

"They are children… not robots, not machines"

The Introduction of Reception Baseline Assessment

Alice Bradbury and Guy Roberts-Holmes, UCL Institute of Education

NUT
ATL the education union
UCL Institute of Education

Contents

Foreword 3
Research team and acknowledgements 4
Executive summary 5
Glossary and abbreviations 6
1. Introduction 7
2. The impact on children starting school 13
3. Identification of particular children and groups 19
4. Relationship to existing assessments 23
5. Perceptions of the accuracy and validity of the data 26
6. The impact on teachers' roles and professionalism 31
7. Parents and Baseline Assessment 37
8. Selection, cost and the role of private companies 42
9. The wider policy context 47
10. Conclusion 52
References 55
Appendix A. Information on Baseline providers' assessments 56
Appendix B. Demographic data on survey respondents 58
Appendix C. Summary graph of respondents' opinions 59

www.teachers.org.uk/baseline

www.atl.org.uk/policy-and-campaigns/campaigns/baseline-assessment.asp

Foreword



Mary Bousted
General Secretary ATL



Christine Blower
General Secretary NUT

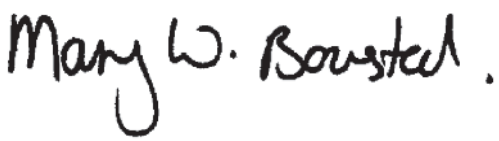
Baseline Assessment was not introduced, as many think, to support a child's transition into school or to help teachers plan for that child's learning; it is better understood as a measure that would supposedly enable school performance to be accurately evaluated.

Yet again the weight of school accountability has been laid on the shoulders of our children. As these children are so young – just four – we have all the more reason to find out what impact the policy has had on these crucial early days in school.

The start of school is a precious and important time for our children. Get it right and the seeds are sown for a love of learning that can carry on through school and into adulthood. Get it wrong and the scourge of low self-confidence and disaffection for school can set in, blighting a child's life chances.

Responding to the disquiet of our members, the ATL and the NUT have worked closely together to test and critique Baseline Assessment. We jointly commissioned this research in order to examine and investigate teachers' experiences and views of the new form of assessment, as well as the claims that the Government has made about how it could be used. The research findings are clear: teachers have major concerns about the negative consequences of Baseline Assessment on children, teachers and schools; we now call on the Government to rethink this misguided accountability measure.

Only 7.7 per cent of the teachers and primary school leaders surveyed for this research agree that the Baseline Assessment is a fair and accurate way to assess children. This statistic alone proves that this policy fails in its stated purpose – to assess the child's starting point at school so that "value added" can be calculated at the end of primary school in year six. Baseline Assessment has critical flaws and should be abandoned; the Government must acknowledge that this policy has failed.





Research team

Dr Alice Bradbury (Project leader)

The research team based at UCL Institute of Education, University College London was led by Dr Alice Bradbury. Dr Bradbury has a background in primary education and specialises in research in early years and primary schools, with a particular focus on issues of assessment and the impact on classroom practices. A previous award-winning research project, funded by the Economic and Social Research Council, focused on the impact of the Early Years Foundation Stage (EYFS) Profile on teachers and pupils, and considered specifically the issue of inequalities in early years education related to assessment. She has also written on the introduction of the Phonics Screening Check (2014) and other policy related topics.

Dr Guy Roberts-Holmes (Co-author)

Guy Roberts-Holmes is an experienced early years researcher and senior lecturer. He leads the internationally successful MA Early Years Education at the UCL Institute of Education. Dr Roberts-Holmes co-led a funded evaluation of the Early Years Professional Status (EYPS) which informed an early years professionalisation national strategy group (2009-2011) and was a researcher for the Department for Education (DfE) funded 'Practitioners' Experiences of the Early Years Foundation Stage (Sept 2010-March 2011) which subsequently informed the revised EYFS (2012).

Emma Jones

Emma is a PhD candidate and recipient of an Excellence Studentship at the University of East London. She started her career as a secondary geography teacher in Birmingham and has worked as a head of department in a number of international schools. Emma has an MA in Education, Gender and International Development and an MRes in Education and Social Science Research Methods from the UCL Institute of Education. She has also worked as a research officer on projects for the Department for International Development (DfID), the British Council and Unicef.

Acknowledgements

The research team would like to thank the many respondents to the survey and particularly the staff and parents at the case-study schools for giving up their time for the project. We would also like to acknowledge the advice provided by a number of colleagues at the UCL Institute of Education, including Annette Braun, Peter Moss, Charlie Owen and Carol Vincent.

Note: This research was commissioned by the Association of Teachers and Lecturers (ATL) and the National Union of Teachers (NUT). However, the analysis presented here is the authors' and does not necessarily reflect the views of the ATL and NUT.

Executive summary

1. Teachers and school leaders have serious doubts as to the accuracy of the Baseline Assessment and its use in measuring progress, in relation to all three baseline providers. Only 7.7 per cent of respondents to the survey agreed the data was an 'accurate and fair way to assess children' because of the wide range of variables in the assessment process.
2. Many teachers and school leaders doubt the use of measuring progress from Reception to Year 6 given the problems of assessing accurately at age four and the variability of children's patterns of progress and development. Only 6.7 per cent of survey respondents agreed it was 'a good way to assess how primary schools perform'.
3. The majority of schools already used informal on-entry assessments to plan teaching and to identify children with particular needs; the Baseline Assessment is not seen as an improvement on these methods.
4. Most schools selected the Early Excellence Baseline due to the similarity to the existing EYFS and the promotion of this scheme as 'early-years friendly'; some felt under pressure to do so from their local authorities or because other providers were removed.
5. Baseline Assessment has an impact on teaching and learning, including encouraging the practice of 'stopping teaching', and was not seen as helping teachers get to know pupils better.
6. Baseline Assessment has little use in terms of the identification of additional needs.
7. There is a significant effect on teachers' work-loads particularly where schools are continuing to use their old on-entry assessment system which fits better with data tracking systems or with the existing EYFS Profile. Survey data suggests this work-load effect occurs with all of the three providers.
8. Few schools plan to provide information on Baseline Assessment scores to parents, due to their concerns. A limited sample of parents revealed that most are unaware of the assessment; some are content, assuming that the assessment will help teachers get to know children better but others have concerns about accuracy and the age of the children.
9. School leaders are uncomfortable with the use of private providers and the related marketing that they have received, and feel vulnerable at a time when there has been a great deal of policy change related to assessment.

Glossary and abbreviations

ATL	Association of Teachers and Lecturers
BME	Black and Minority Ethnic
CEM	Centre for Evaluation and Monitoring, a DfE-approved Baseline Assessment provider
Development Matters	Non-statutory guidance to support the implementation of the Statutory Framework for the Early Years Foundation Stage
EAL	English as an Additional Language
Early Excellence	A DfE-approved Baseline Assessment provider
ELG	Early Learning Goals
EExBa	Early Excellence's Baseline Assessment
EYFS	Early Years Foundation Stage (children aged 3-5)
EYFS Profile	Early Years Foundation Stage Profile, the statutory assessment at the end of Reception
KS1	Key Stage 1: includes Year 1 and 2 (children aged 5-7)
KS2	Key Stage 2: includes Year 3 to 6 (children aged 7-11)
LA (or LEA)	Local Authorities (previously Local Education Authorities)
Leuven Statements/Scale	A five-point scale to assess 'well-being' and 'involvement' in the early years
NFER	National Foundation for Educational Research, a DfE-approved Baseline Assessment provider
Nursery	Class for children aged 3-4, may be based in a primary school
NUT	National Union of Teachers
Ofsted	Office for Standards in Education
Reception	Class for children aged 4-5; first compulsory year of school
SATs	Standard Assessment Tests at end of each Key Stage
SEN/SEND	Special Educational Needs/Special Educational Needs and Disabilities
SIMS	School Information Management System
Summer-born	Children born between April 1st and August 31st
TACTYC	Association of Professional Development in Early Years
Target Tracker	Assessment package to support the monitoring of pupil progress
(W)	Written survey response

Section 1
Introduction

Baseline Assessment was introduced in primary schools in England in September 2015. The assessment is conducted during the first few weeks of the autumn term with children entering the Reception year, aged 4-5, and is designed to produce a 'baseline' figure on the basis of which their progress during the primary years can be measured. The introduction of this assessment followed a consultation on how accountability measures for primary schools could be reformed. According to a DfE report, from 2016 the Baseline Assessment will be 'the only measure used to assess the progress of children from entry (at age 4-5) to the end of key stage 2 (age 10-11), alongside an attainment floor standard of 85 per cent' (DfE 2015c). Although other measures could still be used to demonstrate progress to Key Stage 2 for children starting reception in the school year 2015/16, many schools have been encouraged to use the new assessment. From September 2016 it will still be technically voluntary but there will be even stronger pressure to 'opt in' as the DfE will no longer recognise other measures as a way for schools to demonstrate 'value added'. At the same time, from September 2016 the existing statutory assessment for Reception children, the Early Years Foundation Stage Profile (EYFSP) will be made optional (DfE, 2015a).

Baseline Assessment marks a significant change in the use of summative assessment in early years education in primary schools, as the current statutory assessment (EYFSP) is conducted at the end of the Reception year and is based on thorough observations of children over the school year. The EYFSP is not formally used in value added measures of children's progress and thus school performance. Baseline Assessment data will form a key part of how schools are assessed in the future. Baseline Assessment will result in a single score for each child; when they reach Year 6, each child in the cohort will be measured against their Baseline Assessment score in order to judge the progress they have made while attending primary school. This is a major shift in approaches to accountability in primary education which involves the early years phase more than ever before.

There are also implications for the allocation of funding, as the Baseline Assessment data will replace the EYFSP data as the basis for the allocation of low prior attainment funding to primary and infant schools from 2016. The DfE states that this funding 'helps schools support pupils whose attainment was below the expected level before reception year' (DfE, 2015a). This funding re-arrangement adds further pressure upon schools to 'opt in' to Baseline Assessment.

Also unprecedented in primary education is the use of private providers to produce the actual assessments. Contracts to provide a Baseline Assessment were put out to tender by the DfE in 2014. Six companies were selected and formed the 'approved list', from which schools were required to select an assessment. In 2015 schools paid these providers directly and were reimbursed by the DfE for 'the basic cost of the reception baselines' (DfE, 2015e). After schools had selected their Baseline providers, the least popular providers were removed from the list, leaving only the following:

- Centre for Evaluation and Monitoring, Durham University (CEM)
- Early Excellence
- National Foundation for Educational Research (NFER)

Further information on these organisations and their Baseline Assessments is provided in Appendix A. Schools that had selected other Baseline providers either had to change, or were not reimbursed for the costs. The use of private providers is unusual in that other primary school assessments, such as Key Stage 2 SATs and the Phonics Screening Check are provided by the Standards and Testing Agency (STA), a government agency. The choice of providers also means that data cannot be directly compared between schools using different assessments.

Over 70 per cent of primary schools selected the Early Excellence Baseline (known as EExBa) (TES, 2015a), which is based on observations, like the existing EYFSP. Reflecting this high proportion of schools using the Early Excellence provider, this report's five geographically-sampled case-study schools all used Early Excellence Baseline (EExBa) whilst the survey data was collected from schools using all three providers. It is important to note that the report is not intended as a critique of Early Excellence Baseline (EExBA) and many of the generic critiques from the case-study schools regarding the baseline process could be applied to CEM and NFER providers. Schools' choice of Early Excellence was reported in the press as a rejection of the 'testing' of four year-old-children (Adams, 2015).

Baseline Assessment has proved to be a highly controversial move in terms of assessment in early years, as the assessment is conducted in children's first weeks of formal schooling while they are settling into Reception and while most children are aged four. There have been concerted campaigns to prevent the introduction of Baseline Assessment from teachers' unions, educational organisations and in the press (for example, ATL, 2015; NUT, 2015; Reclaiming Schools, 2015; TACTYC, 2015). As well as the age of the children involved, there has been concern over the use of a single numerical score, the issue of children moving school and the impact on nursery education (age 3-4).

Research conducted for the DfE by the National Federation for Education Research (NFER) in July 2015, which was intended to 'inform the approach to implementation of the reception baseline and to identify effective ways of communicating the results to parents' (DfE, 2015c:3), raised particular concerns, which have informed in part our approach in this report. Through a small-scale survey, interviews with teachers and focus groups with parents, the NFER report found:

- Lack of understanding of the assessment;
- Existing on-entry assessments were used widely, often through observation;
- 'Some evidence' of gaming results in order to maximise progress measures;
- A desire to report results to parents orally.

The policy context for Baseline Assessment

There are antecedents to the new Baseline Assessment, as previous similarly-timed assessments have been used in the past (1997-2002) and many Reception classes already carry out their own form of 'baseline' using teacher observations and comparisons with the EYFS early learning goals. However, as stated earlier, the introduction of Baseline Assessment is unprecedented in that it is the first use of a Baseline Assessment for accountability purposes.

The origins of Baseline Assessment can be traced to comments from Ofsted's Chief Inspector, who said in 2013 that the EYFS Profile was 'too broad an assessment' that did not link effectively to subsequent Key Stage assessments, and was 'a weak basis for accountability' (Gaunt, 2013). One month later these comments were included in a DfE Press Release, paving the way for a consultation on reforming primary school assessment and accountability. This consultation included the proposal for a Reception baseline, alongside other alternatives, based on the principle that 'measures of progress should be given at least as much weight as attainment' (DfE, 2013: 4), and the government received over 1,000 responses. The reaction to the idea of a baseline was largely negative, with 51 per cent of respondents answering 'no' to the question 'Should we introduce a baseline check at the start of reception?' (34 per cent Yes, 16% Not sure), on the basis of concerns that it was too young to test children, there would be a negative impact on the start of school, and it would undermine the existing EYFS Profile (DfE, 2014: 16). When asked 'Should we allow schools to choose from a range of commercially available assessments?', 73 per cent of respondents said 'no'.

The Coalition Government's response to the consultation was the Reforming Assessment and Accountability for Primary Schools document (DfE, 2014) which clearly stated that the primary purpose of the baseline was accountability and not assessment:

> The purpose of the reception baseline is for an accountability measure of the relative progress of a cohort of children through primary school. (DfE 2014:1)

> We will use a reception baseline as the starting point from which to measure a school's progress. (DfE, 2014: 7)

This document also made clear that the result of the baseline must be a single score:

> The purpose of the reception baseline is to support the accountability framework and help assess school effectiveness by providing a score for each child at the start of reception which reflects their attainment against a pre-determined content domain. (DfE 2014:1)

In terms of content, the DfE specified that:

> The clear majority of the content domain must […] demonstrate a clear progression towards the key stage 1 national curriculum in English and mathematics. (DfE, 2014: 1)

This contrasts with the existing EYFS Profile, which covers all areas of the curriculum and particular 'learning dispositions'.

The justification for Baseline Assessment, in common with other uses of progress measures such as 'value added', was 'to make sure we take account of: schools with challenging intakes [and] the important work in work in reception and Key Stage 1' (DfE, 2015a). Thus Baseline Assessment is presented as sympathetic to schools with lower attainment on entry and as recognising the value of early years education. The value and accuracy of predicting results from Reception to Key Stage 1 or 2 has been a concern for campaigns against Baseline Assessment (Reclaiming Schools, 2015).

This emphasis on progress is the epitome of a trend apparent in recent years in primary education, where concerns over the 'value added' by schools are paramount (Bradbury and Roberts-Holmes, 2014). Extensive research on the use of assessments for accountability has shown the potential impact on pedagogy and curriculum, classroom organisation, teachers' work-loads and feelings of professionalism, and most significantly, on the issue of manipulation of results. Previous work on the EYFS Profile found that there was pressure to deflate results to maximise later progress measures (Bradbury, 2013), even though this was not a formal function of the assessment. Note that Early Excellence trialled a Nursery Baseline in 2015 in an attempt to prevent such deflation occurring at Reception Baseline.

As reflected in the various campaigns against Baseline Assessment, the use of a 'high stakes' assessment at age four/five also contrasts with considerable socio-cultural research which has demonstrated that children learn through sets of social relationships (Broadhead, 2006; Fleer, 2010). This body of work argues that authentic, holistic and developmentally appropriate assessment, based upon teachers' observations over time in a range of contexts, makes visible what young children are capable of learning in supportive and collaborative relationships. It is argued that a particularly useful time to engage in such observations and listening to children is when they are participating in rich and meaningful play activities (Fleer and Richardson, 2009) and this can be used to build up a 'learning journey' (Carr and Lee, 2012). Such formative and summative assessment practices aim to make children's learning ever more stimulating, rich and successful. A child's well-being and the characteristics of effective learning, such as resilience, perseverance and self-regulation learnt in the context of meaningful play are seen to be more reliable predictors of later academic achievement (Bodrova & Leong, 2007; Whitebread & Bingham, 2012; Siraj-Blatchford & Kingston 2015) rather than 'short-term academic results' which may not last (TACTYC, 2015).

The prominence of this research and related views of holistic learning among early years practitioners is likely to be the reason why such high number of schools (12,000, over 70 per cent) signed up to Early Excellence as their Baseline provider of choice (TES, 2015a). Early Excellence have promoted themselves to early years educators as retaining holistic teacher observations, including characteristics of effective learning, whilst dissociating themselves from 'testing children'. They describe themselves as providing the only 'non-invasive' Baseline Assessment (Schoolsweek, 2015). However Early Excellence has to work within the Government's guidelines for Baseline, so any maths, literacy or characteristics of effective learning are measured with a simple "yes" or "no", resulting in a number of points. As argued by one of the campaign groups against Baseline, reducing a range of curriculum areas and the complexity of learning characteristics and dispositions to a single score remains problematic in terms of the current dominant views of learning and child

development (TACTYC, 2015). Approximately 3,000 schools are reported to have decided not to do a Baseline in 2015 (TES, 2015b).

As the majority of schools selected Early Excellence, it is worth noting here the form of this assessment. The EExBA is made up of 47 statements to which the teacher has to answer simply 'Yes' or 'No'. There are nine statements for 'Characteristics of Effective Learning', 22 statements for the 'Prime Areas' (Personal, Social and Emotional Development; Communication and Language; and Physical Development), and 16 statements on 'Specific Areas of Learning and Development' (Literacy and Maths). This language mirrors that of the EYFS Profile, making it more familiar to Reception teachers. Statements are labelled either A or B and a child is only assessed on the B statements if they achieve the A statements. Examples of statements are included in Appendix A. Results for Early Excellence were submitted by teachers in the first six weeks of term and returned by Early Excellence to schools with some basic analysis in the first week after half-term (early November). The two alternative Baseline Assessments, as discussed in Appendix A, use either a tablet-based short individual assessment (CEM) or a combination of tasks and some observation (NFER).

