

more
than a
score

Baseline
assessment:

Why it
doesn't
add up

Baseline assessment: Why it doesn't add up

Early years education is at risk. Not just from spending cuts and staff shortages, but from policies which will change how teachers teach and children's opportunities to learn. These policies are not based on research and evidence but on opinion and prejudice. They threaten the quality of education, and the well-being of many children.

At the heart of these risks is the government's decision to resurrect reception baseline assessment – the formal testing of all children in their first six weeks of school.

The government says that a baseline test is required in order to judge the progress children have made at the end of primary school. They want to use the test data to 'hold schools to account'. We say such a test will be damaging to young children. They will be pushed into a world of high-stakes assessment, which is at odds with young children's learning and development.

We are experts in early years education. We are not opposed to the assessment of children's learning. On the contrary, we

support it. Assessment is essential to good teaching and to helpful conversations between teachers and parents. But baseline testing is not good assessment. We will show in this dossier that:

- baseline assessment cannot provide a valid account of the learning of four-year-olds
- it cannot therefore provide a trustworthy basis on which to measure progress
- it will damage the early years curriculum and hold back the learning of many children
- it will adversely affect primary education as a whole
- its costs are not justified.

Alice Bradbury (UCL Institute of Education),
Pam Jarvis (Leeds Trinity University),
Cathy Nutbrown (University of Sheffield),
Guy Roberts-Holmes (UCL Institute of Education),
Nancy Stewart (TACTYC),
David Whitebread (University of Cambridge)

The Baseline Timeline – (1997 – 2017 September)

1997 – Baseline assessment introduced by Labour government

2002 – Baseline withdrawn, in favour of Early Years Foundation Stage Profile

2015 – Baseline reintroduced by coalition government

2016 – Baseline dropped

2017 – Conservative government announces plans to restore baseline.

The Baseline Timeline – (December 2017 onwards)

31 January 2018 – Bidding closed

27 March 2018 – Announcement of successful bidder

9 April 2018 – Commencement of contract

2018/19 – Trialling of the test material

2019/20 – National pilot tests

2020/21 – Introduction of statutory Baseline Assessment

2027/28 – Baseline scores used as basis on which to measure pupil progress, Reception to end of KS2.

The assessment we already have

Teachers already assess children during the reception year, but in more sensitive and fitting ways. Early assessment is important for understanding what children know and can do. Teachers also need to understand children's interests and motivations in order to support their

learning. At present, early years teachers use a rounded approach based on observation of the child across many activities and experiences. They consider a child's well-being, and gain insights from parents and nursery staff. All this is well established throughout the Early Years Foundation Stage (EYFS), starting in nurseries and leading seamlessly to the *EYFS Profile* at the end of the reception year.

The DfE's plans for baseline assessment

The DfE will pay £9.8 million for the first two years to a commercial provider to develop a Baseline assessment. The tender details include:

- The results can be used by schools ('it can be used to supplement schools' existing on-entry assessments')
- It will last up to 20 minutes
- It must be accessible for 99% of the cohort
- The test must ensure that the full range of attainment is appropriately distributed across the range, with fewer than 2.5% of children achieving full marks.
- The assessment is to be delivered in English
- If possible, self-regulation should be assessed.
- Bidders are asked to 'include references to any published research or reports on the validity of the assessment'.

Not a valid way to ensure a good quality of education

Baseline testing: an exact science or myth

Baseline testing is meant to give schools credit for the progress that children make, taking into account their different starting points. It is based on a belief that test data can provide fair comparisons between the start and end of primary school. Accountability data has the aura of science, and it is easy to assume that the numbers deriving from official tests are trustworthy. This is not the case.

All assessments are limited in one way or another, but the younger the child, the greater the flaws are likely to be. It is impossible to capture the dynamism of young children's learning through a single uniform test. Young children respond to tasks very differently in different situations, and from one day to another. Their responses will depend on mood, tiredness, and how settled they are at school. They may not stay focused, or understand what is expected of them.

This is one of the reasons why experts are sceptical about baseline. The statistical data does not support the DfE's confidence that it is possible to create an assessment of reception age children which is suitable for that age group and "sufficiently granular and well correlated with later outcomes such that it could be used as a baseline from which to assess progress." (DfE, 2017)

'If you are testing a four-year-old,' Professor Harvey Goldstein told the House of Commons Education Committee, 'you cannot get a very reliable test.' Most children at such a young

age will not be able to show their true abilities in a test taken out of the context of familiar relationships and practical experiences.

The most recent report from Peter Tymms and colleagues (2017) shows a correlation of only 0.57 between the start of reception and the end of primary school. Earlier research by Tymms and colleagues (2012) showed that baseline tests were entirely unsuccessful at identifying children with special needs.

