of lowering ‘low ability’ children’s scores can be seen as the beginning of a process through which some children are
repeatedly deemed to be ‘underachieving’. These scores make a real difference because, as Gillborn explains in relation to Black pupils:

It is these scores that schools will use to judge the progress of the students in later assessments. Potentially, the lower attainments of Black students in subsequent stages of the education system will no longer be viewed as a relative drop in performance; they may simply be viewed as performing in line with their lower starting points. (Gillborn 2006b, 10)

Thus low EYFSP scores have a potential impact on a child’s entire school career; they may set these children up for many years of low expectations.

Jim also commented in relation to the removal of EYFSP points:

I don’t know what other schools are like and this is the issue. We don’t know what kind of other scores, other schools are like. We have about, we have a fair few nines, across the board.

This shows the wider pressure to produce intelligible results, not just for the LA, but in terms of comparisons with other schools. Jim knows that his school cannot intelligibly score higher than schools in more affluent areas, and so is worried that they have too many nines. Here, the teachers’ complex positions as professionals contribute to the production of results: while at times they feel trusted to get the right results by the ‘teacher knowledge’ discourse, this can be taken away in a moment by the LA. This reflects what Sachs (2001) terms ‘managerial professionalism’, where ‘teachers are placed in a long line of authority in terms of their accountability for reaching measurable outcomes that stretches through the principal, to the district/regional office, to the central office’ (p. 152). Jim fears that he will not play his part in this line of accountability adequately, that the other schools will make his results look incorrect; without access to this information, he is left guessing what will be seen as realistic. This mirrors Ball’s description of the uncertainty inherent in many accountability systems, which involve being ‘ontologically insecure’, ‘unsure about whether we are doing enough, doing the right thing, doing as much as others’ (Ball 2003, 220).

The situation at Gatehouse reveals the intricate balancing acts involved in transforming early years into an ‘auditable commodity’ (Shore and Wright, 1999 in Ball 2003), and the necessity of engaging in what Ball calls ‘fabrications’ (Ball 2003):

Fabrications conceal as much as they reveal. They are ways of measuring oneself within particular registers of meaning, within a particular economy of meaning in which only certain possibilities of being have value. However, such fabrications are deeply paradoxical [...] Fabrications are both resistance and capitulation. They are a betrayal even, a giving up of claims to authenticity and commitment, an investment in plasticity. (Ball 2003, 225, emphasis in original)

In submitting to the school management’s concerns to change the results, Jim is engaging in ‘both resistance and capitulation’. On the one hand, the deliberate changing of results is a resistance to the detailed, year-long build up of ‘knowledge’ prescribed by the EYFSP; it shows he can assess without collecting information into folders, and that he is prepared to cynically produce results which he understands to be inaccurate (though he attempts to deflect this) and therefore undermine the principles of teacher assessment. However, changing the results is also capitula-
tion; there is no doubt that he will produce the results, and that they will be as acceptable as he can manage given the ambiguity of what is ‘accurate’. Jim gives up what Ball calls ‘claims to authenticity and commitment’ by changing the results; his awkwardness suggests that this is not without its costs.