During the period of the research study and immediately after, further issues relating to Baseline Assessment have been reported in the press. First, there were reports on the decision of a high-profile headteacher not to conduct Baseline Assessment, which was seen as a rejection of government policy (TES, 2015a; 2015b). There was also speculation that Baseline Assessment would be abandoned, though this policy did not materialise (Griffiths, 2015). Finally, since schools have received their results from providers in early November, issues relating to accuracy with all three providers have brought the policy into question (Mansell, 2015). This has opened up further debate about the use of baseline scores at age four to predict children's attainment at age 11 (Mansell, 2015; Reclaiming Schools, 2015). These developments bring the issues of accuracy and purpose identified in this report to the fore.

Research questions

It is within the context described above that this research project sought to answer the following questions:

1. What impact does the Baseline Assessment have on the start of school for the children concerned?
2. To what extent does the Baseline Assessment support or undermine existing transition and assessment strategies used in schools (especially the EYFS Profile)?
3. How do parents and teachers perceive and experience the assessment?
4. How can this assessment be adapted to the needs of different groups of pupils, such as children with SEND or EAL, looked-after children, children from BME groups and from disadvantaged backgrounds?
5. Do teachers and school-leaders believe it to be a valid assessment and how do they intend to use the resultant data?

The Research study

The research was carried out in the autumn term of 2015, using a mixed methods approach involving a nationwide survey and five case studies of primary schools.

Online survey

The survey was distributed via the NUT and ATL email databases using the Bristol Online Survey service, and was completed by 1,131 people. Fifty per cent of respondents were Reception teachers, 38 per cent EYFS or Phase Leaders, 7 per cent Senior Leaders and the remainder were support staff or 'other'. The survey analysis of teachers included finding out those who experienced the introduction of the EYFS in 2007 (DfES, 2007), compared to those who were newer to teaching and had only experienced the recently 'slimmed down' EYFS in 2012 (DfE, 2012). The survey found that (47 per cent) of Reception teachers and headteachers had been teaching prior to the original 2007 edition of the EYFS and a further 42 per cent of the teachers

and headteachers had taught at least since its introduction in 2007. So, 89 per cent of respondents were very experienced teachers who had gone through significant policy changes regarding the early years curriculum and assessment during their careers. The majority of respondents (80 per cent) worked in state-funded non-academy schools and 16 per cent worked in Academies. Most worked in schools graded by Ofsted as 'good' (63 per cent) or 'outstanding' (21 per cent). The Baseline Assessment provider used by survey respondents reflected the proportions nationally: 76 per cent used Early Excellence, 10 per cent CEM and 11 per cent NFER. The remaining 3 per cent used a discontinued provider or had opted to do no Baseline Assessment (1.9 per cent). Results were analysed with this demographic data to see if there were differences in responses due to differences in school situations, roles, length of service and Baseline Assessment provider (full details in Appendix B).

The survey involved a number of questions for teachers and school leaders on their views and experiences of Baseline Assessment. For some questions respondents were asked if they 'agree a lot', 'agree a little', 'disagree a little' or 'disagree a lot' with key statements on Baseline Assessment. The survey was piloted with serving teachers before distribution. The data generated by the survey was conducted using the Bristol Online Survey service and exported and analysed using SPSS, a quantitative data analysis software programme. The majority of respondents answered all the questions. In cases where there were non-responses, percentages are reported as a proportion of those that answered the question rather than of the whole number.

Case studies

Case studies were conducted in five primary schools across England, as detailed below. Purposive sampling ensured that the five case-study schools represent different types of school and different areas of England. At each school, Reception teachers, EYFS co-ordinators, headteachers and other school leaders, and parents were interviewed using standard interview schedules. The parents were interviewed as a group but all others were interviewed individually with the exception of one interview with two teachers. Interviews were recorded and transcribed professionally for analysis. Qualitative data was analysed using the themes generated by the research questions. A total of 35 people were interviewed, comprising five headteachers; two assistant headteachers or EYFS co-ordinators; 13 Reception teachers; and 15 parents.

Region	Status and size	Most recent Ofsted category	Baseline Assessment chosen
London	Community, three form entry	Good	Early Excellence
North	Community, three form entry	Outstanding	Early Excellence
South England	Community, two form entry	Good	Early Excellence
South West	Community, four form entry	Good	Early Excellence
West Midlands	Church of England VC, two form entry	Good	Early Excellence

Table 1: case-study schools

The sample is limited in that we were unable to gain access to an academy school, and all of the case-study schools had chosen the Early Excellence Baseline Assessment, reflecting the trend across the country. Our sample of parents is also limited in terms of gender and social diversity, as most respondents were white mothers. For reasons of anonymity the schools are referred to as School A-E and this ordering is not linked to the list above.

Ethical considerations

The research was conducted within the ethical guidelines provided by the British Education Research Association and the UCL Institute of Education. Care has been taken to ensure anonymity of all respondents and the security of data. For ethical reasons relating to maintaining anonymity, the case-study schools are identified only by a letter and the letters are not linked to the demographic information listed above. Schools were recompensed with funding for either a half day or full day of teaching cover (depending on the number of interviews) to reduce the impact of the research on the children.

Section 2
The impact on children starting school

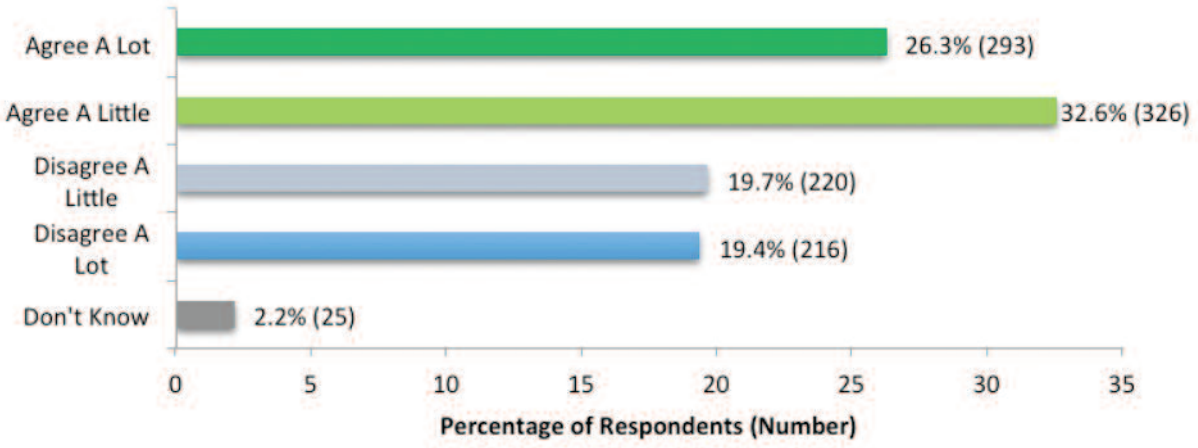
In this first section on our findings we discuss the varied responses of teachers to Baseline Assessment in terms of teaching and curriculum and the potential disruption to the 'settling in' period at the start of Reception.

2.1 Impact on teaching and curriculum

Our overall finding is that Baseline Assessment has had a distinct impact on the start of school in terms of teaching and curriculum. The extent to which this impact was positive or negative and how much Baseline Assessment disrupted the important 'settling in' period of the first few weeks in Reception was unclear from the survey and case-study data, however. This ambiguity was confirmed by the interviews, which suggested that the variety of different approaches taken resulted in different feelings towards the assessment.

The survey data (Figure 1) shows that a majority of respondents (59 per cent) agree that 'The Baseline Assessment has disrupted the children's start to school', although the proportion that 'agree a little' is the highest response. This does not vary with length of time teaching or by provider.

Figure 1: Responses to 'Baseline Assessment has disrupted the start to school for reception pupils'



This finding of disruption to the start of school for many teachers is significant given that the assessment is not intended to be useful for a number of years. Survey comments were mixed, with responses including both 'quick and not too time consuming' and 'very time consuming', but with the majority identifying negative impacts. The divergence of views here may reflect the wide range of responses to Baseline Assessment and the different styles of the three providers.

Nonetheless, in the case-study schools, which all used the EExBa, there were also mixed views on the disruptive nature of Baseline Assessment. At School D, where according to one teacher 'we carried on doing exactly what we did', another commented that the familiarity of the observation based assessment limited the impact:

> *I don't think it has been disruptive because – I think with this Baseline it was our ethos of what we wanted to do. […] I haven't really changed my practice …. I think if we had gone with a different provider than Early Excellence then I think it would have made a huge difference. (Teacher 3, School D)*

At other schools we found there were changes to practice which resulted from Baseline Assessment, mainly providing extra activities and 'stopping teaching'. In all of the case-study schools, teachers described organising provision of activities around the statements included in the Early Excellence Baseline Assessment.

> *This year it is very much a case of find our way through it I suppose, working out how to change our curriculum to meet the needs of what we need to do for Baseline. […] We are having to do activities to find that information, so that has taken away from being able to get into the areas of provision and work with the children, and showing them what types of things they can do with the new resources they have got. (Teacher 1, School A)*

> *[We] thought that actually we haven't really looked at that [some areas] so we had better do something about that, so we might put out a particular activity to almost cover those points. (Teacher 1, School B)*

> *Rather than go with the children's interests – of what they were interested in – I have geared what I have been setting up in the class to try help me gather information for the purpose of this assessment. (Teacher 3, School C)*

> *Most of our observations were just from what we had seen in everyday situations, unless we had to create a situation because we have got a couple of children who weren't showing us that they could do it. (Teacher 2, School D)*

> *We assessed usually one activity by activity. So we set up an activity that we knew would fulfil the criteria for certain statements and then watched the children access it and then we just went down [the list of names] and said, yes they can do it. (Teacher 2, School E)*

This practice existed alongside observation of children for the other statements, but did still have an impact; some teachers commented on the impact of not following children's interests on building relationships.

Another more dramatic impact was the decision to 'stop teaching' until the Baseline Assessment was complete, in order to provide a 'true baseline':

> *We have stopped teaching. You know normally we would do very gently settling in, rules, all of that. But then we would start with some of the more formal teaching activities, you know a little bit of phonics, but we have not, we have held back because we don't want to influence judgements. (Teacher 1, School A)*

This practice was described as being advised by Early Excellence, and in one school (School E) it was seen as a positive development which allowed the teachers to focus on children's personal and emotional development.

> *I liked the way that they said that, that you have got to look for their well-being and involvement before you worry and you had the six weeks to do it. I mean we did try and do it before the six weeks was up because then we got into the whole debate of EExBA would say don't do direct teaching. I mean even though there is teaching going on with any interaction, there will be teaching. But in terms of like the phonics we didn't start until sort of three or four weeks in. (Teacher 2, School E)*

The range of different responses here suggest real confusion over how schools should adapt their practice during the Baseline Assessment period, and over the extent to which the assessment should be a 'true baseline'. As we discuss in the next section the different responses will also have an effect on the results, as some children were given more opportunity to show they fulfil the criteria.

It is also important to note in terms of impact on practice that there were some survey comments on the use of the Baseline Assessment results to group or stream children (as previous on-entry systems had been used):

> *Also there is no time given to these poor little children to settle in before they are assessed and in our school they are put into ability groups based on these results! (W)*

> *It helps us group the children in differentiated maths and phonics groups. (W)*

The potential for grouping and labelling children on Baseline Assessment data is a worrying development especially given that many respondents queried the accuracy of the Baseline Assessment:

The data is skewed and it didn't give an accurate result of children's ability on entry. (W, CEM user)

The CEM Baseline Assessment is inaccurate and so the data we have received as a result is unreliable. (W, CEM user)

Baseline Assessment is the biggest farce I have undertaken during my entire teaching career, the potential for children to guess at answers or to misinterpret things is too vast to give a clear representation of where children are at. (W, NFER user)

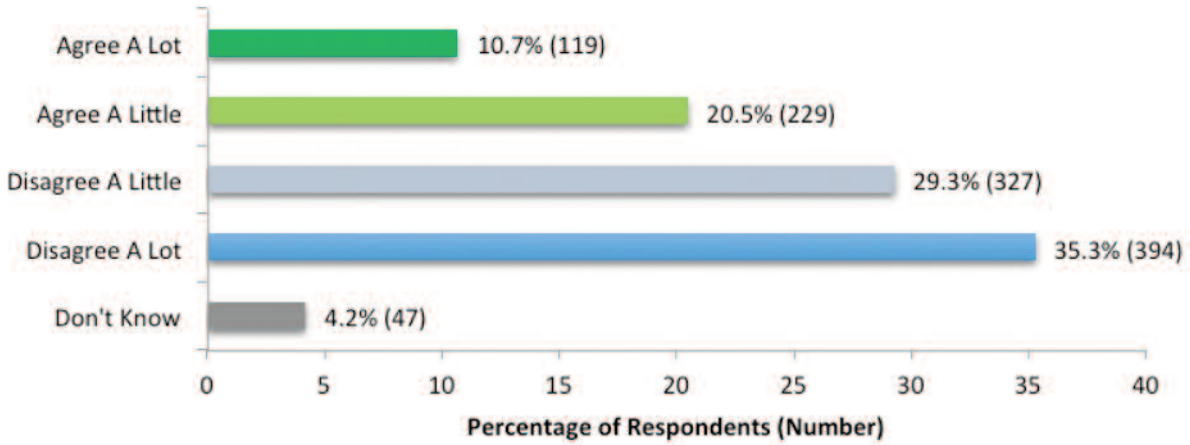
CEM has a definite element of luck/ chance so it doesn't provide enough accurate data. (W, CEM user)

Throughout this research teachers and headteachers using all three providers voiced serious doubts about the accuracy of the Baseline Assessment data collection process and its analysis. This is further discussed in Section 5.

2.2 The impact on building relationships

The establishment of relationships between teachers, support staff and pupils is usually regarded as a key focus of the first few weeks of Reception, as it may be some children's first experience of a formal educational setting. Therefore one important question in relation to Baseline Assessment is the impact on the teachers' ability to build effective relationships with children while conducting the assessment. The survey data, as shown in Figure 2, shows that respondents disagree that there has been a negative impact (65 per cent) and a majority strongly refute the suggestion. However, over 30 per cent of respondents agree there had been a negative impact, which is notable.

Figure 2: Responses to 'Baseline Assessment has negatively affected the development of the relationship between pupils and teaching staff'



Similarly, some comments on the survey and in the case-study interviews suggested the focus on assessment did have a detrimental effect on the 'settling in' period:

> *I feel that the Baseline Assessment has to be completed too early in the year and means that teachers are madly trying to collect evidence, rather than concentrating on the welfare of their new pupils and helping to create a calm and relaxing environment which is vital for a positive start to their school life. (W, EExBA user)*

> *The Baseline Assessment took me away from getting to actually know the class. We didn't gain anything from it. Some children looked at me and said "I can't read" when asked to read parts of the assessment. It was heartbreaking to see their reaction to it and I spent a lot of time reassuring children. (W, CEM user)*

> *If you have got 60 young people coming in through the door and in six weeks' time you have got to tick 47 boxes about all of them, of course your mind is going to be on that rather than on talking to them about their nice shiny shoes and about their pet rabbit at home and all those things that give young people a sound, secure start to learning. (Head, School B)*

For some, this detrimental effect was directly linked to changes in the activities provided or the time taken:

> *You had to set it up so you could actually assess these things which makes it very difficult to make, to start forming a relationship with the children, which is so important at this time of year. (Teacher 2, School C)*

> *It has had a negative impact on the children in Reception as their first few weeks are so important to establish routines, rules and relationships with staff and this was hugely impacted by the time it took to administer the assessment. (W, NFER user)*

Interestingly, the survey comments suggest that potentially there was also a negative impact on relationships even when the shorter alternative Baseline Assessments were used:

> *Although I agree with baseline in principle, the NFER package prevented us from building relationships with children at a crucial stage. (W, NFER user)*

> *A complete waste of time assessing the children as they start their very first days at school. We have not had the time to build relationships with them. (W, NFER user)*

> *It has to be done as early as possible at the crucial time when you should be developing relationships with them not ignoring the majority and looking at a screen one-to-one with one child. (W, CEM user)*

> *The reception teacher in our school found it impossible to administer the base line test and teach a class of new reception children. A supply teacher had to be drafted in so that the tests could be carried out. The result was that our reception teacher was not able to work with her new class in setting expectations for behaviour, rules, routines or getting to know the children. The CEM test caused a lot of disruption to learning and had an negative impact on relationship building. (W, CEM user)*

One final point is that, as with the issue of disrupting the start of school, responses to this question varied depending on the approach taken at the school. At School E, where they 'stopped teaching' and focused on settling children in, there were positive responses:

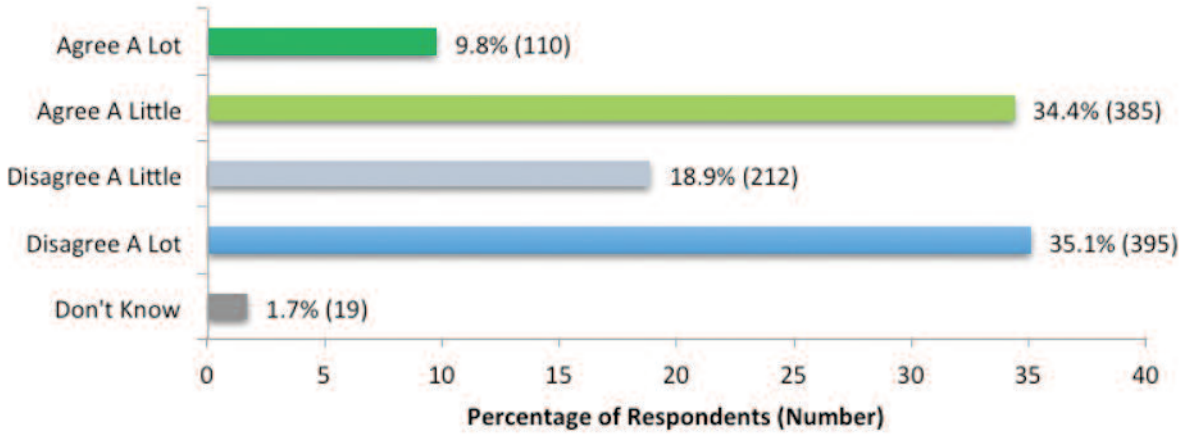
> *I found it quite enjoyable actually, you know just being able to focus on settling in, not worry about doing phonics too early and all the carpet time, the circle time and settling in, so from that point of view I think it was good. (Teacher 1, School E)*

Although the findings on these two related issues are mixed, we can conclude that Baseline Assessment is having an impact on the teaching and curriculum and on relationship-building during this first half-term of Reception. Teachers and school leaders are negotiating how they respond to this new demand; this is in keeping with previous research on policy enactment in schools, which emphasises the importance of school context and the ongoing process of adapting to new policy (Braun et al, 2012). Nonetheless, whatever the response, it is clear that Baseline Assessment as a policy is having a significant impact on classrooms. These findings, particularly on the social and emotional impact, add to the recent research on the impact of accountability measures on children throughout the education system (Hutchings, 2015).