If Baseline cannot reliably represent children's achievements at age 4, then it is difficult to see how it can be used to measure their subsequent progress. The British Educational Research Association (BERA), responding to the DfE's 2017 consultation on primary assessment, noted 'the marked variation in the pace and sequence at which children acquire a range of literacy and numeracy skills under normal conditions'.

It is a mistake to attribute this variation to school performance. Socioeconomic background is a much stronger influence on progress than any difference between schools. Children growing up in poverty tend to have lower Baseline scores and to make below average progress once at school. The children of well educated parents will tend to make above average progress, as they are likely to have wider experiences, conversations which help them to pick up a wider vocabulary, and might get more help with school work.

All of which led BERA to conclude that 'too much reliance is being given to test data that cannot bear the weight of interpretation placed upon them'.

Measuring the wrong things

The DfE wants the test to focus on 'skills which can be reliably measured and which relate to attainment in English and mathematics at the end of key stage 2, most notably early literacy and numeracy'. A large body of evidence suggests, however, that Baseline assessments focusing on literacy and mathematics are not strong predictors of later achievement. Other aspects of development have been shown to be more predictive of children's development as learners and their later achievement – including spoken language, play and playfulness, and 'self-regulation' (Whitebread 2014). These are not so easily tested as items of knowledge – but they are the real foundations of future development.

Children may score highly at an early age on tests of elements of literacy such as letter recognition – if they have been taught those specific skills. But an early start to formal literacy teaching does not necessarily lead to improved reading later on. A large-scale analysis of the evidence concluded that 'the short-term effects of more academic programmes wore off after a few years in primary school', but 'cognitive-developmental approaches emphasising children's choice, autonomy and self-regulation' produced a long-term effect on educational and social adjustment outcomes (Chambers et al 2010).

Aspects of 'self-regulation', on the other hand, lay strong foundations for later attainment. One research study showed that control of attention at age 4 predicted literacy at age 21, regardless of how good a reader a child was at age 7 (McLelland et al 2013). Assessments of inhibitory control, working memory, and flexibility at ages 4-5 have been shown to predict later competence in reading and mathematics (Blair et al 2007).

The important foundations of emotional well-being and cognitive self-regulation are included in the Early Years Foundation Stage (EYFS) framework. Although the DfE tender suggests the possibility of Baseline assessment including aspects of self-regulation, it has not thought this through: self-regulation is a complex area that cannot be assessed through questions in a test, and cannot easily be turned into data.

Problems with 'value-added'

Far from serving as a useful indicator of school quality, Baseline assessment will leave schools at the mercy of unreliable data. It is unfair to judge a school and its teachers in this way.

It is misleading to judge a school by its pupils' combined value-added data. Many children will be at a different school by the end of Key Stage 2. Some schools have very high levels of mobility, with many pupils joining part way through: this mobility often results in children making weaker progress. Many schools do not have a nationally typical population, in terms of the proportion of children speaking languages other than English (EAL), affected by poverty, or with special educational needs. While nationally EAL children form 20% of the population of primary schools, this swells to over 50% in the case of Inner London. Nationally 30% of children are living in poverty. In parts of London, Manchester, and Birmingham the figure rises to over 50% (End Child Poverty 2018).

Real children, real schools

The DfE's progress measurement system claims that baseline would allow comparison of how schools 'add value' to children's education. This profile of a Year Six cohort at a Birmingham primary school shows how unfair this approach is. Only half the Year Six pupils have been at the school since Year 3 – so how could the school be accountable for their progress? And with so much happening in the children's lives, how can a 'school effect' on their progress be fairly identified?

(The national averages are shown in brackets.)

63% disadvantaged (25% nationally)

54% English as an Additional Language (20% nationally)

29% Special Educational Needs (12% nationally)

13% close to permanent exclusion

53% at the school since Y3

25% arrived in Y4

13% arrived Y6 (including new to the UK)

33% with significant family difficulties

79% with significant social, emotional and family issues

4% life threatening medical condition

20% diagnosis of autism spectrum disorder

<http://www.birminghammail.co.uk/news/midlands-news/desperate-schoolchildren-caught-filling-pockets-13881509>

Gaming the System

If Baseline tests are used to measure schools' performance, schools will have an incentive to keep Baseline results low in order to show more progress later on. This was recognised as a potential problem by the National Foundation for Educational Research, research for the Department for Education before the 2015 pilot; the report noted 'some evidence' of gaming results in order to maximise progress measures (DfE, 2015). In the voluntary pilot year itself (2015), research again identified this as a problem. As one teacher commented, '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' (Bradbury and Roberts-Holmes, 2016b). These findings chime with a wide body of research on education policy which demonstrates that when tests are 'high stakes', distortions occur (Stobart, 2008). This was admitted by Robert Coe, the director of Centre for Educational Measurement, Durham University, which had provided one version of Baseline assessment. He commented in his evidence to the Education Select Committee:

When you introduce high stakes into that and compulsion, the whole thing changes and that is something that we need to think quite carefully about... The incentive is to want to do badly on a Baseline if it is progress you are going to be measured on. (Education Select Committee, 2017)

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. (teacher in Bradbury and Roberts-Holmes, 2016b)

Where there is pressure to deflate results, classroom ethos is also affected. As one teacher noted in 2015:

The desire to minimise rather than celebrate what children can do when they start school has the unfortunate effect of distorting the thinking and priorities of school leaders and teachers in reception.

Headteachers' wishes for low Baseline

The problem is at the moment the accountability component dominates everything else and it distorts the curriculum, it distorts learning, it distorts children's behaviour. There is lots of evidence now about the stress that children go under. Assessments should not be doing that to children. Assessments should be encouraging children to learn.

Harvey Goldstein, Professor of Social Statistics, University of Bristol, evidence to House of Commons Education Committee, 18th January 2017

The harm to children

Baseline labels children from the start

Enormous damage can be done to children when adults believe that their 'ability' and potential can be defined at a young age. The return of Baseline testing risks worsening the harm caused by belief in 'fixed ability'.

It is clear that baseline data will be used by schools to make judgements about individual children. The DfE state that the test is 'not intended to provide detailed diagnostic information about pupils' areas for development' – but adds that it will be 'used to supplement schools' existing on-entry assessments', and there may be 'some brief commentary on pupils' relative strengths and weaknesses'. (DfE 2017)

Grouping by ability is becoming more common, even in the Reception year, and teachers are likely to use Baseline to determine children's groups. In the 2015 pilot of Baseline (Bradbury and Roberts-Holmes 2016b) teachers stated:

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

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)

Grouping implies a prediction of potential progress. Early predictions can become a self-fulfilling prophecy. Children placed in so-called 'low ability' groups are likely to experience a more limited curriculum and increased pressure, depriving them of the foundational experiences they need. Low expectations may limit children's horizons. This will have a disproportionate impact upon children already experiencing stress at home.

Negative labelling and placement in lower groups are also serious problems for children who are born in summer (and so nearly a year younger than others in the class), those for whom English is an additional language (EAL), and children growing up in poverty whose life experiences may have been more limited (Jensen 2009). Summer-born children consistently receive lower scores in early assessments, and are over-identified as having Special Educational Needs, when they are actually just younger rather than less able.¹

Baseline assessment using a universal test is likely to be insensitive to difference and diversity amongst children so that

'everything can be reduced to a common outcome, standard and measure. What it cannot do is accommodate, let alone welcome, diversity – of paradigm or theory, pedagogy or provision, childhood or culture' (Moss, Dahlberg, Grieshaber et al. 2016).

¹ The Government has now indicated that parents may delay their summer-born child's start in reception by a year. This may be of benefit to those children whose parents take up this option, but it will only widen the age range of the children being measured in a one-size-fits-all model. It is also possible that delaying entry may be more likely for more advantaged children whose parents may be more informed and have the resources to arrange alternatives to school entry.

The proposed Baseline test does not even recognise diversity of language: the tests will be entirely in English. Such an intolerance of linguistic diversity negates EAL children's potential, by refusing to tune into the learning they could demonstrate in their home language. Not surprisingly, many EAL children have had less exposure to English and will score lower on a baseline test (Andrews 2009). With the right sort of help, they can follow the curriculum well and make strong progress (Strand 2016) but low baseline scores could easily lead to low expectations.

Disrupting the vital settling-in period

A settling-in period is important for establishing supportive, encouraging and nurturing relationships between reception teachers, support staff, children and their families. This is an emotionally crucial time for young children. They need to build trusting relationships in their early months at school. Parents may also feel vulnerable and apprehensive.

During the trial run of Baseline in 2015, 59% of teachers agreed that 'The Baseline Assessment has disrupted the children's start to school' (Bradbury and Roberts-Holmes, 2016b). As one teacher commented:

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.

The DfE specifies that the tests (probably computer-based) will be conducted on a one-to-one basis according to a set procedure, with children individually withdrawn from sociable and enabling

activities with other children. This process is highly disruptive to the teacher's need to establish a sense of trust and community. Baseline tests take the teacher out of the classroom. The demands of Baseline assessment and data collection take priority.