Jim’s decision to remove marks from some pupils reveals his investment in a fabrication, but by arguing this I do not mean to apportion individual blame; it is clear that Jim is severely limited in his ability to resist the pressure to produce a fabrication. Having been told in previous years that his assessments were ‘wrong’ because they were too high or low, Jim is familiar with the ways in which his assessments can be deemed incorrect on the basis of the ‘difficult intake’. His comments throughout the year suggested he was resigned to the need to ‘produce a figure’. Indeed, later in the same interview he commented that despite the changes, ‘they’ll probably say something like they’re too high’, suggesting he feels that whatever results he produces, he will be criticised. Within his professional context, Jim is unable to resist the need for a fabrication due to the pressures of the LA and school management; he is constrained by local and national discourses which define how well pupils can do so that the results remain intelligible. However, he is not willing to admit to this capitulation: he claimed that ‘I didn’t take away stuff that wasn’t fair to take away’. This can be seen as an attempt to justify a process which he knows to be questionable, but it is also, perhaps, an attempt to deny the extent to which he is constrained by external pressures when producing the EYFSP results. Nonetheless, his willingness to tell me about changing the results at the beginning of our final interview, when he could have simply avoided this issue, suggests he has serious misgivings about the practice, however he presents it later. Any discussion of these assessment practices must be informed by an awareness of the contradictory effects of the EYFSP on the teachers and their professionalism. While a ‘teacher knowledge’ discourse legitimises their positions as early years experts, the EYFSP simultaneously devalues the teachers’ status by requiring them to produce evidence and designating them ‘right’ or ‘wrong’ in their assessments. What the EYFSP seems to give, it also constantly threatens to take away; it gives them the power to assess, but not the power to decide if these assessments are accurate. They need to engage in the performance of assessment, because that is an important part of their teacher professionalism, but it is ambiguous exactly what performance is required. Teachers’ main concern is to fulfil their roles as deliverers of intelligible results, but they cannot be sure what this will mean. And all the time, their investment in the fabrication is immense: it takes up significant proportions of their time, and it is a judgement on a year’s work, their professionalism and their pedagogy.

The example of Gatehouse shows how the need to produce acceptable and intelligible results is a powerful determinant of the scores that are allocated; however, it also shows how performative technologies require constant work and maintenance. The ‘enacted fantasy’ (Butler 1990) does not happen in one moment, but must be built up and performed throughout the year, in the production of folders of evidence and through continual observation. The fictive accuracy of the EYFSP must be resignified though the regular collection of data and the allocation of points, up until the final scores are awarded, even if they are frequently changed in response to LA demands. Even then, success (in the form of getting intelligible results) is only ever fleeting – it will need to be reproduced, refabricated again next year, and the next. Furthermore, value-added measures and long-term analysis of results demand coherence between fabrications; there can be no respite.
Producing inequality

The data discussed here suggest that the EYFSP plays a part in producing inequality in the early years. With a few individual exceptions, results are kept deliberately low for the children in these schools, the vast majority of whom are on free school meals and/or from minority communities, and a new ‘truth’ is created – that significant proportions of these children are already ‘behind’ expectations after just one year of school. In turn, these results are taken as evidence of inequalities both in school and in society more generally. In order to explore what the case of the EYFSP in these classrooms means in terms of how assessments contribute to systematic disadvantage in the education system, I return to Gillborn’s initial analysis of the national FSP results.

In his 2006 article Gillborn used a CRT-style counterstory to examine the key issues relating to the then FSP. In this story, he described an imaginary ‘deeply racist society’ in which a despised group succeed on a test, and so therefore ‘The test must be redesigned so that the despised group no longer succeed’. He concludes the story:

But, of course, such a crass and obviously racist set of events could never occur in the real world. There would be an outcry. Wouldn’t there? (Gillborn 2006b, 326)

Gillborn does not use this story to suggest that the EYFSP was deliberately introduced in order to make children from minoritised groups do badly; instead, he offers this analogy to show that the results of this policy are the same as if they had been intentionally racist. In the same manner, I offer here a story which focuses on the processes I have observed at St Mary’s and especially at Gatehouse; this story takes as its premise the same racist society:

A new test is introduced; the despised group do badly, and the status quo is preserved. But, the new test is for very young children, and so has to be based entirely on teachers’ judgements and observations. There is a risk that the despised group might start to get better scores. How to control the teachers, and ensure the scores are kept down? Two processes ensure that the despised group continue to get low scores. Firstly, the idea that this group will and should do badly must become commonsense in the school system. This can be achieved by creating policies which single these groups out as problematic, such as designating areas in which they live as special zones, to emphasise their low attainment. Because the entire test is based on what the teachers ‘know’ about the children, any lower expectations of this group will not be too obvious; the teachers can give lower scores to the despised group, without it ever seeming unfair.