2.3 Teachers' knowledge of the children from Baseline Assessment

In the survey, opinion was also divided about whether the Baseline Assessment had helped teachers to get to know pupils better (see Figure 3). However, although a high proportion 'agree a little' (35 per cent), many more express strong negative opinions (35 per cent) than strong positive opinions (10 per cent). There was some variation by role for these opinions: a smaller proportion of Reception teachers (7 per cent) agreed a lot than EYFS co-ordinators or Senior Leaders.

Figure 3: Responses to 'Baseline Assessment has helped me get to know reception pupils better'



The detailed data provided by the case studies revealed the complexity of views of how useful Baseline Assessment was in getting to know pupils. In general, the teachers interviewed were not enthusiastic about the role of Baseline Assessment in this process, especially given that in all case-study schools and for the majority of survey respondents, there were already on-entry assessments in place. The knowledge provided by the Baseline Assessment was often compared negatively to these other on-entry assessments. For some, Baseline Assessment was a barrier to getting to know the children:

> *I'm not sure it has given us any more information than we have always gathered. (Assistant Head, School C)*

> *So for the children who can't do things … with Development Matters [alternative system] they would have just given a band and perhaps a description, now they are just being told they are a 'no'. […] So you are not seeing as much, you are just marking it against the set statements it has got there. (Teacher 3, School A)*

> *I think the only problem I found was that because it was so important to get it in by a deadline you spent most of your time concentrating on baseline things rather than getting to know the children. (Teacher 2, School C)*

However, given a different approach as taken at School E, it could also be seen positively:

> *I actually feel compared to this time last year I have got a much better understanding of where all my children are … it gave me that time to really get to know them and particularly what was good with the Baseline from EExBA was the Characteristics of Effective Learning … I never looked at them before. (Teacher 1, School E)*

Written responses in the survey were also very mixed on this issue but there were more negative views expressed than positive. Typical comments included:

> *We have our own more detailed assessment arrangements which then feed directly into our teaching provision for each individual child. (W, CEM user)*

> *We had to do our usual baseline checks in addition to the baseline as it didn't tell us what we needed to know. (W, NFER user)*
>
> *We have a nursery intake so we know our children really well already, the Baseline Assessment didn't give us any information about the children that we didn't already know. (W, EExBA user)*
>
> *We have not been happy with our CEM experience and it did not tell us anything we did not already know (W, CEM user)*
>
> *We are a very large Primary with a lot of staff therefore our reception teachers continued to complete our own useful observational assessment whilst myself and two other colleagues pulled the children out to do the test. (W, CEM user)*

The issue of information from nurseries further limits the usefulness of the Baseline Assessment. For the case-study schools with nursery classes, Baseline Assessment was simply a time-consuming exercise that was not useful, and the attempt to assess children 'on entry' when they have already been in the setting for a year calls into doubt the purpose of assessing 'value added' by the school. One result of this is that schools may conduct a form of baseline on entry to Nursery classes (at age 3); indeed one survey respondent commented that they 'also had to use EExBA to baseline the nursery children' (W). This service is promoted by EExBA as complementing the Baseline Assessment. This is an example of the drift of policy into younger age groups, as seen with the Phonics test and the use of assessment data more generally (Bradbury, 2014; Roberts-Holmes and Bradbury, in press).

Overall, the data suggests that there are limitations in the use of Baseline Assessment to help teachers get to know pupils better, although it has be to remembered that this is not the stated purpose of the assessment.

Section 3

Identification of particular children and groups

This section considers the usefulness of Baseline Assessment in developing teachers' knowledge of children and identifying particular needs.

3.1 The limited contribution of Baseline Assessment to identification

In keeping with the findings relating to teacher knowledge of the children, Baseline Assessment was not seen as a useful tool in identifying children with particular needs. As shown in Figures 4 and 5, few survey respondents saw the assessment as helpful in identifying the needs of children with Special Educational Needs (SEN) or with English as an additional language (EAL). Figures for summer-born and looked-after children were similar. There was no variation between providers in answers to this question.

Figure 4: Responses to 'Baseline Assessment has helped to identity the needs of SEN children'

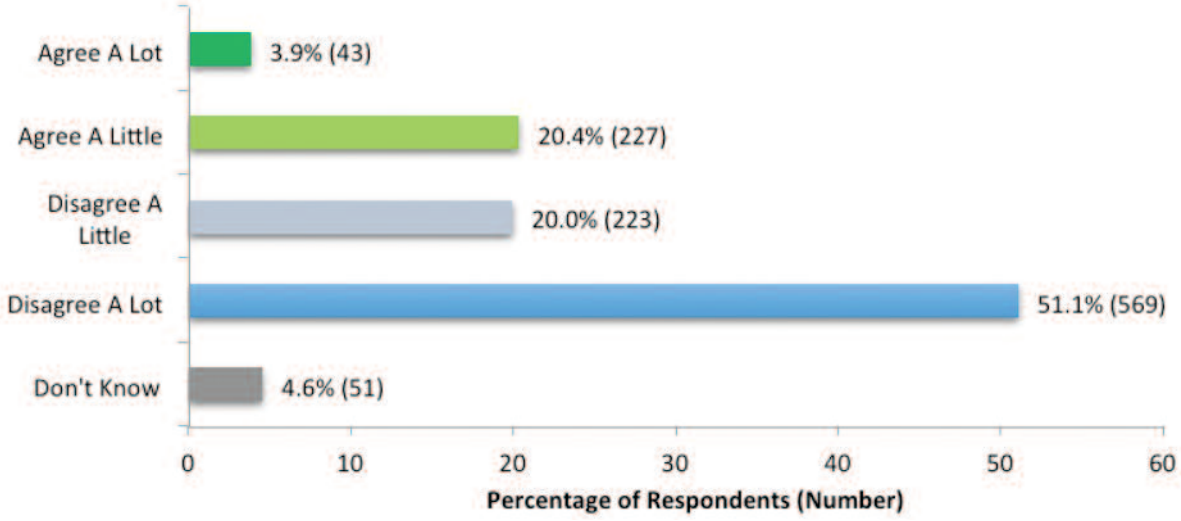
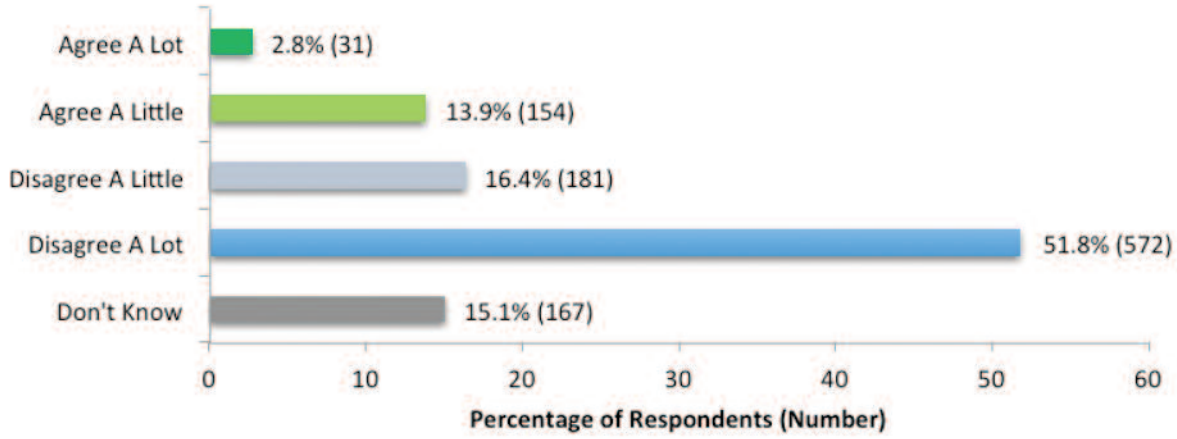


Figure 5: Responses to 'Baseline Assessment has helped to identity the needs of EAL children'



This data shows that teachers feel Baseline Assessment has a very limited role in identifying children with particular needs. The teachers and headteachers we interviewed and comments on the survey were dismissive of the idea that Baseline Assessment would help with identification:

> *We know – not really – we know if you are EAL don't we? We don't need a test to tell us that, we know if you can't speak English. (Head, School A)*
>
> *I think my practitioners' observations and assessments are particularly useful for identifying groups of children and their needs. Whether we need 47 questions that we answer yes or no to that have a score related to them to tell us that. (Head, School B)*
>
> *We don't need this to identify groups – we know our children and the 'groups' they fall into due to our observations and interactions. (W, EExBA user)*
>
> *We had to do our usual baseline checks in addition to the new Baseline Assessment as it didn't tell us what we needed to know about the children. (W, NFER user)*
>
> *We ended up doing our own assessments and then CEM, and have gone purely on our own assessments in order to then identify particular children and help them progress. (W, CEM user)*

A small number of comments suggested some advantages in terms of identification, such as higher achieving children and those with dyslexia. Overall, however, Baseline Assessment appears to have a limited use as a tool for identifying particular needs, especially where existing arrangements or information from nurseries already identified these children.

3.2 Concerns over the impact on EAL children

Information from the case-study schools suggested there were serious concerns about the assessment of children with EAL using the new Baseline. Unlike the EYFS Profile, Baseline Assessment has to be conducted in English, so there are significant effects for those children who do not speak English confidently. This is the case both with observation-based assessments which rely on children speaking and the more formal assessments which may include instructions which children cannot understand.

Teachers in our case-study schools commented:

> *The problem we have found this year is because we have to do it – it has to be conducted in English. I think that has impacted quite a bit. And normally for our on-entry assessment we*

would – most of our support staff are bilingual and for the majority language here – so we would use that to inform our on entry assessments. Now with the Baseline having to be conducted in English I think it will impact on the results. (Assistant Head, School C)

For EAL children [Baseline Assessment] was disadvantaging wasn't it, because you couldn't do it in another language, it had to be in English. (Teacher 2, School E)

It is important to note that by not allowing the baseline to be conducted in children's home languages means that what these children are capable of doing and their understanding is not accounted for. By not being able to assess EAL children's abilities, comprehension and achievements serves to further undermine the accuracy of Baseline Assessment and its credibility as a measure of progress. These concerns were also reflected in the survey comments, which showed that this was a problem beyond our case-study schools:

Baseline must be fair for EAL children – I had 6 this year, 3 of whom achieved less with Baseline Assessment but did well in my own assessments as I allowed them to use their own language. (W)

I come from an inner city school with very high numbers of EAL children. The prescriptive questions detailing what you can/can't say meant that some things I knew my children would be able to do with a simplified instruction, they couldn't do as part of the Baseline. (W, NFER user)

Not having any opportunity to note children with EAL is a concern as this will clearly mark them lower than necessary. (W)

Children who are EAL or didn't understand the question in a certain way were unable to answer questions correctly, affecting their scores. (W)

This is a serious concern relating to Baseline Assessment, as the lower attainment of EAL children would be both inaccurate in terms of measuring progress and risks setting lower expectations for a particular group. Potentially there are long term consequences for these children, who may be seen as having made expected progress even when they have low attainment later in their school careers.

3.3 Baseline Assessment and children with SEND

As shown above, over 50 per cent of survey respondents 'disagreed a lot' with the suggestion that Baseline Assessment had helped them identify children with SEND, and other information was seen as the main source for identification of additional needs. There were also concerns raised as to how to assess children with SEND, the accuracy of these assessments, and a wariness about labelling children very early on in their school career, which questioned the idea of 'identifying' children with SEND altogether:

I feel it just excluded children who have SEND and does not give you a clear idea as to where any child actually is! (W, EExBA user)

Some children with SEN scored above their true level in areas such as speaking. (W EExBA user)

Children that can't answer back [are disadvantaged] because they are either SEND or they are EAL so in that case, yes, it does disadvantage children because they communicate in other ways. (Teacher 2, School C)

And we know if you have got special needs, it may well have already been identified in Nursery, so we will already have that information. And we are very, very wary of labelling children as having learning difficulties when they start school. (Head, School A)

There is also great risk that children will be inaccurately assessed, with potentially damaging effects throughout their school careers.

3.4 Concerns over the impact of Baseline Assessment and low expectations

Further concerns, not raised in the survey questions but prominent in some of the interviews, related to the potential class impact of Baseline Assessment. This was related particularly to the impact of parents on the assessment:

> *If parents, like the parents you have spoken to this morning, pick up on the fact that this is happening every year there is the likelihood that some parents will be spending the whole summer holidays teaching, you know getting hold of this Baseline Assessment, looking at the criteria. Because there is this natural need for your child to pass tests you will be coaching them to pass the test. Other parents won't, so that is not a narrowing of the gap; that is an immediate widening of the gap that we already have. (Head, School A)*

This issue, where Baseline Assessment becomes an assessment of parental involvement and knowledge of the system, could have significant implications for those parents without the appropriate social capital, and thus cause further disparities in attainment by social background – an issue which the current and former Coalition Governments have been keen to address.

Finally, serious concerns were raised by some interview respondents about the potential for low expectations in general, based on Baseline Assessment. One headteacher commented:

> *You know you don't want limiting judgements at this point. Because obviously what you are looking to do is open potential up, and I know that sometimes by measuring that and saying you could have issues here it might enable you to do that, but actually it can lead to low expectations as well. So obviously what we are in the business of trying to do is identify the needs as early as possible, but what you are not wanting to do is to say this is happening here, therefore this is what we expect of you here; that might be too low. But actually that worries me, I think there has been a real issue around target setting in terms of lower expectations, actually. Because you have hit your target you don't need to go higher, that somehow people don't work beyond that. […] If you are looking at staff in schools and expectations of children, yes, I just have questions around that kind of information setting expectations for the future at that very early stage. (Head, School C)*

The potentially damaging effect of Baseline Assessment is identified here: that for children with low scores ('below typical'), even if they make good progress, it will be seen as acceptable for them to remain low-attaining at age 11. This is a problem inherent in any 'value added' measure where the baseline is known and is more likely to affect those groups who are lower attaining within the system in general, such as ethnic minorities, children receiving free school meals, children with SEN and EAL and some summer-born children (Bradbury, 2011). Although schools have not been advised to set targets or make predictions based on Baseline Assessment, the form of the assessment makes this inevitable.

One of the major frustrations that teachers noted was that data from Baseline Assessment did not fit into existing tracking software. The nature of this software encourages the use of predictions based on expected progress, and thus the danger of low expectations. Therefore the risk of particular groups of children being systematically under-assessed in Baseline Assessment is significant for their long-term educational trajectories. The individual reporting mechanisms to parents provided by NFER and CEM may strengthen this tendency.

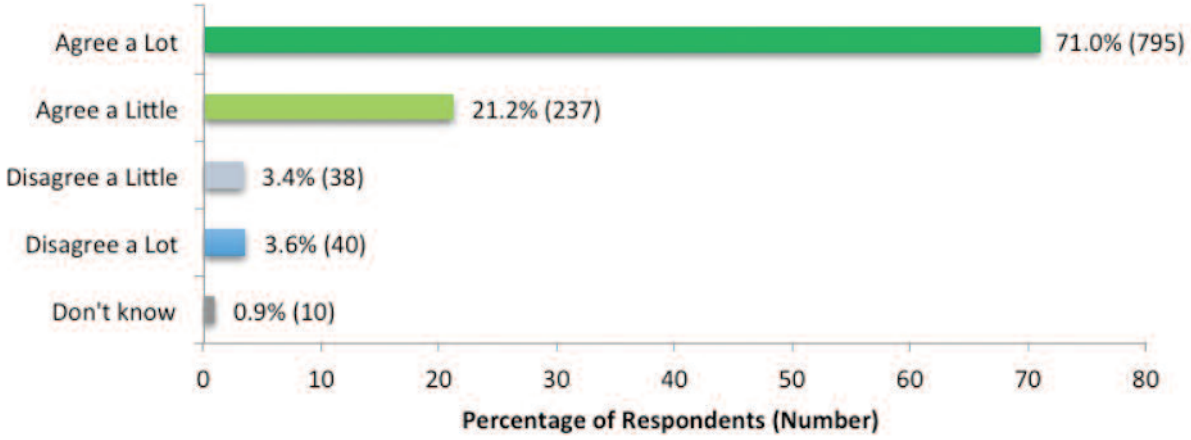
Section 4
Relationship to existing assessments

In this section we examine how Baseline Assessment relates to the existing assessments in Reception classrooms, which many teachers see as sufficient and more useful.

4.1 Existing on-entry assessments

As mentioned, all the case-study schools and the vast majority of survey respondents' schools already had in place existing assessments which were conducted on entry to Reception. Over 90 per cent of survey respondents agreed that these arrangements supported teaching and learning, as shown in Figure 6.

Figure 6: Responses to 'My school has existing assessment arrangements at the start of reception which support teaching and learning'



These assessment arrangements were described in both the written comments and in interviews as less time-consuming and as sufficient for gathering information and identifying children's additional needs. Many of the teachers stated that the children are presently assessed against the EYFS Development Matters criteria using detailed observations, discussions with parents and reading nursery reports. Under their existing and effective on-entry systems the children are assessed at the end of the spring and summer terms and their progress checked by comparing this against their Autumn Development Matters baseline. Respondents saw the use of these existing baselines as more in keeping with the ethos of early years, or as evidence of their experience and expertise. Existing baselines were also more compatible with the EYFS Profile and other tracking systems:

> *We all do a baseline across all 17 strands of the Foundation stage curriculum and we will age band. We already did that. […] We would then repeat those assessments at Christmas and at Easter and then at the end of the year. (Teacher 1, School A)*

> *The results were of no real value as they did not pinpoint which areas the children could/couldn't do, it just gave an overall judgement. In order for me to see where they were going right or wrong I would have had to record their answer on a separate piece of paper, causing even more work-load. Some of the most important areas just are not assessed for example the ability to hold a pencil! (W, CEM user)*

> *As it is not linked to the EYFS through which we are monitoring the progress of our children this baseline meant extra work to make it 'fit' our LA assessment system and took extra time to complete as we could not use it for tracking in our own room. We still completed our own baseline to allow us to track the children through all 7 areas of learning. Not a good experience, results we cannot use and extra work-load again. (W, CEM user)*
>
> *We have our own more detailed arrangements which then feed directly into our teaching provision for each individual child. (W, CEM user)*

Moreover, survey comments suggested that existing arrangements were seen as more useful for future planning:

> *Our in house baseline told us what percentage of children were achieving well below, below at or above age related expectations across each area of learning and within each strand. Also the resulting NFER analysis was less than useless and told us nothing. Their data is so thin it tells us nothing and is incomparable with our own. (W, NFER user)*
>
> *We already make our own Baseline Assessments to help us teach effectively. The compulsory and rigid nature of the NFER Baseline Assessment did not tell me anything I would not have assessed anyway and was restrictive. (W, NFER user)*

There were some comments which suggested that formal Baseline Assessment was an improvement on existing on-entry assessments, because it encouraged teachers to think more broadly. For example, as discussed above, the teachers at School E saw the assessment as allowing them to think about the characteristics of effective learning in more depth. For others, the speed of the assessment was an improvement.

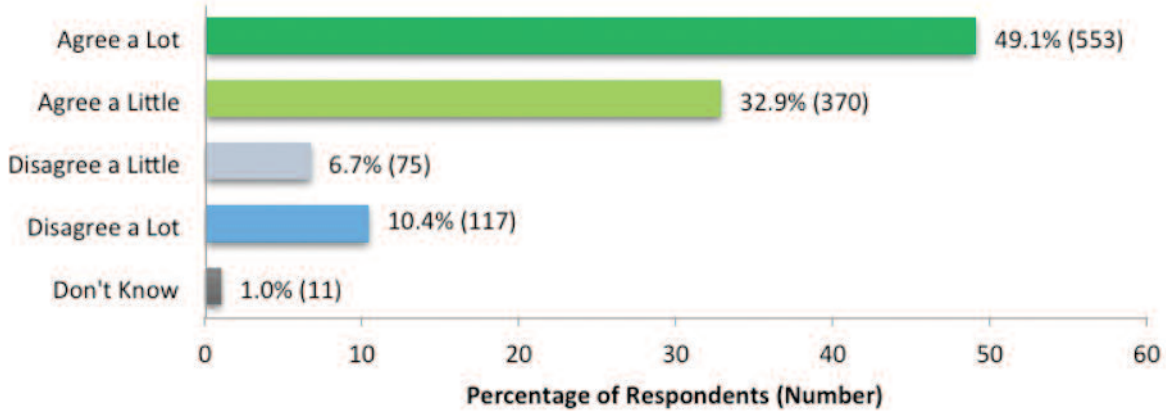
> *As a teacher in a large EYFS unit we used to carry out on-entry assessments. Having this Baseline Assessment to work to seemed to make it simpler and quicker. (W, NFER user)*

Again, the extent to which Baseline Assessment was an improvement in terms of time partly depends on the provider selected and the approach taken by the school. Nonetheless, the majority of respondents appeared to prefer their existing assessment arrangements, particularly in terms of use for future planning and teaching.

4.2 The effectiveness of the EYFS Profile

There was strong support for the existing EYFS Profile assessment among both survey respondents and the teachers we interviewed. As shown in Figure 7, 82 per cent of respondents agreed that 'The EYFS Profile helps me to monitor the development of the reception class and plan for their learning'.

Figure 7: Responses to 'The EYFS Profile supports monitoring and planning in reception'



This strong support was based on the extensive nature of the Profile and the fact that it could be used throughout the year to show how children have progressed, in contrast to Baseline Assessment:

> *The EYFS allows us to develop a holistic understanding of each child through observation. (W)*
>
> *I find the EYFS profile a useful tool to assess the starting point of the reception children. As it is ongoing throughout the year this is the most useful way to monitor the children's development. (W)*

There were however some negative comments in the survey responses, largely related to the scale and content of the EYFS Profile:

> *The EYFS Profile does help to plan for learning but there are many gaps, especially in maths. (W)*
>
> *The EYFS Profile is too vast to track so many children. (W)*
>
> *There are definitely some gaps in the EYFS Profile as it stands, e.g. maths and understanding of the world. However if tweaked it is a good tool. (W)*

It should be noted, however, that the problems of the EYFS Profile here are in some ways relevant to all observation-based assessments in early years; they are in keeping with research on the previous iteration of the EYFS Profile, for example (Bradbury, 2013). As early years have only used statutory assessments since 2003, there remains general concern about assessing children formally at a young age, especially given the difficulty of making definitive decisions on subjective statements. We return to this issue of validity in the next section.

Section 5

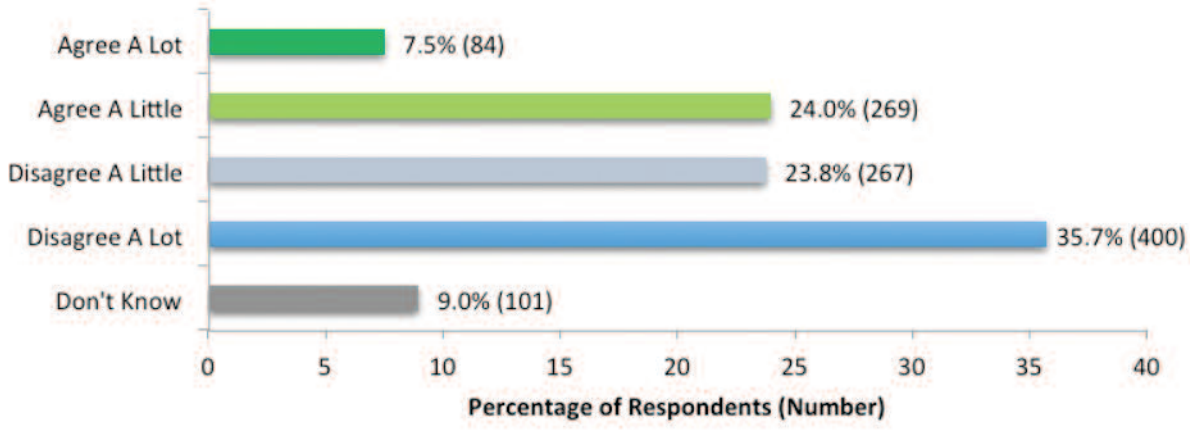
Perceptions of the accuracy and validity of the data

With all three Baseline Assessment providers there were serious concerns raised about the accuracy and validity of the assessment process. Issues such as guessing correctly on multiple choice computer-based statements whether or not children had attended nursery; whether they took the test in week one or later on and the interpretation to a binary statement of a complex learning process all brought the accuracy of the assessment into question.

5.1 Fairness and accuracy of data

A further important finding is that there were serious concerns about the accuracy and validity of the data produced by Baseline Assessment, for a number of reasons. Only 7.7 per cent of survey respondents agreed with the statement 'Baseline Assessment is a fair and accurate way to assess children'. When asked if they agreed or disagreed with the statement 'Scores obtained by the Baseline Assessment are an accurate reflection of children's attainment at this stage', a total of almost 60 per cent of respondents disagreed, as shown in Figure 8. The proportion that 'disagreed a lot' was high at 36 per cent; suggesting there was real doubt as to the accuracy of the assessment.

Figure 8: Responses to 'Scores obtained by the Baseline Assessment are an accurate reflection of children's attainment at this stage'



There was some variation by provider in answer to this question, with the 'agree' results for Early Excellence much lower at 28 per cent, compared to the other providers where there was a more equal split between agree and disagree. As discussed later in this report, this may be due to the problems of subjective statements on the EExBA assessment.

	Early Excellence (n = 849)	CEM (n = 120)	NFER (n = 121)
Agree	28 per cent	43 per cent	42 per cent
Disagree	62 per cent	47 per cent	49 per cent
Don't Know	9 per cent	9 per cent	7 per cent

➢ Percentages may not equal 100 per cent due to rounding

Table 2: Responses by provider to ''Scores obtained by Baseline Assessment are an accurate reflection of children's attainment at this stage'

Reasons for these doubts from the survey and case-study data included: the children are too young; varied and inconsistent training; subjectivity of the statements; and the lack of moderation and need for evidence.

5.2 The age of the children and previous experiences

Some teachers argued that a standardised assessment in the first six weeks of school when most children are four years old was developmentally inappropriate, and therefore they questioned the assessment's accuracy and validity. Some commented that the new and unfamiliar school context and routine would not be conducive to children showing their 'true potential', and that young children in their first six weeks of school lacked the confidence to demonstrate what they were capable of doing:

> *I did have children that were crying and I just couldn't get anything out of them at all because they were too upset to do anything, even when I left it later on. Some children just refused or just weren't ready and I know they said you only assess them when they are ready, but some children, well, you got to the point where you had to assess them because it had to be done whether they were ready or not. And obviously then it is not accurate because they weren't at a stage when they wanted to say things. (Teacher 1, School C)*

> *It's ridiculous. It's not a fair representation of children. Many young children are not yet confident enough to show their new teacher what they can do when put on the spot. (W, EExBA user)*

> *I think doing any sort of reputable assessment of very young children is dodgy because the children are so young. You know if those children were in Denmark they wouldn't have had to pick up a pencil yet. (Head, School D)*

This final comment reflects the ongoing debate about the compulsory starting age for school in England, which is considerably younger than many other European countries. As referred to here, Danish children do not start school until six years old, and a range of studies has shown how English school children, who start at four years old, are constantly assessed and tested compared to European children (Hutchings, 2015).

Teachers and school leaders also commented that the results would be unreliable because they did not take into account the children's prior experiences, particularly whether or not they had been to nursery:

> *Some children, they settle in much more quickly, so if we just talk about our nursery children they probably settle in much quicker than the other children. So you are doing things at different times anyway and the timing of it all is a bit tricky. (Reception Teacher 2, School B)*

> *It is not a level playing field – it is a snapshot. It is done with very young children when they have just joined this thing called school where everything is new and different. (Headteacher, School A)*

These differences in experience and the young age of the children invalidate the results of the assessment, and therefore make the measurement of progress from this point meaningless.

5.3 The problems of binary yes/no judgements and subjective statements

The DfE's guidance and approval process (DfE, 2014) stipulated that 'each assessment item must require a single, objective, binary decision to be made by the scorer'. Hence all three Baseline Assessments use binary statements requiring either a 'yes' or a 'no' answer. The Reception teachers in our case-study schools and comments from the survey were frustrated with such simplistic and unsophisticated binary phrasing of the complex learning statements. They strongly argued that children's learning was, in contrast, complicated and contradictory:

> *It is a binary yes/no! And children aren't like that, children are more complicated than that… those contradictions exist within the child and that is a true reflection of an individual, a unique child, who can be complicated. (Teacher 2, School B)*

This need for a binary judgements contrasted with the established 'best fit' approach using the Developmental Matters framework. This was seen as more accurate and reliable because it demonstrated what the children were capable of and how their learning might be developed:

> *With Development Matters it is about finding a best fit […] So I guess with the yes or no there is no room for that is there? There is no wriggle room. (Teacher 1, School B)*

> *With this, it is just a yes or no. Whereas we can show progress with Development Matters […] But there is no way of doing that with EExBa, it is either yes or no. (Teacher 2, School B)*

For these teachers, the 'best fit' description allowed flexibility within broad categories, demonstrating what the children could do, whereas the simplistic binary logic of the Baseline Assessment does not allow such subtleties. Survey comments also included discussion of this focus on the negative:

> *Baseline does not support our knowledge of children, as it merely states whether or not they can independently do specific things; this does not provide information as to where the children ARE, simply where they are not. (W, CEM user)*

> *It is difficult to assess whether the child should achieve it or not (i.e. they may achieve part of the statement, but not another part of it, meaning the official answer is 'no') and so the data does not give the whole and true picture of a child's abilities in that aspect. (W, EExBA user)*

For some teachers this lack of flexibility in making judgements about a child's abilities and potential led to considerable frustration. The highlighting of what the children could not do and negative impression of the child created by focusing upon their deficiencies was seen as an unreliable method of assessment. Reception teachers are used to a range of detailed, comprehensive narrative-type assessment practices, which enable a teacher to demonstrate what a child is capable of and reflects their competencies and skills. The contrast between Baseline Assessment and this approach, as used in the EYFS Profile and Development Matters, was stark. The two methods of assessment represent different paradigms in terms of early years philosophies, which were difficult for the teachers to reconcile.

A further major problem was the subjective nature of the binary statements used to different extents by all three providers and the multiple ways in which they could be interpreted; for many teachers, this rendered the whole assessment inaccurate. The need to produce a yes/no response to statements linked to complex and context specific characteristics of effective learning, such as, 'risk taking', 'curiosity' and 'persistence', were seen by one headteacher as 'rather crude and bit silly' (Head, School B). Another headteacher went through some of the characteristics of effective learning statements with the researcher to demonstrate the wide range of interpretations that were possible in deciding whether or not a child should be awarded a 'yes' or 'no' judgement. She argued that teachers' different experiences, understandings and interpretations would lead to subjective judgements that were unreliable and invalid:

> *So look at this one: 'Curiosity, shows curiosity about objects in the world around them and has particular interests'. That is open to such a wide range of interpretations. And also you are judging children according to what you think is typical, aren't you? You might be judging children according to the types of children you have had in the past or your own child or – so you see how vague it is and very subjective. (Head, School A)*

This experienced headteacher was doubtful about the accuracy of Baseline Assessment; she also commented that assessment at this age produced 'wildly different judgements from school to school, teacher to teacher'. Similar concerns were reflected throughout the case studies and survey sample:

> *Observational assessments are subjective, they are only as good as the teachers' knowledge of what they are assessing. (Teacher 2, School E)*
>
> *The interpretation of some statements and exemplifications are open to misinterpretation and have caused uneasiness among our reception team. (W)*
>
> *The statements are still open to interpretation which is why the data reports are not showing an entirely accurate picture of children. (W)*
>
> *There have still been huge inconsistencies about the practitioner expectations for judging how children have met the statements. (W)*

The survey responses confirmed that these concerns about inaccuracy applied to all three providers. Of particular concern regarding the computer based assessments of CEM and NFER was the potential for children to guess answers correctly from multiple choice questions and to misinterpret computer instructions.

> *The CEM test was easy to implement but didn't give an accurate result of children's ability on entry. (W, CEM user)*
>
> *CEM has a definite element of luck/chance so it doesn't provide enough accurate data. (W, CEM user)*
>
> *Baseline Assessment is the biggest farce I have undertaken during my entire teaching career, the potential for children to guess at answers or to misinterpret things is too vast to give a clear representation of where children are at. (W, NFER user)*
>
> *It took a long time out of my teaching time and was open to lucky guesses, particularly in the maths part. (W, NFER user)*

As well as this statistical flaw with such computer testing, respondents also pointed out how four year old children, if tested on a computer at the end of a long day, were tired and would do less well than those tested in the morning:

> *Due to having to do one at a time, some children were assessed in the morning and others at the end of the day when they were tired and easily distracted. (W, NFER user)*

Similarly, respondents pointed out that if children were tested at different weeks then their computer test results would also be distorted because young children learn at a rapid rate in their first few weeks in Reception:

> *In addition 1 member of staff administered the whole test so those children who did in the first week of starting generally had lower scores than children who it in weeks 5 and 6 into the term. (W, CEM user)*

The danger here is that even so called 'objective' computer tests are in fact distorted, inaccurate and invalid due to the multiplicity of testing variables in individual school contexts not being taken into account such as the conditions, timings and arrangements of the tests. Interestingly such concerns around the inaccuracy of Baseline Assessment have already been raised in one survey school during an Ofsted inspection:

> *Our schools concern is that we don't feel the test reflects the actual level of children, i.e. the children in my school have always in our opinion been well above national average yet the test shows otherwise and that children are below the average. We had Ofsted in this week who queried the accuracy of our test results as they felt children were above average...but we have simply followed the test. (W, NFER user)*

Thus CEM and NFER which are less reliant on teacher judgement (although they too have teacher observation sections) were seen as inaccurate and invalid. This is in keeping with the argument that suggests it may be simply impossible to design an assessment which accurately assesses children at this age.

5.4 The need for evidence, training and moderation

The lack of evidence required for judgements made as part of Baseline Assessment, in contrast to the detailed evidence required for existing assessments in Reception, was seen as a serious flaw:

> *Well the no evidence to me makes it less solid, doesn't it. Because there is no evidence, it is just a practitioner's point of view. (Head, School A)*

> *And I do think that is important and to know exactly where the child is and having evidence to prove that is important, but something crude that is just yes and no, I don't know if that is any use long term or not. (Teacher 2, School A)*

Without evidence to support the judgements being made these teachers doubted the reliability of Baseline Assessment; as a result, many planned to keep their existing assessments.

The above inconsistencies around interpretation were exacerbated by the limited training teachers received. Often only one teacher attended training and then 'cascaded' the information to others, but some schools paid for extra staff to attend. There were also concerns about consistency of the Early Excellence training:

> *Even though there has been training, not everybody has gone on the training…and it would appear that not all the training has been cascaded to the practitioners and there seems to be quite a bit of confusion…For some children the baseline is not being done until six weeks, for some it is done on week one or week two, and I am not sure of the validity of that. (Assistant Head, School C)*

> *Statements need to be specific and not open to individual interpretation. Training needs to be consistent across the country! (W)*

One solution to these problems of lack of evidence and variations in training was the use of moderation between schools; this is an established practice in relation to the EYFS Profile. Moderation of teachers' judgements was not required by the providers, but teachers wanted to ensure that their judgements were as accurate and fair as possible and so had organised moderation meetings within their schools and cluster groups:

> *So it is open to interpretation, as I say it would have been very hard to do without having another teacher to talk to and sort of sound off ideas and what do you think is the best way to make sure this evidence is a true reflection. (Teacher 1, School E)*

> *I know we are not meant to moderate but I think there is a place for moderation if only to ensure that everybody is singing from the same sheet when they are doing the assessments. (Assistant Head, School C)*

For these teachers, the need to produce 'accurate' data was more important than the instruction not to moderate and the sacrifice in terms of time. Even so, for some moderation did not solve the problem: one written comment argued 'Having moderated between cluster schools it is still very subjective' (W). One headteacher was frustrated with the associated costs in terms of time and money especially when the teachers should be engaged in forming relationships with the children:

> *How you are going to moderate it without wasting huge amounts of money and time? You know I could pay thousands of pounds to have all my reception teachers spending a month moderating when really what they should be doing is nurturing the children and making sure that the children are settled and getting to know the children. (Head, School A)*

Thus effective moderation would have additional costs and still remove teachers from the classroom. Overall there was a strong desire to 'get it right', even if it meant increasing their work-load. It was important to teachers to assess as accurately as possible, and therefore, as seen in the next section, the problems caused by Baseline Assessment affected their sense of professionalism.