A second back-up process ensures that if the teachers do start to give children from the despised group high scores, this can be monitored and prevented. Each teacher’s scores must be ‘checked’ by the local authority ‘expert’ advisors, who have the power to deem teachers’ scores right or wrong. They are feared by the teachers, because the teachers know they are judged on these scores. They give contradictory advice, and so the teachers are always unsure about quite what the advisors want; some become so confused with the system that they just want to produce what the advisors want to hear. When the advisors tell the teachers that their scores are too high, the teachers respond. So, when results from a school with a large number of children from the despised group are above the national average, the advisor can simply declare them ‘too high’ and the school will change them. In turn, this is helped by the official designation of schools like this into the special zones; being in this zone provides further
proof that the teacher must have simply got the scores wrong. Perhaps soon, the advisors will not even have to say the scores are wrong because the head teachers will begin to understand the system, and make sure that their teachers score low to avoid the pressure from the advisors.

Through these two processes, the group gets low scores. The results are published each year, and this backs up the idea that it is only commonsense that the despised group do badly, and no one is to blame. Everyone is quietly pleased with the wonderfully circular, self-perpetuating way in which the group stays lower than the other groups in this test, especially since it provides the benchmark for all of the children’s further progress. As long as the advisors keep checking, the teachers keep feeling that they need to get it ‘right’, and the idea that the group will inevitably do badly keeps circulating, the despised group will get poor scores and keep getting poor scores.

It is important to reiterate here that I do not believe that the EYFSP is a deliberate, planned strategy to ensure that minoritised groups or children from disadvantaged families do badly in schools. However, as with Gillborn’s story, the results are the same: the data discussed here suggest that the teachers expect their schools to do badly because of their ‘difficult intake’, and the EYFSP scores are changed if they are too high for a school in an area with low educational attainment. Indeed, the children’s final scores in the classes at Gatehouse and St Mary’s were lower than the national average: the results are the same as if it had been deliberate, and this is deeply concerning. Gillborn writes:

But there is no evidence of conscious intent: there is no conspiracy. It is more frightening than that. Rather than being generated by a deliberate strategy (one that is readily open to exposure and reversal) these changes appear to have resulted from the normal workings of the education system – a system that places race equality at the very margins of debate and takes no action when black students are judged to be failing. (Gillborn 2006b, 334)

The idea that the ‘normal workings’ of the system result in minority groups being disadvantaged is a familiar concept to CRT scholars who take racism as an endemic aspect of society, but it is a shocking idea to many educators. I am arguing here that it is the ‘normal workings’ of the EYFSP that can work to disadvantage some children (both from minority groups and other ‘inner city’ identities), without anyone ever consciously intending to do so. My intention is not to blame the teachers for their situation either; as I have argued, they are tightly constrained by the discourses surrounding them and their professional context.

In contrast to Gillborn’s comment above, I would argue that it is the centrality of issues of race and class within educational debates, not a position at the margins, which allows the normal workings of the system to produce classed and raced outcomes. The repeated connections made between ‘disadvantaged’ pupils and low attainment, most recently in discussions of the ‘pupil premium’, make these disparities inevitable; all other results are not intelligible. I am not suggesting that these issues should not be discussed (that would be hypocritical in the extreme), rather that it is important to recognise that concerns, particularly in the media, over the lower attainment of children from minorities or on FSM also have an impact on expectations, and on the discourses that operate in schools.
This exploration of the EYFSP in Reception tells a familiar story about the negotiation of policy into practice and its unintended effects (Ball 1993). However, other policies are implicated too in this case, particularly value-added measures and the creation of Education Action Zones, with similar unintended effects. Value-added scores provide a real incentive to keep results low; it is perhaps surprising that any school gives their children high scores on the EYFSP, considering the high expectations that would be set up for the next six years. Moreover, the incentive to give low scores is greater for schools that are unlikely to do well on raw results because value added is their chance to show their success. Unintentionally, this measure may work to lower the EYFSP results of schools that are already seen as low attaining; it risks perpetuating educational ‘failure’ for these children in particular. Similarly, the designation of areas as Education Action Zones, although intended to alleviate educational inequalities, may perpetuate this ‘underachievement’ by making it the only intelligible option. There is clear link in the Gatehouse teachers’ explanations between this specific label and low results. EAZs ended in 2005, and yet in 2009 Jim was still talking about the area as an EAZ; we see how the negative effects of the label linger on even after the policy and the money have disappeared. The idea that policy which is notionally redistributive and should contribute to social justice could also have the effect of making sure that results remain low is worrying; if, as I have argued in relation to Reception, other teacher assessments are informed by the need for results to be intelligible within the context of the school, then this has implications for many schools which are involved in similar schemes or other policies which aim to reduce inequality.