Section 6

The impact on teachers' roles and professionalism

All three providers' Baseline Assessments had a significant impact on teachers' work-loads, particularly where existing arrangements were used alongside the new assessment. Teachers' feelings of professionalism were challenged and undermined by Baseline Assessment. However, they begrudgingly accepted the policy.

6.1 Increases in work-load

In the survey, teachers overwhelmingly reported an increase in work-load both inside and outside the classroom (see Figures 9 and 10). A very high proportion of respondents (82 per cent) agreed that Baseline Assessment had increased their work-load within the classroom, and a slightly higher proportion (84 per cent) agreed it had increased their work-load outside the classroom.

Figure 9: Responses to 'Baseline Assessment has increased work-load within the classroom'

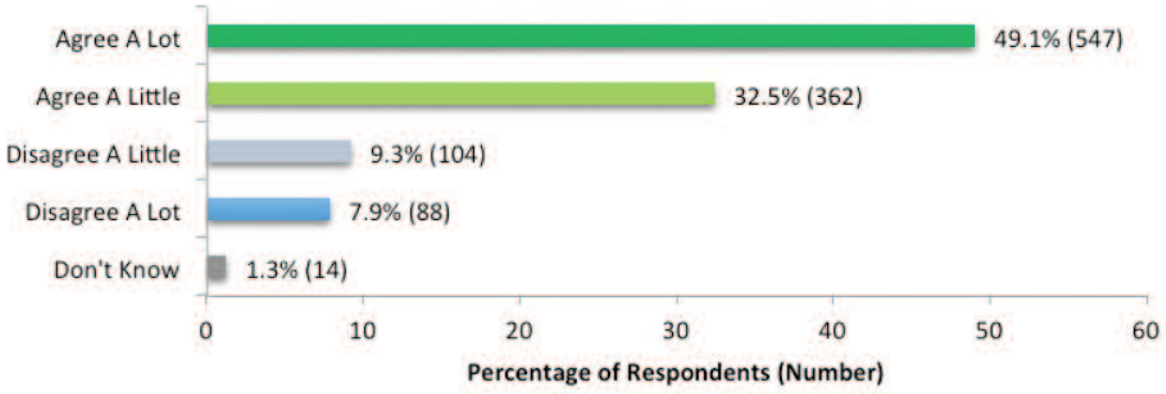
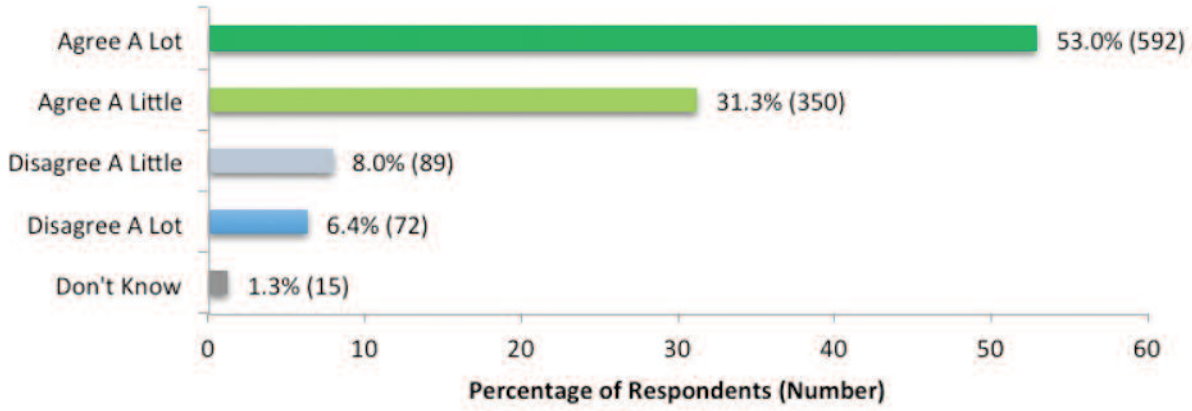


Figure 10: Responses to 'Baseline Assessment has increased work-load outside the classroom'



The variation between different providers (shown in Table 3) shows some difference but perhaps less than would be expected considering only one Baseline Assessment (EExBA) is fully observation based.

	Early Excellence (n = 849)	CEM (n = 120)	NFER (n = 121)
Agree A Lot	48 per cent	50 per cent	54 per cent
Agree A Little	33 per cent	24 per cent	34 per cent
Disagree A Little	10 per cent	15 per cent	4 per cent
Disagree A Lot	8 per cent	8 per cent	6 per cent
Don't Know	1 per cent	0 per cent	0 per cent

Note: Percentages may not equal 100 per cent due to rounding

Table 3: Responses to 'Baseline Assessment has increased my work-load within the classroom'

The percentage agreeing with the statement about increased work-load is actually higher for the test-based instruments used by CEM and NFER. On a separate question, 75 per cent of survey respondents agreed with the statement 'Baseline Assessment is an additional burden on Reception teachers'.

Similarly, the Reception teachers we interviewed and the written responses to the survey overwhelmingly reported increases in work-load. Baseline Assessment generated extra work through additional training, the time taken to make observations and judgements, discussions with colleagues and inputting data. This additional work was combined with the time taken to conduct the usual on-entry assessments in the first half-term of Reception, so that teachers argued they were doing a 'double job':

> *At the moment it feels like we are doing a double job that is a huge increase in work-load. (Teacher 1, School A)*
>
> *The new system is currently running alongside our proven system increasing work-load on staff with no extra benefits. (W)*
>
> *We did a comparability study with CEM and EExBA and they were very different but both very time consuming. We have been disappointed with the data from both baselines. (W, CEM user)*
>
> *It took much too long for each child to complete the assessment and they often became tired and irritable and it took a LOT of time out of class for teachers. (W, CEM user)*

Was very time consuming and wasn't accurate. (W, NFER user)

Baseline Assessment was seen as the cause of an unnecessary increase in work-load as the existing systems were preferable, and as a duplication of teachers' assessment-related work-load. Some teachers even reported that they carry out three baselines; the Development Matters baseline, a Local Authority baseline and now the new Baseline Assessment:

We are repeating our current work, we do 3 baselines; a school one, an LA one and now one for national early excellence. (W)

At the beginning of the school year we find that we assess our children 3 times! On entry to inform planning, baseline and LA requirement. This significantly affects our work-load. (W)

It could be argued that assessment in Reception is currently in a transition period, where some duplication of assessments results from the overlap of different policies. It is important to note that Baseline Assessment was never designed to give teachers information that would be useful to them in their daily planning and preparation; it was designed as an accountability tool. However, since the new assessment was unhelpful to teachers and provided no new knowledge about the children, it seems highly likely that teachers will continue to use their own systems alongside the Baseline Assessment in the future. Such attempts to marry the new Baseline Assessments to their existing assessment arrangements were not an intended incrementalism, but nevertheless teachers were trying accomplish it.

One particular source of concern and stress for Reception teachers was the stipulation that Baseline Assessment had to be completed within the first six weeks of the autumn term:

90 children in six weeks to get to know them inside out and upside down is a big ask. (Teacher 1, School A)

Essentially you are having to make the same judgements twice. And in a more hurried way. (Teacher 1, School B)

This time pressure was compounded by the need to become familiar with a new assessment system and the data entry involved. The latter task was seen as particularly onerous and not in keeping with the priorities of early years teachers as expressed in the survey comments.

The system for submitting data to CEM has been a nightmare, with some assessments having to be done three times before the data uploaded. This has caused considerable stress and has added to the work-load of FS staff. (W, CEM user)

The baseline took up huge amounts of time. (W, NFER user)

The Reception teachers at our case-study schools reported that they were using their PPA (Planning, Preparation and Assessment) time to collate, moderate and input the data into online systems rather than engaging in planning, and that they were inputting the data online at home in the evenings and weekends just to stay 'on top' of all the data. In the largest case-study schools, the Deputy Head and/or the EYFS lead took responsibility for inputting data online as they were concerned that the extra work-load might adversely impact upon teachers' well-being. At one of the larger case-study schools, the increased data work-load was so great that the headteacher agreed to buy in supply cover to prevent the teachers from 'exploding', and at another it increased the assistant head's work-load in turn:

Because there is an increase in the work-load for our teachers, I had a meeting before the holidays and they were about to explode and so then I spoke to the Head and said look we have got to get them some more release time. So then we have put two extra days of supply in so they all get half a day and all of a sudden that is another £380. We don't have to do that you know but we don't want them to go under, so you have to do it for their well-being don't you? (Deputy Head, School D)

So at the moment I have got the two things having to run together, so it has increased my work-load because I have got an additional set of results to input and send off. I have taken it on to try to give the teachers a bit of relief, you know relief to the teachers. (Assistant head School C)

> *Baseline has delayed the start of important curriculum work, the children are unsettled by the constant changes in staff and has had a negative effect on staff well-being as we have all been sleeping badly after administering it. (W, NFER user)*

Thus the increases in work-load has potential financial and organisational effects for the schools, particularly where there were large cohorts:

> *A further time-consuming activity was the need to find ways to use the Baseline Assessment data alongside existing assessment tracking systems. Despite the teachers feeling that Baseline Assessment did not provide them with any new information, and the fact that this was not a requirement, they spent time trying to align their current Developmental Matters baseline with the new baseline. Thus teachers wanted to make the new data useful to them, and they were frustrated that such issues had not been thought through ahead of implementation. (W)*

> *Additional work-load is caused by the mis-match between the Baseline statements and the Ages and Stages we are required to report to our LA. (W)*

> *I was disappointed that the assessments within the baseline do NOT directly correspond with Development Matters. Since that is the document we work with and the childrens' progress is ultimately measured upon, what is the point of an on entry assessment that doesn't directly take skills from this document and instead writes new ones instead? (W)*

> *It would be good if things joined up! That is to say, that if there was a clear set of statements which could be used to track from entry to the ELGs, our life would be easier and it would make better sense (W)*

This frustration was also evident at the case-study schools. This problem, again, reveals the extent to which teachers want Baseline Assessment to be a useful tool for their planning and monitoring, rather than a policy which has no relevance beyond accountability. The time taken on Baseline Assessment and increased work-load make this failure to be useful even more unpopular. In a context where the Secretary of State has committed to reducing teacher work-loads (DfE, 2015b), this finding of an increase in work-load for Reception teachers and other staff is significant.

6.2 Baseline Assessment and Early Years Pedagogy

Early years teachers' professional identities are often based upon their child-centred approaches such as observation and listening, drawn from their training and the wider culture and philosophy of early years. This approach is manifested in the Early Years Foundation Stage (DfE, 2012), which is underpinned by a child-centred sociocultural approach to teaching and learning and places children's play at the centre. Thus for many of the teachers we interviewed, Baseline Assessment was a disruption to their professional culture, and in some cases teachers felt deprofessionalised by collecting data for accountability rather than focusing upon the child's best interests:

> *I am concerned that this will feed into a league table. I feel no longer trusted as a professional. (W)*

> *To be honest a trained chimp could have performed the Baseline Assessments! (W, NFER user)*

One case-study teacher described how her sense of professionalism was affected by having to carry out Baseline Assessment, which took away her enjoyment of settling the children and made her feel 'fake'. She felt guilty about putting aside child-centred principles and instead had to attend to 'this thing hanging over me':

> *If I was sitting in the role play area talking to the children about what they are making and you know engaging with them in that way, I would have to say, 'Oh I have got to go and do some Baseline Assessment', it would make me feel guilty and it would just be this thing hanging over me and by the end of the six weeks I just thought, there it is. I handed it in and it was like enough was enough. (Teacher 3, School C)*

This teacher's experience demonstrates the tension between the need to complete Baseline Assessment for the purposes of school-based performativity and longer-held notions of professional skill. A similar tension was felt by one headteacher:

> *So you know a large part of me cares about the individual child and do the best for them but because of my job I have to think about what the impact is upon the school and how we are going to look compared to the school down the road and how Ofsted are going to judge us and how the Local Authority are going to view us and what our data is going to look like when we are compared with the whole country…where is the child? But that is a very small part because all the SATS results are basically used to punish or to praise the school. The individual child and their performance is really quite low down on everybody's list, which is ridiculous. (Head, School D)*

This headteacher was articulating the tension between the high stakes data needed for accountability and the needs of the child as an individual. Thus Baseline Assessment had an effect not only on the Reception teachers, but also on school leaders' sense of how to manage the new policy.

One aspect of Baseline Assessment which was seen as undermining Reception teachers' professional status was the simplicity of some judgements. This approach was seen as mechanistic in contrast to the complex professional process of an EYFS Profile assessment. For one teacher below, professional dialogue was replaced by 'thirty seconds of gut instinct':

> *As professionals we know there may be more about a child than yes/no. Nothing from the 'testing' told us anything we didn't already know or were learning (but had to rush the child to find out). (W, CEM user)*

> *When the Early Excellence baselines came back the results didn't correlate at all to my own baseline. I've had to justify my results to the head master and I've questioned my own judgements. Surely there are others in my position? (W)*

> *At the end of the year we deliberate quite a lot. […] Whereas for this if you have to think about it longer than thirty seconds you are thinking too much. You have got to go with your gut instinct. (Deputy Head, School D)*

These teachers' professional approach to assessment is being questioned by Baseline Assessment. When teachers' 'gut instinct' was not clear, and they had to reflect on judgments, they were instructed by Early Excellence to 'err on the side of caution', and say no. Here again, teachers' professionalism was undermined as they were steered towards making a negative judgement about a child's potential.

> *If you get to a statement and you are not sure and you are really like discussing it with your colleague or you are thinking, can she or can't she, "err on the side of caution" and say no. We were told that on the training course. So if like with some particular children you are like, 'hmm, can they or can they not?', just "err on the side of caution" (Teacher 3, School D)*

This approach, which can be characterised as 'no reflection, no thinking, no dialogue' undermined early years teachers' sense of professionalism, because it contradicts how they have been trained to operate. As Moss (2014) has argued, there is a risk that early years professionals are reduced to 'grey technicians' through the adoption of crude quantitative approaches. Baseline Assessment replaces diversity, complexity and contingency with the solidity of 'facts' and numbers to demonstrate progress; for the teachers involved, this is a major shift in how they operate and the underlying principles of their profession.

6.3 Baseline Assessment and teacher morale

Many teachers felt frustrated, disrespected and felt their professionalism undermined; they 'lost control' of the data they collected as it was submitted to the providers. They were unsure what would happen to their judgements, which were then returned to them a few weeks later as scores, colour-coded charts and percentages.

> *It does feel a bit odd that you have given something away having no idea of how it is to be used. (Head, School C)*

> *There are so many unanswered questions about what is going to happen, what it is going to look like, what it is going to be used for? (Teacher 1, School A)*

Other teachers mentioned they were uncertain and anxious about the future of the EYFS Profile, and that the lack of training left them burdened and removed the enjoyable aspects of Reception teaching. Thus the policy shifts have caused great anxiety and uncertainty, which leaves teachers feeling powerless and undermined.

Although many of the teachers saw the introduction of Baseline Assessment as further eroding their professional control and autonomy, they 'begrudgingly accepted' it (Selwyn 2015), and were usually prepared to engage in the process and attempt to make it useful for them, rather than resist it entirely. Nonetheless, the teachers were cynical about Baseline Assessment as being yet another 'unnecessary' change in an endless cycle of accountability reform:

> *I can tell you, we head teachers just sighed, we just kind of had a group hug at the meeting, rolled our eyes, and thought here we go again. (Head, School A)*
>
> *In the back of my mind I knew I had to do this Baseline and it wasn't really for my purpose it was for the government, so that kind of made me a bit anti to start with. (Teacher 2, School C)*
>
> *I have always taken the philosophy that as a teacher you know you have to do things you don't necessarily want to do or you might not see a purpose for but it is just one of those things that you have to do. (Teacher 2, School A)*

These teachers and school leaders are inured to the problems caused by continual changes to assessment systems, and therefore simply try to find a way to negotiate a new system. As discussed, the failure of Baseline Assessment to provide any useful information for their planning and the disruption to their professional expertise involved, meant this negotiation was more challenging than with other policy changes. A comment from the survey summed up this feeling of frustration: 'Serves no other purpose than to give the government another tool with which to bash teachers'. (W)

We can conclude from the data discussed in this section that the impact of Baseline Assessment on teachers' roles and professionalism has been significant: their work-loads have increased as a result, they feel frustrated by the problems of accuracy inherent in Baseline Assessment, and they feel deprofessionalised by the imposition of an assessment which challenges their principles and serves no purpose for them.

Section 7
Parents and Baseline Assessment

Parents' views of Baseline Assessment were mixed, although they did have concerns about labelling and the age of the children. There is differing advice from all three providers on what schools should share with parents. Some parents did want to know the 'scores'. However, teachers and headteachers noted the danger of parents wanting to prepare their children for Baseline Assessment and potentially 'widening of the gap'.

7.1 Parents' understanding of the assessment

We interviewed 15 parents in groups about Baseline Assessment, and this was a limited sample in terms of number, their background, gender and ethnic group, due to these being the parents that were available after school and willing to participate. As such, our findings on parents' views must be treated with caution. However, we can identify some themes across the five schools in different areas of the country which are relevant.

The main finding was that most parents were not familiar with Baseline Assessment, and very few understood that the purpose was to measure school performance over time. For some parents, this was a source of anger:

> *Last night [at parents' evening] was the first time I had heard about it. (Parent, School A)*
>
> *I feel a bit bad because my child's Reception teacher said that it was in the news quite a bit that this was the new assessment and I didn't even know anything about it. (Parent, School B)*
>
> *I actually felt angry that there was an assessment of my child being done without my knowledge and that the parents weren't made aware, that we weren't told what the assessment was. (Parent, School C)*

This lack of knowledge is logical given schools' reluctance to inform parents about Baseline Assessment. Given a simple explanation of Baseline Assessment, the parents had a range of views, both positive and negative. Advantages included being able to identify children with particular needs earlier, helping parents to know how to support their child, and that Baseline Assessment might result in additional funding. (These supposed advantages are however contradicted by teachers elsewhere in this report.)

> *This Baseline Assessment will just give teachers a little bit more freedom to pick up on things and discover it before they are into Year 3, 4 and 5 when it is so much harder to deal with. So yes I 100 per cent agree with this whole new system being put in place. (Parent, School E)*
>
> *If somebody is dyslexic that can be picked up a lot earlier rather than later on, so you know if it is on that guideline that would be great. (Parent, School E)*
>
> *It might be helpful because I can always do my bit at home if she needs anything else. (Parent, School A)*
>
> *If you could imagine it could be used to help target additional funding, you know, if we can actually invest a little bit here because they are starting from a lower base, and invest a little less there because that is already a good school. (Parent, School A)*

For these parents, Baseline Assessment was a positive development, particularly as a tool for identification. The teachers (as discussed in Section 2) were dismissive of the idea that Baseline Assessment helped identify additional needs, but clearly parents assumed that the assessment would help the teachers in the short term, rather than merely serving the purposes of accountability.

Parents had a number of concerns about Baseline Assessment, based on how it would affect classroom

practice, and the underlying principles of testing and labelling at a young age. In two of our case-study schools the parents had been informed about the Baseline Assessment and phonics teaching at the same meeting, and this led to concerns that Baseline Assessment might be used for grouping. For others, the age of the children and the difficulty of relating the assessment to Year 6 were the main concerns:

> *All of a sudden it felt like streaming for the phonics and the assessment, and it is obviously a bit daft because obviously they need to be in the right group for phonics and the Baseline Assessment isn't judging but you can't help feeling a little bit sensitive about it. (Parent, School A)*

> *Let that pressure pass by. Because she is four I would much rather almost not get myself into that target environment […] She is only four, so it is ridiculous, I acknowledge that. (Parent School A)*

> *At the same time when they are so young, to me it doesn't matter if my daughter is not reading yet. (Parent, School A)*

> *I am not quite sure on how you then match it against Year 6. It seems such a long span doesn't it, they change so much over the year. (Parent 2, School B)*

7.2 Parents of Summer-Born children

Parents of summer-born children were sensitive to the fact that carrying out Baseline Assessment in the first term of Reception was too early, and parents were also aware of the impact of attending a nursery:

> *I have a younger one, she is one of the younger ones, she is a June baby and for me I do see a difference with my older one who was a January baby. That six months is a big difference. (Parent, School E)*

> *My son he has only just turned four […] He is very young in his year so obviously he is going to be compared with children who are a bit older. It does seem very early. […] He did go to nursery at the school here as well so he was used to some of the routines, but some of the children they may not have had that experience at all … it might be all very new to them so it does seem quite an early time to be doing this. (Parent, School B)*

> *I think the four to five gap is phenomenal. And so maybe assessing those who are all new to it and just four by a few days, maybe against kids who are five in a few days might be extremely different. (Parent, School A)*

For these parents, a system of assessing all children at the same time, despite the significant age gap between children with summer and autumn birthdays, was problematic. As one further parent noted, this is exacerbated by the fact that children are just settling into school: 'measuring it now is a bad idea, because it is such a – everything in their lives has been chucked in the air, and now we are going to measure you!' (Parent, School A)

7.3 Parental anxiety over judgements

Even those parents who were generally positive had doubts about categorising children; nonetheless, they did want to know their child's results:

> *I think as long as whatever is found out is seen in a positive way […] Yes, I would like to know I think but then there is always the danger that children then get put into categories – 'oh he is a bit slow', or 'he is a bit overactive' and 'she is very eager to learn'. (Parent, School B)*

> *They gather all this data about kids and it is one thing if they use it to assess a school's performance and it is another thing if it suddenly ends up being used on their CV one day. (Parent, School A)*

> *When it comes in about an overall evaluation for a child's starting place and maybe even speculation for further potential and then it becomes a benchmark for a qualification? Because*

> *there are so many things that are subjective so I just have more and more questions. (Parent, School C)*

This second quote came from a parent who reflected further on the problems of presenting parents with definitive results on their child, and the potential for this to cause anxiety and anger:

> *It is kind of a double-edged sword because on the one hand I think that the outcomes are that the assessment flags things that I wasn't aware of, and maybe I agree with them, maybe I disagree with them. It is much more helpful for me to have a conversation with his teachers and to see if our observations are in line and then to kind of work together. So something that I am not really part of the equation, I'm just given something, I think it could do a disservice because it could create an anxiety in me or some kind of bias, or I could also say, you know what these teachers or whomever did this test don't know my child at all, and I could be really angry and frustrated. (Parent, School C)*

For some parents, being simply presented with the information about their child's Baseline Assessment would be problematic, and a dialogic approach was preferred. This is important to parents because Baseline Assessment could be seen as a judgement on their parenting, and this is the moment they first deliver their child into the state education system. One parent acknowledged the inevitable pressure parents will face in the build up to Baseline Assessment, saying 'there will be an app on the market within weeks' to assuage middle-class competitive anxiety. For those parents with more awareness of the long-term purpose of Baseline Assessment, there was sympathy with the use of data to measure more accurately the 'value added' by a school:

> *You worry sometimes that the school gets this reputation for achievement that is not necessarily just to do with the staff it is to do with the background or the support that kids are getting out of school. So if you can find a way of measuring what the school adds that's great. (Parent, School A)*

> *Everybody should be held accountable and I appreciate that in theory it is one way to kind of judge what kind of service a school is doing. (Parent, School C)*

For these parents, who are perhaps sceptical about the continued high achievement of some schools in more affluent areas, Baseline Assessment might provide more useful information. However, these same parents also identified the concerns above, about the age of the children and parental preparation, and overall in both their interviews the judgement was that the negative aspects of Baseline Assessment outweighed the advantages.

A final point on the range of parental views is that some were acutely aware of the tensions faced by teachers who had to conduct the assessment at the same time as settling children into school. This parent felt guilty that their child's teacher was having to spend time cuddling their child rather than doing paperwork:

> *In Reception there are still a lot of children who find it difficult to settle in. So my daughter wanted somebody to cuddle when she first comes in when I say goodbye, so I am thinking that the teachers have a pile of papers on their desk waiting and they have to think will I be there for the children or will I be there for the paperwork? So I can see that they get taken away from actually being there with the children when I just want them to play with the children and keep them safe and cuddle them. (Parent, School B)*

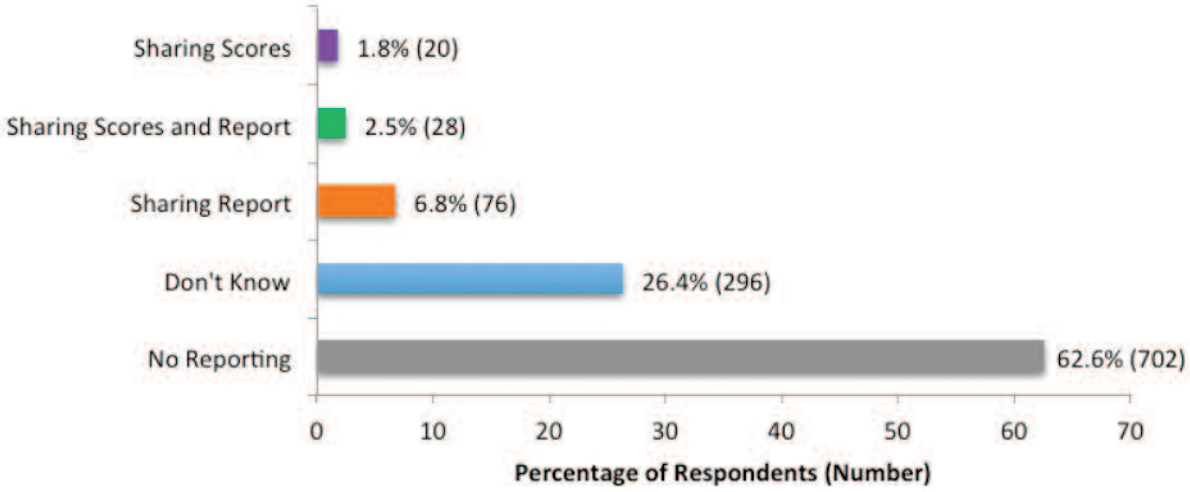
This quote reveals the main priorities of parents of Reception children: that they want the staff to keep them safe and care for their emotional needs, rather than be 'taken away' by assessment. This finding coheres with the teachers' sense of being drawn away from their real purpose by Baseline Assessment.

7.4 Reporting results to parents

There is differing advice from all three providers on what schools should share with parents. Early Excellence advises not to share, whilst the NFER and CEM provide a report to share with parents.

Despite many parents wanting to know their child's Baseline Assessment results, only a very small proportion (11 per cent) of survey respondents had plans to report the results to parents, either in the form of scores, a report or both (see Figure 11). 63 per cent of respondents had no plans to report information to parents, suggesting that schools are reluctant in this first year to involve parents in the assessment.

Figure 11: Reporting Baseline Assessment results to parents



The teachers and school leaders in our case-study schools similarly had no plans to report information from Baseline Assessment to parents. This was facilitated by Early Excellence, who do not stipulate reporting requirements, unlike the individual reports to parents provided by CEM and NFER. Professional organisations such as TACTYC (2015) have noted how sharing results with parents may have a negative impact on teacher-parent relationships and on the children. Similarly, teachers were ambivalent in reporting Baseline Assessment information to parents because of general feelings of uncertainty about it. For some, there was concern that some parents might want to 'prepare' their children for the assessment, thereby creating unnecessary anxiety for both parent and child. There were also concerns that parental preparation practices might 'widen the gap':

> *We could print out an individual thing for parents if we want, but we have yet to decide whether we want to do that. We haven't yet decided because all our assessment policies are all up the spout. (Deputy Head, School D)*

> *I don't think it would be right to introduce it to the parents because that would probably create some anxiety and might feedback to the children because the parents might be saying, 'You are going to go for your test' or whatever. (Assistant Head, School C)*

> *Parents really pushing, pushing, because they don't understand necessarily however much we try and talk to them about it, that might not be the best thing for your child. It is okay – we are teaching them in school, we support them, but enjoy your reading, don't be pushing it and pushing it and pushing it. (Head, School C)*

Reflecting on the 'ridiculous' number of assessments including SATs, phonics and now Baseline Assessment, one headteacher said that she had to remind parents at a recent meeting that the central purpose of the Reception class was to learn to love learning itself:

> *[I said] "Please don't panic, please don't worry on the playground if somebody is saying your child got more yesses on their Baseline Assessment than yours. Let's love learning, let's love reading, let's remember they are young". (Head, School A)*

Despite this reluctance to inform parents, teachers and school leaders were aware that it might become necessary in future to do so. The Deputy Head of a case-study school predicted that the data might be used

to generate web based 'league tables', which would have the effect of ranking and comparing Baseline Assessment between schools, so that parents might use it in making judgements about their choice of school. This teacher was concerned that such comparable data would simply reflect the school's socio-economic catchment area:

> *So potentially you could have a league table of Baseline Assessment information for parents. Which shows them that the schools, potentially if over time the Baselines are pretty consistent well actually it is a reflection of your intake isn't it? So it might then reflect parental judgement about where they want to go. (Deputy Head, School D)*

As we discuss in the next section, concern over the future use of results caused great anxiety for school leaders; the use of information by parents was a key part of this.

Overall, we can identify a mismatch between what baseline is for, that is, a cohort measure for school accountability and what parents want from assessment. This tension reflects the schools' views of the problematic nature of Baseline Assessment.

Section 8

Selection, cost and the role of private companies

School leaders reported that the process of choosing a Baseline, though welcomed, was not simple, and there were criticisms of all three providers in the survey. There were also concerns about the use of private companies, and the cost to schools of administering the Baseline.

8.1 Choice of Baseline Assessment provider

At the case-study schools, Baseline providers were most commonly selected by the early years teachers themselves, in consultation with the headteacher. All five schools used the Early Excellence Baseline Assessment (EExBA), in line with national trends, often because of a perception that this provider was more in keeping with the ethos of early years. Similar comments were made on the survey:

> *Our response then as responsible practitioners for young people is that we must choose the one that we think is better in terms of application and rounded in terms of practice, and my understanding is that is the one that is least linked to tests and is more about practitioner type judgement. (Head, School B)*

> *For me it wasn't really much of a decision because we wanted to go with the EExBA because it was the ethos that we already follow. It wasn't a particular change from practice that we already do. (Assistant Head, School C)*

> *We chose EExBA to use and it fits the usual EYFS statements. The other baselines I viewed would not have done so in my opinion. (W)*

> *EExBA was chosen as it is reflective of our existing Baseline Assessment methods – in line with EYFS principles, although school based assessment reflects ALL profile areas, not those limited by the Government. (W)*

> *I am happy I chose EExBA as we were able to decide the most appropriate activities and observations in order to gather information towards the statements. (W)*

There were however some tensions between headteachers' and early years teachers' priorities in choosing a Baseline Assessment provider, and some more general doubts about EExBA. At School A, the head wanted to choose an alternative provider on the basis that it would minimise disruption:

> *[We] selected another provider because we saw it as quick, quite simple. It was a kind of a little bit of a laptop based […] It was click on a few buttons, it was very oral, it had been devised by Speech and Language Therapists for young children. It wasn't going to damage anyone and it was going to be fairly quick and fun to do. Because we were typically thinking 'Why are we doing this?' (Head, School A)*

This headteacher went on to explain her rationale further: she commented that since they were already doing a baseline and the results were only for the DfE, she thought 'let's do it in the easiest, less onerous, cost effective, way of doing it'. However, following a presentation by the local authority which heavily promoted Early Excellence, she felt under pressure to choose EE:

> *By that time busy, busy heads, early years co-ordinators who like to make decisions in groups … The wave was, you know there was this wave of early years, Early Excellence, and you almost felt like if you were going with somebody else you know you were a traitor to good early years pedagogy. (Head, School A)*

This headteacher decided to stick with her earlier choice, however, only to find that the provider had been removed; then, she commented, 'I was absolutely exhausted by it all but I just thought fair enough we will just do the Early Excellence'.

At School D, a similar tension arose between the headteacher's desire to minimise disruption and the early years teachers' need for something 'meaningful':

> *So as you know we were given six choices. I will be honest I wanted to go for the computerised one, it was pretty quick. […] The Reception team felt that if they were going to do it they wanted to do something that was meaningful to them. Which I can fully understand, so they have gone with the system that was recommended by early years as being a meaningful assessment. (Head, School D)*

These tensions were apparent in some survey comments, such as one teacher who noted: 'In schools where headteachers are not as clued into reception I fear a test will be chosen as it appears easier to somebody who doesn't truly know how to work with and assess this unique age group' (W).

At School E, the headteacher reflected on the topic of having choice, and described himself as 'cynical' about it:

> *I imagine their intention is to make us feel we have got some control over it and some sort of ownership. And I guess if it goes the way of Early Excellence which is the one we feel is quite good then we will be quite happy. […] But if they go with another one then we won't be quite so happy. […] I think people have appreciated the ability to choose one system and try what they think is the best way to match our own practice. (Head, School E)*

Choice is therefore not a simple advantage to the current system of Baseline Assessment, because the range of providers differ in how they are perceived. Although Early Excellence appears to have a monopoly on the position of 'early years friendly' Baseline Assessment, their selection is not always straightforward. There are real tensions between different staff in terms of priorities; for some headteachers, damage limitation and speed are more important than finding an assessment which coheres with other early years practice. If the choice of providers were to be taken away, however (or Early Excellence removed as an option) there would be considerable opposition from these schools.

The comments from School A and School D reflect the problematic nature of an assessment which only serves the purpose of accountability, rather than providing useful information for teachers. The selection of Early Excellence as more 'early yearsy', as they have promoted themselves, has proved extremely popular, even though the results are used for the same purpose as the other assessments from CEM and NFER. This depiction of Early Excellence obscures the fact that it also produces numerical scores for children which will be used to measure progress. It may appear more useful or 'meaningful' in that the format is familiar to teachers, but is not actually intended to be a guide for planning. There are also limits to its usefulness, as discussed in other sections, because it is not consistent with the EYFS Profile or compatible with data tracking systems, and does not provide more information than existing baselines.

Finally, in relation to survey respondents who chose an alternative provider, we have limited information but there appear to be a number of motivations, including coherence with existing baseline systems and not needing any moderation. One respondent commented:

> *We did 2 baselines CEM and NFER. CEM was simpler to administer and more child friendly. The tasks for NFER were very time consuming and they were not adaptive enough for the less able children. It seemed like a test, whereas CEM did not. (W)*

There were several survey comments on the other two providers, which were in general negative. The dominance of Early Excellence in press reports and its effective promotion as coherent with early years practice may have contributed to this. The problems listed were very similar to those experienced with Early Excellence, including the time taken, the lack of new information, problems with accuracy and the inappropriate content. A small sample is included here:

We chose NFER as we felt on balance it would be the least disruptive. We have continued with our usual assessment and collection of data. It has been a waste of time, money and resources. (W, NFER user)

The compulsory and rigid nature of the NFER Baseline Assessment did not tell me anything I would not have assessed anyway and was restrictive. It took a long time out of my teaching time and was open to lucky guesses (particularly in the maths part). Due to having to do one at a time, some children were assessed in the morning and others at the end of the day when they were tired and easily distracted. Observational assessment would have been much more accurate and "natural". (W, NFER user)

Although I agree with baseline in principle, the NFER package prevented us from building relationships with children at a crucial stage and the literacy tasks in particular were wholly inappropriate. (W, NFER user)

We have not been happy with our CEM experience and it did not tell us anything we did not already know. (W, CEM user)

The CEM base test asked many ridiculous questions that a Reception child wouldn't even be expected to know. (W, CEM user)

Our existing methods of baseline are both holistic and far more detailed than this CEM baseline. (W, CEM user)

The CEM one is inaccurate and so the data we have received as a result is unreliable. Children were able to score for areas I know they did not know. (W, CEM user)

To be fair all children enjoyed coming to play on the computer (as they saw it) and many children asked if they could have more than one go. The results were interesting and broadly in line with our judgements for most pupils. Our biggest gripe was the technical issues with the CEM system as that needed a lot of time to send off data. (W, CEM user)

CEM doesn't take into account anything other than literacy and maths. What happened to "the unique child"? (W, CEM user)

These comments reveal the complexity of a system with a choice of providers; different issues with accuracy and impact exist depending on the provider chosen.