In this paper, I have argued that the EYFSP assessment system in early years education has a role in producing inequality, and that results cannot be taken merely as records of inequality. The relationship between assessment and inequality is complex and changing with the introduction of different forms of assessment and different monitoring systems. The increase in teacher assessments as replacements for tests in both primary and secondary education in recent years makes these findings relevant beyond early years education: what if other teacher assessments are only recognised as accurate if they mirror the patterns of previous test results? Gaps in attainment and inequalities will be the only intelligible results and will be reproduced each year, while the possibility for change is completely removed.

This issue is important in a policy context where the value of different forms of assessment is contested: the problems of formal statutory testing are well-documented (Gillborn and Youdell 2000; Stobart 2008) and there have been major campaigns by teaching unions to abolish statutory tests (NUT and ATL 2010). Within these campaigns, teacher assessment is posited as the ‘responsible alternative’ (NUT and ATL 2010). However, we should be wary of assumptions that the removal of tests will return us to an idealised past of a pre-testing regime where there were no inequalities. Any ‘return’ to teacher assessment, if it can even be characterised as such, would be affected by the practices and discourses established during the intervening period of statutory testing. The data I have discussed from early years classrooms demonstrates that the use of teacher assessment can have a powerful role in producing inequality; the potential impact of teacher assessment should be an issue of great concern for those working to eliminate inequalities in the education system.
Acknowledgements
This research was funded by the Economic and Social Research Council (ES/G018987/1). I would also like to thank David Gillborn, Deborah Youdell and the reviewers for their helpful comments on this article.

Notes
1. The term ‘early years’ commonly refers to the Nursery and Reception classes in a primary school, formally now known as the ‘Early Years Foundation Stage’.
2. Different assessment systems operate in different parts of the United Kingdom.
3. This benchmark is reached if a pupil gains a score of at least six (out of nine) in all seven of the EYFSP scales for Personal, Social and Emotional Development and Communication, Language and Literacy.
4. Although differences in attainment by gender are significant at age five and this is an important issue in the sector, I concentrate mainly in this article on issues of class and race inequality.
5. League tables for EYFS results have been proposed by the Conservative-Liberal Democrat government, however. The government has also announced a review of the entire EYFS framework.
6. At Gatehouse, this observation also involved children from the parallel Reception class because the two classes often mixed. The other Reception teacher, Liz, was a new teacher and did not participate in the project.
7. Paul appears to be referring to the three levels (high, middle and low) of the sample folders he has to provide; usually a boy’s folder and girl’s folder are submitted for each level.
8. This phrase is usually used in terms of escaping from the dire consequences of poverty – for instance, one might take an unpleasant job in order to ‘keep the wolf from the door’, to avoid starvation or eviction. Paul appears to use it to mean avoiding the attentions of the LA advisors.
9. At St Mary’s, I was not able to observe how Paul finalised his EYFSP results.

Notes on contributor
Alice Bradbury is a senior lecturer in the Department of Education at Roehampton University, London.

References