8.2 Costs

In the interviews, headteachers commented on the extra costs involved in Baseline Assessment, beyond what is provided by the DfE, mainly relating to training and cover to allow teachers to moderate and input data. This was also highlighted in the survey comments, which also mentioned the relationship between choice of provider and cost:

Yes I think we have spent £420 on a one day training course for our Early Years Lead teacher. I literally couldn't afford to send all three. So she has been on the course and rolled it out and I think there is another training event for headteachers and co-ordinators to go along – that will have a cost and then the actual materials. (Head, School A)

We have had to get some supply in to release them to actually moderate and input data. So I think it has cost us about possibly, off the top of my head, probably about four days of supply. So that is the best part of £800. In a school this size with a budget that is not a criminal loss but it is obviously £800 that could be spent otherwise. (Head, School D)

It has taken a ridiculous amount of time to complete – and also has cost the school a lot of money in terms of getting supply teachers to cover classes whilst the Baseline assessment was undertaken. This is not value for money for schools – even if the government paid for the test itself. (W, NFER user)

Moreover, as one head pointed out, there is always a cost of additional assessment at some point:

> *All these systems cost money to administer don't they? Who picks up those costs, whether it is top sliced from our budget before it comes through to us or whether it is a learning bill to us I'm not sure. But obviously there is a cost to the tax payer that has to be met somewhere. (Head, School B)*

The cost of training appeared to be a particular issue, which resulted in many Reception teachers receiving training only from the one teacher who had attended the course, leading to further confusion over how to assess the children. Further training on analysing the data was also offered to schools, so that they could interpret their own information once it had been processed by Early Excellence. There were also concerns about the costs of additional materials, and the fact that those who went on EExBA training were only allowed one manual, which they were instructed not to photocopy. These problems of cost seemed to be a particular issue at a time when budgets are being reduced in general at school and local authority levels, and new assessment systems ('assessment without levels') are being implemented elsewhere in primary schools (DfE, 2015d):

> *The financial cost to schools in terms of assessment without levels-assessment with levels cost me £800 a year, assessment without levels costs me £4,000 to £5,000 a year because of what I have had to buy and commission. […] You are paying the private sector for the joy of delivering your own assessment. Because the government has said that it is all in schools' hands and the local authority haven't got the capacity to do it anymore because they have been stripped of their funds. (Head, School A)*

> *It has cost my school additional money at a time when budgets are getting tight! The extra work has had an impact on my team as a whole to give us no extra information than we would otherwise gain from our own practice. (W)*

Again, the issue of usefulness was a key part of establishing if Baseline Assessment provided value for money. One survey respondent commented, for example, 'This Baseline Assessment has told us no more than the profile and our own assessment systems. So, a waste of money'. This seems particularly unacceptable to teachers and school leaders because other budgets are being reduced, including local authorities', and other policy changes have created additional costs related to assessment. This perception that Baseline Assessment is a 'waste of money' is an important part of the devaluing of the policy among school staff.

8.3 Private companies

A further issue which damaged the credibility of Baseline Assessment for many respondents was the use of private providers. Headteachers and some teachers commented on the 'sales pitches' they were now subject to, and the value for money of using outside companies. There were also some doubts expressed about the expertise of providers:

> *A lot of email contact – we get bombarded with things […] I do really object actually to the way anyone latches on and sends you more and more stuff. (Head, School E)*

> *It would be relatively simple to provide accurate teacher judgements after the children have been in school for a full half term […] This would not require paying companies that have very different assessments monies that could go directly into schools to benefit children in need of extra support. (W)*

> *So EExBA have tried to give us some examples, but who is to say that they are interpreting this the way the DfE want it interpreting? [The DfE are] not taking responsibility for this but passing it out to a private company and then affecting what your teachers are going to do, because this private company is – who are these people – they are affecting the way that your highly trained experienced professionals are performing. (Head, School A)*

As we see in this final comment, for some the use of private providers is seen as an abdication of responsibility from the DfE. There are also related concerns here about accuracy and how well the providers' assessment

cohere with the Government's aims. There were also several comments in the interviews and survey about the problems associated with having different providers, particularly the lack of consistency, which built on doubts about accuracy more generally:

> *It would be better if all schools completed the same assessment. (W, CEM user)*
>
> *I think our comments would differ significantly if we had chosen a different Baseline Assessment. (W)*
>
> *Needs to be standardised, not four different ones producing different data in different areas!! (W)*

Within this context of doubt about the effectiveness of Baseline Assessment, some headteachers commented that the combination of policy changes of 'assessment without levels' (DfE, 2015d) and Baseline Assessment left them vulnerable to the promotional activities of private companies:

> *If I am sitting here as a headteacher with over 20 years' experience saying 'I don't really know what to do' somebody – other people – lots of people, even I am vulnerable for somebody to come in and say, 'I can solve this problem for you' […] I think it creates even more scope for people to actually make a buck out of this whole thing if I'm honest. (Head, School D)*
>
> *I feel extremely concerned about the increased use of private companies in the entire education world, but also within assessment at the moment. I feel that by removing and taking away all the known assessments and I am not saying they were great, but taking away all of those, what has happened is we have been opened up to a completely free market and we are being bombarded with sales pitches. And actually that is very hard when what you are trying to do is focus in on what you are doing for children. […] I think that the companies at the moment can really capitalise on the fear factor in schools and with headteachers and it is not healthy really. (Head, School C)*

For these headteachers, whose concerns were echoed by others also, Baseline Assessment is part of an assemblage of policy change which has left them vulnerable to the private sector.

Section 9
The wider policy context

This section explores the wider policy context where guidance on assessment to primary schools has been reduced. There were significant concerns raised about the use of the Baseline Assessment data by schools themselves and others, and particularly its use within measures of school performance.

9.1 Assessment without Levels

The introduction of Baseline Assessment came at a time when there was already a great deal of policy change related to assessment in primary education. The policy of 'assessment without levels', announced under the Coalition Government in 2013 and implemented in 2015, removed the familiarity of national curriculum levels. For some of the headteachers we interviewed, this combination of removing an established system and introducing the uncertainties of Baseline Assessment was contradictory. As mentioned earlier in relation to private companies, this was seen as leaving headteachers vulnerable to those who wanted to 'capitalise on the fear factor in schools and with headteachers' (Head, School C).

> *I don't think I have ever come across a situation where heads feel so at sea. You know when levels came in nobody, if I am honest, nobody really liked it but having had the system for so many years and everybody knowing what it means, every teacher knows what it means, we are now back in a situation where nobody really knows what they are doing. (Head, School D)*

Overall, there was a sense that assessment was an area of real risk and worry for school leaders:

> *It just feels like we have got a whole assessment pit at the moment that nobody is quite sure how we are all going to come out at the end of this year really. What that is going to look like, which is not good. [The situation] is not good. Especially when the stakes are so high and you are looking at how we are judged and all the Ofsted… (Head, School C)*

> *As you know we were given six choices […] So I guess I went into headteacher defence mode and said we will do the quickest system that tells the least so then whoever is here in seven years will be least punished by it. […] We are bound down and broken by those judgements and the way people view us. (Head, School D)*

As one headteacher commented, assessment policy appeared to serve the purposes of showing improvement for Government:

> *I think there is a huge drive politically to show that education is moving forward. And there will be I'm sure at the next election a very clear drive to show that standards have been pushed up. And Baseline Assessment is one measure I presume of doing so…. we are a system under cosh in that respect and there is an awful lot of weighing of the pig isn't there and measuring it? (Head, School B)*

This headteacher refers to the saying that weighing the pig doesn't fatten it; in other words that assessing children doesn't improve their learning. Similarly, another headteacher commented 'This is an equation that testing leads to better teaching' (Head, School E). For all these headteachers, there is a feeling that the situation is damaging – they feel 'punished', 'under the cosh', 'bound down and broken', 'at sea' and in a 'pit'. This emotive language reveals the affective impact of assessment policy, which is seen as a tool with which to criticise schools; one headteacher referred to Baseline Assessment as 'giving us enough rope to hang ourselves' (School C). This description of intense pressure chimes with findings from recent research on other sectors of education and accountability, which described schools as 'exam factories' (Hutchings, 2015).

The uncertainties of the future leave the headteachers feeling anxious and lost, particularly at a time when their support from the local authority has been reduced:

> *Once upon a time something new like this would have come in and the local authority would have solved the problem for us and they would have got us all together and said 'This is what's what'. Because they had the staff, they had the personnel, they had the expertise. They would have solved that problem for us and said 'This is what we are going to do as a Local Authority. It doesn't matter what the one next door is going to do; this is what we are going to do'. That doesn't happen now because the local authority has been so cut back that there isn't the level of expertise or people with the time to actually solve these problems so it is down to us. (Head, School D)*

This wider move of reducing LA funding and therefore support for schools exacerbates the problems of new policies by leaving individual schools to find appropriate solutions on their own. School leaders' response is to be defensive and reduce the threat as much as possible, as the quote above suggests, by finding the solution which will leave them 'least punished', thus limiting the potential damage caused. Here Baseline Assessment is part of an assemblage of policy moves which have left school leaders feeling vulnerable, worn down and uncertain.

9.2 Use of Baseline results as part of a high stakes accountability system

There were significant doubts about the use of Baseline Assessment as a measure of schools' performance: only 6.7 per cent of respondents to the survey agreed it was 'a good way to measure schools' performance'. First and foremost, teachers and school leaders doubted the underlying principle of measuring progress from Reception to Year 6:

> *I don't think you should [use it to measure progress], I don't think you can, because they are children and they are not robots, not machines, they are children. You don't know what influences they have got from outside, what is going to happen in those seven years, so I think it is ridiculous. (Teacher 3, School C)*

> *I am very concerned that it will be used as a way to predict the children's KS2 attainment and assess their progress. I feel that is a great pressure on us. (W)*

> *So much can happen in a child's life between EYFS and the end of KS2. (W)*

> *You can't basically put a kid in a box when they are four and say because you did quite well on your Baseline you will be a high achiever when you are eleven. It is too young to say. Or because you didn't do so well on your Baseline that's it, you are disadvantaged, you are not going to achieve. (School A, Head)*

> *It is damaging that this score will stay with the 4 year olds for the rest of their primary schooling. It labels the pupils and also creates pressure on staff to push children who aren't ready/are plateauing which is normal in academic progress. It will lead to inflation of scores further along the line. (W)*

Here we see concern that outside influences will render the data inaccurate as a measure of the 'value added' by a school, and the dangers of labelling children and creating pressure on schools. There were also comments which challenged the idea that any method of tracking progress from age four to eleven was worthwhile, and which questioned where the policy had emerged from in terms of research:

> *Ofsted have never judged progress from Reception to Key Stage 1 or from Reception to Key Stage 2, because all the research showed that there was no correlation between a very young child's development and attainment at age five to where they would be at age seven or age eleven…All of a sudden that has changed. […] So where has that research come from? Why has somebody changed their mind? Was it the DfE? (Head, School A)*

> *Not long ago the idea of any kind of formal assessment for Reception age children was considered unacceptable. How can it suddenly be ok? (W)*

> *Any educator knows, any parent knows, young people's journeys from the age of four and five can have huge ups and downs … if we have a Baseline Assessment that it is looking at, or attempting to look holistically as the skills of a child and we link that to a Key Stage 2 assessment that is looking at SPAG, reading, writing and maths, how do the two correlate? […] It's a bit bonkers. (Head, School B)*

> *We don't learn in a linear way. Learning is stepped. You know we learn a bit, we stop, we learn a bit, we stop. So it doesn't necessarily mean that these children are all going to be in a linear line, the same as they are now when they get to Year 6 […] You are sort of setting from now a goal that they are going to achieve in Year 6, I think that is a bit harsh actually. (Teacher 2, School E)*

Baseline attempts to construct a linear relationship for progress from age four to age eleven, even though the content of the assessments are different; this was seen as a major flaw in the system. School leaders were concerned that they were now being unambiguously told that there was such a relationship between the early years and Year 6 and they questioned why this had suddenly become possible. Furthermore, the variation between children and their rates of progress meant that a simple correlation between Reception and Key Stage 2 was impossible:

> *Children's progress is going to be judged against how far they have gone in seven years. Now to my mind that is an almost impossible thing to do because you can't test children at 11 about the same things you were testing them at four. It just doesn't make sense. (School D, Head)*

Because of these doubts about the principle of tracking and prediction from Baseline Assessment data, there were also concerns raised about the future use of data by Ofsted and, as discussed in an earlier section, the publication of the data online and comparisons with other schools by parents. These developments were seen as inevitable: as one survey comment put it: 'It is certain to be used as yet another stick with which to beat schools by the likes of Ofsted'.

9.3 'Problems of low scoring'

One of the concerns raised in the DfE's own research into Baseline Assessment, before the 2015 cycle began, related to the temptation to deliberately deflate results in order to increase progress measures, known as 'gaming' (DfE, 2015c) which is a direct result of the pressure of the 'high stakes' accountability environment. It is important to note that any assessment including the test-based forms provided by CEM and NFER is open to some manipulation as all three providers have some elements of subjective teacher assessment.

Although we did not ask directly about this in the survey, there were some comments on this issue, which echoed some wider tensions between headteachers and Reception teachers:

> *This pilot has shown that individuals have manipulated scores in their favour, for example, marking children lower than they are so that more progress will be shown. (W)*

> *Schools want their baseline scores to be low in order to maximise the progress they can show. This means there should be some moderation in place in order to ensure schools are being accurate with the data they submit. Headteachers' wishes for low baseline scores also means that we are beginning the year looking for the negatives in children-what they can't do and how low they can be scored in order to make our scores low. This is the very antithesis of the philosophy behind the EYFS, which makes me very uneasy about completing such an assessment. (W)*

> *I agree that there needs to be standardisation as some schools inaccurately score children as significantly below on entry to inflate progress but a test is not the answer. (W)*

In the interviews school leaders similarly recognised the possibility and pressure to deflate results:

> *[At the meeting about providers] there was a sort of tongue in cheek response from headteachers that if it is going to be linked to progress perhaps we should choose the most difficult one. (Head, School B)*

> *Obviously you are not going to shoot yourself in the foot, okay, if you are assessing a child with Baseline and it is a simple yes or no and we all know that there is a lot of grey in the middle of yes and no. Are you really going to tick the yes box if you are in any doubt? You are going to tick the no box aren't you? (Head, School A)*

Nonetheless, respondents claimed that they had not taken this approach for a range of reasons, including the moral imperative to 'get it right', the resulting uselessness of the data, and the danger of 'coming unstuck' when compared to other similar schools:

> *You have got to be honest, again you will really shoot yourself in the foot if you are working in a very leafy lane middle-class school with highly educated parents and the children are coming in, you know broadly in line with national averages or slightly above and suddenly you are publishing data that you would expect to see in a slightly different type of school. (Head, School A)*

> *Although we are bound down and broken by those judgements and the way people view us we believe in what we do and hopefully we will get good results by doing what we do, which we think is right, so we don't play those games. I know a lot of schools are forced to do that, particularly schools that are in trouble. (Head, School D)*

> *[There is] this tongue in cheek idea of headteachers saying lets choose the most difficult system that the children are more likely to fail on. So you do have those moments cross your mind because there is the pressure there, of course there is pressure there, it is hugely high stakes and yet you have to do this job with a very strong moral compass. (Head, School B)*

Discussions about the pressure and high stakes accountability environment were more prevalent in those schools which were visited later in the process and which had access to the Early Excellence national results, which showed almost 50 per cent of children had been graded as 'below typical' or 'well below typical' nationally. These results may be due to the instruction from Early Excellence to 'err on the side of caution' which was mentioned by many of the teachers interviewed:

> *But it does seem like people pushing results down in order to maximise progress. Which is obviously a temptation. [But] if you suppress them, push them down a little bit then when Ofsted come and look at it and say, well this is rubbish for this sort of school, they can see the type of school, the type of area. So then they will say that your assessment isn't valid really. (Head, School E)*

> *It is all this pressure going on instead of just saying, the children are here and they are going to make the progress that is expected or slightly better. But everybody is so stressed about it all and so people start erring on the side of caution because they are worried. (Teacher 2, School E)*

> *They said if you are umming and ahing and if you are not sure it is a no. (Teacher 1, School E)*

> *I have said to the staff, say no [if you're not sure], because in one sense it is – I feel it is better to say no, because when you are then looking at that as part of assessment to inform your planning, it is better to say no and go over something again than say yes and shelve it and move on when the children might not be there yet. (Assistant Head, School C)*

This practice of always saying 'no' if there is any doubt a child can achieve a point facilitates the deliberate deflation of results, and even legitimises the low scores that schools have produced with EExBA. Further evidence of this was seen in this headteacher's comments:

> *[Teachers] may feel that they are under pressure for the children to appear more able and developed than they actually are, because there is that natural thing to want to do well in assessments and tests isn't there? We don't normally go into assessments and tests wanting to fail them, we normally want them to go in to shine and to celebrate where we are at. What I have explained is that if it is as crude as 'yes you can or no you can't' and you would only say yes you can if that child actually could. If they nearly can you say no, that is shooting yourself*

> *in the foot. Because we are all very positive people and I don't know, if one of the statements is; Can recognise all letters of the alphabet, and you sit down with a child and they miss one, there is this natural feeling to say, 'Oh yes they can'. But actually that is shooting yourself in the foot, because technically they can't. (Head, School A)*

This problem of low scoring within a high stakes accountability culture is not unique to Baseline Assessment; similar tactical responses have been found in relation to the EYFS Profile (Bradbury, 2013). Within a culture of high stakes accountability, Baseline Assessment, with its primary role being as an accountability measure, is particularly vulnerable to low-scoring practices. This undermines the accuracy and credibility of the assessment considerably, even if schools do not deliberately deflate results.

For one headteacher, the implications of the problem of low scoring due to high stakes accountability could be important in the future of the policy:

> *Well it worries us slightly but on the other hand it runs the great risk of the whole thing being rubbished by Nicky Morgan or whoever, who is going to say look this isn't worth the paper it is written on because clearly schools are not assessing it properly. Or they are going to bring in some sort of strict test regime because they will say this is too subjective, because that could be the criticism of the Early Excellence one, it is fairly subjective, it relies entirely on the professionalism and the viewpoint of individual practitioners. So she might want to bring in something that we don't really agree with, more of a yes/no test type of thing because she feels that she can't trust people to do it honestly. (Head, School E)*

This concern is in keeping with the feeling that Baseline Assessment is another 'tool to bash teachers' discussed in earlier sections; for many teachers, Baseline Assessment is part of a game which teachers cannot win, whichever provider they pick and whatever tactical approach they take. The issue of trust in observational judgements is clearly an important part of any future discussion about Baseline Assessment as a policy, as the majority of schools chose Early Excellence. However, it is important to note that any assessment including, test-based forms, is open to some low scoring, and as the previous section demonstrates, Baseline Assessment is seen as having further underlying flaws.

Section 10
Conclusion

This research has reinforced the view that Baseline Assessment is a controversial policy which lacks the support of teachers and school leaders and has serious limitations in terms of measuring school performance.

Only 7.7 per cent of survey respondents agreed that Baseline Assessment was an 'accurate and fair way to assess children' and only 6.7 per cent agreed it was 'a good way to assess how primary schools perform'. As a teacher comments in Section 9, Baseline Assessment is problematic as a progress measure 'because they are children and they are not robots, not machines, they are children'.

Many teachers doubt the use of measuring progress from Reception to Year 6 given the problems of assessing accurately at age four and the variability of children's patterns of progress and development. These quotes from the survey illustrate some of these concerns:

> *The data is skewed and it didn't give an accurate result of children's ability on entry. (W, CEM user)*
>
> *Baseline assessment is the biggest farce I have undertaken during my entire teaching career, the potential for children to guess at answers or to misinterpret things is too vast to give a clear representation of where children are at. (W, NFER user)*
>
> *Appalling form of unnecessary assessment. Goes against the principles of ethical and purposeful assessment in the EYFS. Serves no other purpose than to give the government another tool with which to bash teachers. Why change the system of EYFS Profile when it was perfectly adequate? I can see how this is already damaging to teachers and children and this kind of poorly considered policy makes me want to leave the profession. (W)*
>
> *I find the awarding of a point score to four and five-year-olds unhelpful and fail to see how it can be an accurate predictor of future performance. I worry it will lead to sample tests being produced which will encourage parents to drill children. I see it as another assault on a play based curriculum. (W)*

For many teachers, Baseline Assessment has had a negative impact on their working lives without benefitting the children they teach. It goes against the principles of good teaching in early years, and at the same time does not assess accurately enough to form the basis of a school performance measure.

Importantly, it may be impossible to produce any kind of assessment which provides this kind of baseline with children of this age.

We can conclude from this research project that Baseline Assessment is problematic at best, and potentially damaging at worst. We summarise our key findings below.

1. **Baseline Assessment is inaccurate and therefore problematic as the basis for school accountability.**

   Teachers and headteachers see all three Baseline Assessment providers as inaccurate, unreliable and lacking in validity and hence its ability to accurately measure the 'value added' by schools. Given the inaccuracy of the Baseline Assessments, teachers and headteachers have serious reservations regarding its use in predicting later assessments.

2. **Baseline Assessment has potentially damaging effects on children relating to low expectations and labelling.**

   There is a serious danger of under-assessing young children's abilities, particularly for groups such as EAL and summer-born children, and setting low expectations which in turn could have an impact on their future educational trajectories.

3. **Baseline Assessment increases teachers' work-loads without providing useful information.**

   Baseline Assessment does not have the support of teachers as it increases their work-loads, does not provide useful information, and has negative impacts on relationship building in the first term of Reception. For some teachers, Baseline Assessment damages the key emotional work they do in the first term of Reception and positions them as data collectors, damaging their feelings of professionalism. It serves little purpose in terms of identifying children's needs and is not an improvement on schools' existing forms of on-entry assessment. The EYFS Profile, in contrast, has the support of many teachers.

4. **Baseline Assessment has cost and resource implications for schools.**

   Teachers and headteachers were very concerned about the increased assessment costs of Baseline Assessment. They resented paying private providers for the training and analysis and the associated costs of teacher supply cover.

Adams, D. (2015) 'Schools opt to replace 'test too far' for reception children'. Guardian, 9 June 2015, Available at: http://www.theguardian.com/teacher-network/2015/jun/09/schools-replace-test-too-far-recpetion-class [Accessed 14 December 2015)

ATL (2015) 'Baseline Assessment', Available at: https://www.atl.org.uk/policy-and-campaigns/campaigns/baseline-assessment.asp [Accessed 14 December 2015]

Bodrova, E., & Leong, D. J. (2007) *Tools of the mind: The Vygotskian approach to early childhood education (2nd Ed)*. Upper Saddle River, New Jersey: Prentice-Hall.

Bradbury, A. (2011) 'Equity, ethnicity and the hidden dangers of "contextual" measures of school performance', Race, Ethnicity and Education, 14 (3): 277-291

Bradbury, A. (2013) *Understanding Early Years Inequality: Policy, assessment and young children's identities*, London: Routledge

Bradbury, A. (2014) "'Slimmed down' assessment, or increased accountability? Teachers, elections and UK government assessment policy", *Oxford Review of Education*, 40(5): 610-627

Bradbury, A. and Roberts-Holmes, G. (2014) " The 'Datafication' of Early Years Settings", British Education Research Association Conference, Institute of Education, September 2014

Braun, A., M. Maguire and S. Ball (2012), *How Schools Do Policy: Policy Enactments in Secondary Schools*, London: Routledge

Broadhead, P. (2006) 'Developing an understanding of young children's learning through play: The place of observation, interaction and reflection' *British Educational Research Journal 32*, no. 2: 191_207.

Carr, M. & Lee, W (2012) *Learning stories: Constructing learner identities in early education* London: Sage

CEM (2015) 'Introducing BASE', Available at: http://www.cem.org/reception-baseline-assessment [Accessed 14 December 2015]

Department for Education (DfE, 2014) "Reception baseline: approval process for assessments". Available at: https://www.gov.uk/guidance/reception-baseline-approval-process-for-assessments

Department for Education (DfE, 2014) "Reforming assessment and accountability for primary schools". Available at: https,//www.gov.uk/government/uploads/system/uploads/attachment_data/file/297595/Primary_Accountability_and_Assessment_Consultation_Response.pdf (accessed 10 March 2014).

Department for Education (DfE, 2013) 'Primary assessment and accountability under the new national curriculum' Available at: https://www.gov.uk/government/consultations/new-national-curriculum-primary-assessment-and-accountability [Last accessed 24 November 2015]

Department for Education (DfE, 2015a) 'Policy paper: 2010 to 2015 government policy: school and college funding and accountability' Available at: https://www.gov.uk/government/publications/2010-to-2015-government-policy-school-and-college-funding-and-accountability/2010-to-2015-government-policy-school-and-college-funding-and-accountability#appendix-2-reception-baseline-assessment [Accessed 24 November 2015]

Department for Education (DfE, 2015b) 'Policy paper: Reducing teachers' workload', Available at: https://www.gov.uk/government/publications/reducing-teachers-workload [Accessed 10 December 2015]

Department for Education (DfE, 2015c) 'Reception baseline research: views of teachers, school leaders, parents and carers'. Available at https://www.gov.uk/government/publications/reception-baseline-research [Accessed 24 July 2015].

Department for Education (DfE, 2015d) 'Speech by Nick Gibb: Assessment after levels'. Available at: https://www.gov.uk/government/speeches/assessment-after-levels [Accessed 10 December 2015]

Department for Education (DfE, 2015e) 'Reception baseline assessment: guide to signing up your school'. Available at https://www.gov.uk/guidance/reception-baseline-assessment-guide-to-signing-up-your-school

Department for Education and Skills (DfES). 2007. Early Years Foundation Stage: setting the standards for learning, development and care for children from birth to five. Nottingham: DfES.

Department for Education (DfE, 2012) Statutory Framework for the Early Years Foundation Stage: setting the standards for learning, development and care for children from birth to five. (Last Accessed 24th November 2015) http://www.foundationyears.org.uk/files/2014/05/eyfs_statutory_framework_march_2012.pdf

References

Early Excellence (2015) 'Reception Baseline Assessment', Available at: http://earlyexcellence.com/eexba/ [Accessed 14 December 2015]

Fleer, M. (2010). *Early learning and development: Cultural-historical concepts in play*. Cambridge: Cambridge University Press.

Fleer, M. and Richardson, C. (2009) 'Mapping the Transformation of Understanding' in Anning, A. Cunning, J. and Fleer, M. (Eds) *Early Childhood Education: Society and Culture* 2nd Edition. London: Sage Publications.

Gaunt, C. (2013) 'Ofsted chief calls for review of early years assessment' Available at: http://www.nurseryworld.co.uk/article/1187290/ofsted-chief-calls-review-early-years-assessment

Griffiths, S. (2015) 'Tories in school tests U-turn', Sunday Times 1 November 2015. Available at: http://www.thesundaytimes.co.uk/sto/news/uk_news/Education/article1627329.ece [Accessed 11 December 2015]

Hutchings, M. (2015) 'Exam Factories? The impact of accountability measures on children and young people', National Union of Teachers, Available at: http://www.teachers.org.uk/node/24299 [Accessed 14 December 2015]

Mansell, W. (2015) 'Reliability of new baseline assessments'. Available at: http://www.naht.org.uk/welcome/news-and-media/blogs/warwick-mansell/reliability-of-new-baseline-assessments/?lister625p=6

Moss, P. (2014) Transformative Change and Real Utopias in Early Childhood Education, A story of democracy, experimentation and potentiality. (Oxon, Routledge).

NFER (2015) 'NFER Reception Baseline Assessment', Available at: http://www.nfer.ac.uk/schools/baseline-assessment/baseline-assessment_home.cfm [Accessed 14 December 2015]

NUT (2015) 'Better without Baseline', Available at: http://www.teachers.org.uk/baseline [Accessed 14 December 2015]

Reclaiming Schools (2015), 'A few anomalies? No, baseline is flawed from start to finish' Available at: http://reclaimingschools.org/2015/11/20/a-few-anomalies-no-baseline-is-flawed-from-start-to-finish/ [Last accessed 30 November 2015]

Roberts-Holmes, G.P. & Bradbury, A. (in press) "Governance, Accountability and the 'Datafication' of the early years in England' *British Educational Research Journal*

Schoolsweek (2015) 'Baseline assessment provider Early Excellence admits data 'anomalies' Available at: http://schoolsweek.co.uk/baseline-assessment-provider-early-excellence-admits-data-anomalies/ [Last accessed 30 November 2015]

Selwyn, N., Henderson, M. and Chao, S. (2015) Exploring the role of digital data in contemporary schools and schooling —'200,000 lines in an Excel spreadsheet' British Educational Research Journal. (Last accessed 12th December, 2015). Available Online at, DOI, 10.1002/berj.3186

Siraj, I., Kingston, D. and Melhuish, E. (2015) Assessing Quality in Early Childhood Education and Care. Sustained Shared Thinking and Emotional Well-being (SSTEW) Scale for 2-5-year-olds provision. London: Trentham Books,IOE Press

TACTYC (Association for the Professional Development in Early Years) (2015) 'Position Statement, Say No to Baseline Assessment', Available at: http://tactyc.org.uk/wp-content/uploads/2014/10/TACTYC-Baseline-position-paper-1.pdf

TES (2015a) 'Four (uncomfortable) truths about the new primary baseline test – from inside the eye of the storm' Available at: https://www.tes.com/news/school-news/breaking-views/four-uncomfortable-truths-about-new-primary-baseline-test-%E2%80%93-inside [Accessed 24 November 2015]

TES (2015b) 'Superhead Dame Alison Peacock snubs baseline tests for four-year-olds' Available at: https://www.tes.com/news/school-news/breaking-news/superhead-dame-alison-peacock-snubs-baseline-tests-four-year-olds [Accessed 24 November 2015]

Whitebread, D., and S. Bingham (2012) TACTYC's Occasional Paper No. 2: School Readiness; a critical review of perspectives and evidence, http://tactyc.org.uk/occasional-paper/occasional-paper2.pdsf.

Appendix A: Information on different providers

1. Examples of statements from the Early Excellence Baseline

Examples provided by Early Excellence on their website (Early Excellence, 2015):

Characteristics of Effective Learning, e.g. Creativity & Critical Thinking:

- Having Own Ideas: Thinks of his/her own ideas and different ways of doing things, uses imagination in play.
- Making Links: Makes links and connections in their experiences, developing ideas of grouping, sequences and patterns.
- Reviewing: Reviews activities as he/she does them and changes the approach as required.

Prime Areas of Learning & Development

- Personal Social & Emotional Development – Managing Feelings and Behaviour
  - A. Is able to take turns, accepts the needs of others and is aware of the consequences of their actions.
  - B. Is able to negotiate with peers, recognise and adapt their behaviour to different situations.
- Communication & Language – Listening and attention:
  - A. Listens to others and stories in small groups.
  - B. Maintains attention and concentration when listening to others.
- Physical Development – Moving & Handling
  - A. Moves with confidence in a range of ways, negotiating space and is able to use one handed equipment.
  - B. Experiments with different ways of moving, is able to adjust speed and direction and is able to form recognisable letters.

Specific Areas of Learning & Development

- Literacy – Reading:
  - A. Knows that print carries meaning and knows how to handle books.
  - B. Links sounds to letters, naming and sounding the letters of the alphabet.
- Mathematics – Numbers:
  - A. Counts at least four objects.
  - B. Is confident to order and ascribe using numbers 1-20, add and subtract using single digit numbers.

2. Comparison of the three Baseline Assessments

Provider	Method	Link to EYFS	Training	Cost	Moderation	Report to parents?	Add-ons
Early Excellence	Observations and 'interactions within everyday experiences'; no tests or pre-set tasks	'As EExBA is linked directly to the EYFS it can be used within existing systems to track development to the end of EYFS using the ELGS and then to outcomes at KS1.'	Required	£213 for 30 pupils (£7.12 per pupil) but reduces with larger cohorts	'Light touch' national moderation	No	Nursery baseline; Analysis and interpretation of data training
NFER	Tasks and observational checklists' 'playful'	'In developing the mapping, NFER referred to the ELGs but also to the development statements in the age/stage bands in Development Matters, for 30-50 months and 40-60 months.'	Not required	£225 for 30 pupils (£8.50 per pupil) but reduces with larger cohorts	No	Yes – report places child on scale, labelled working towards, expected or above, for Maths, Literacy and 'general readiness for school learning'. Age adjusted.	None
CEM	Computer based, takes 15-20mins; no data to input	Yes, as part of 'learning curve' for whole of primary	Not required	£126 for 30 pupils (£4.50 per pupil), whatever size of cohort.	No	Yes – report showing child on a scale (without figures) and explanation.	Nursery baseline and later assessments Nursery baseline and later assessments that link. 'Inspection ready' version with more features.

Appendix B: Summary of demographics of survey respondents

1131 Respondents in total

Reception Teacher	49.9 per cent
EYFS Co-ordinator/Phase Leader	37.8 per cent
Senior Leader	7.2 per cent
Support Staff	0.7 per cent
Other	4.4 per cent

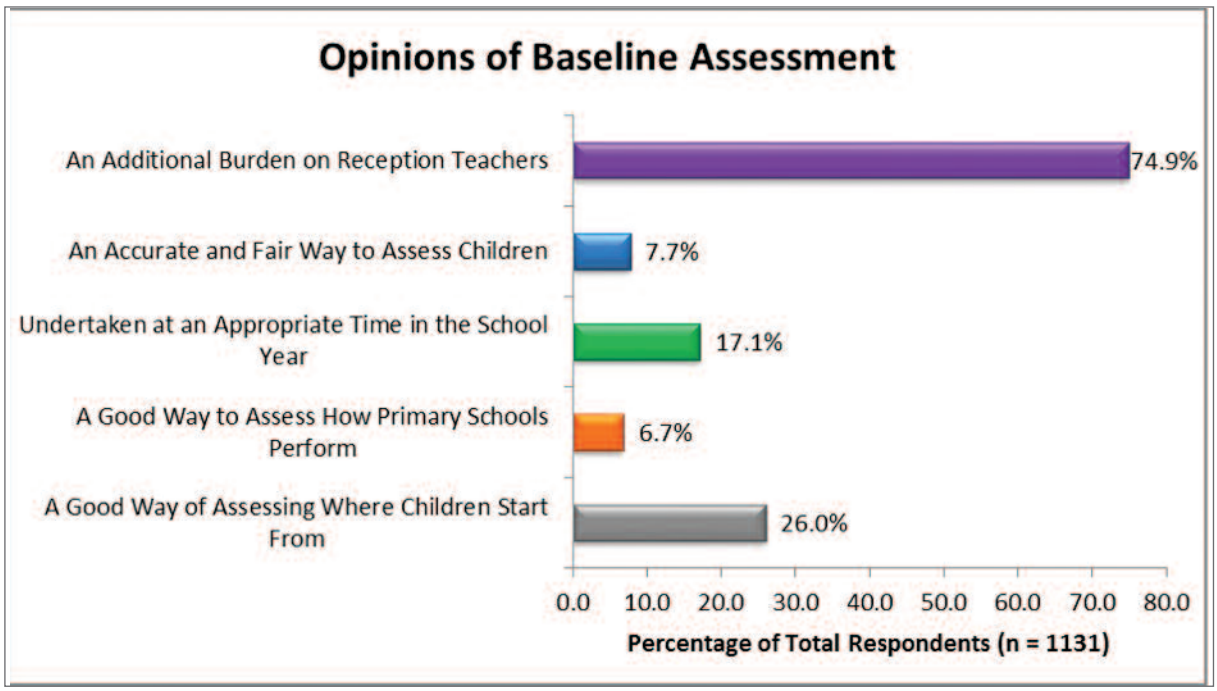
Under 3 years	11.2 per cent
3-12 years	42.3 per cent
Over 12 years	46.5 per cent

Outstanding	20.5 per cent
Good	63.2 per cent
Requires improvement	15.1 per cent
Inadequate	1.2 per cent

State funded (non-Academy)	79.5 per cent
Chain Academy or multi-academy trust	12.3 per cent
Standalone Academy	4.0 per cent
Free School	0.6 per cent
Independent School	0.7 per cent
Other	2.8 per cent

ATL	34.0 per cent
NUT	59.0 per cent
None	7.0 per cent

Appendix C: Summary graph of survey respondents' opinions



February 2016



teachers.org.uk



atl.org.uk

Designed and published by the Communications Department of The National Union of Teachers – www.teachers.org.uk
Origination by Paragraphics – www.paragraphics.co.uk – 10404/02/16