Baseline tests will be applied to 99% of pupils, with very little regard to special needs and circumstances. In the 2015 trial of Baseline Assessment young children were very well aware that they were being tested, and for some children the test was an emotionally damaging experience:

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, Cedar)

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, Centre for Educational Measurement, Durham University, user) (Bradbury and Roberts-Holmes, 2016b).

This is ethically unacceptable, especially at a point when young children's confidence needs building up, not undermining. These tests will be far removed from the reciprocal informal interactions with familiar adults that enable a child to relax and participate freely.

Undermining partnerships with parents and early years settings

Because children demonstrate different aspects of their learning and capacities in different contexts, assessment must include the insights of parents, as well as staff from nurseries or children's centres. Early years centres all hold detailed information about individual children, which is commonly shared through passing on records and sometimes through discussion. Parents' insights are gained in introductory meetings and through shared records before the children start school, and this is continued through on-going partnerships. One of the problems of Baseline testing is that it sets this wealth of reliable information to one side.

If schools feel pressured to 'under-assess' – to engineer situations conducive to low scoring in baseline tests – this is likely to result in on-going tensions between nurseries and reception units.

Ignoring parents' knowledge of their children undermines the home-school partnership. Some schools may keep Baseline secret. Some parents may see Baseline as a judgement on their parenting and 'teach to the test', as this head teacher describes:

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. (Bradbury and Roberts-Holmes, 2016).

Harming children's learning and development

Because Baseline assessment focuses on a narrow range of knowledge and skills, it is likely to encourage reception and nursery staff to concentrate on providing narrow experiences. It will limit the rich exploratory, playful, creative, and intellectual experiences which benefit children in the early years (Moyles 2015).

Staff in nurseries and preschools will be under pressure to prove that children are 'ready for school' and so may focus on the particular knowledge and sub-skills that will be tested. Reception teachers will be expected to show progress in these narrow measures. Ofsted's Bold Beginnings report, whose recommendations make not one mention of play, is an unfortunate move in this direction (TACTYC 2017). Children will be increasingly subjected to inappropriate and excessively formal teaching, while missing the crucial input of adults in supporting children's thinking about their play (Jarvis 2018).

Children become confident, resilient learners through experiences which are right for their age group. In the early years, play and playful experiences offer the richest context for children to meet challenges, take risks, make decisions, solve problems, re-group and bounce back when intentions are not met, build concentration and intrinsic motivation, and think creatively and critically. Alongside secure attachments with key adults, playful interactions with other children help children learn about their own and others' emotions and develop perspectives which underpin their progress through life. (Whitebread,

Kuvalja, & O'Connor, 2015). This is the best way to support children's development and school success.

Children in England begin school earlier than in most other countries, so it is essential that approaches appropriate to older children are not used too soon. Children who are introduced to formal learning at the age of four or five often fail to benefit long-term, and may also develop higher anxiety levels, lower self-esteem and less motivation to learn (see Elkind 2001).

Elsewhere – for example, in Finland – early childhood education is focused upon health and well-being rather than cognitive achievement (Sahlberg 2014). The Scandinavian emphasis upon educative play until the children are aged six and seven, combined with very low rates of child poverty, means that by the age of 11 Finnish children perform better than their English counterparts. Sahlberg (2014) notes that the success of Finnish schools is partly the result of not ranking children according to educational success, so any sense of student 'failure' is removed.

Damaging mental health and well-being

Escalating concerns about children's mental health should make the Government wary about introducing Baseline tests. A Baseline test on entry to school adds yet another stressor to the existing load which many children already experience.

The 'Adverse Childhood Experiences' research (Felitti et al 1998) revealed that, when placed in stressful environments, young children are likely to exhibit abnormally raised levels of the stress hormone cortisol. This hormone mobilises the 'fight or flight' response with glucose pumping into the

muscles. Hence cortisol disturbances in young children can lead to suppressed growth, anxiety, depression and less memory capacity for intellectual development (Guilfoyle and Sims 2010). Whilst some amount of stress is a normal feature of human life and can prevent boredom and inertia, constant stress can cause them to become chronically anxious. This is particularly so when it is rooted in insoluble problems and paired with low levels of physical activity. Long term anxiety can in turn become a contributing factor to physical and mental illness.

A high-stakes Baseline test will inevitably raise the assessment-fuelled stress load on teachers (TES 2017a online). 'Baselined' children will move up into subsequent year groups, taking their (unreliable) 'Baseline' statistic with them to be applied to successive teachers' performance management.

Social media, assessment and poverty pressures have already become a toxic trio and young people are far more likely to have mental health problems than previous generations. These include depression and eating disorders (Patalay and Fitzsimmons 2017) and self-harm – for which hospital admissions have increased 42% in a decade (The Guardian 2016, online).

The pressures are particularly severe for children growing up in poverty. Schools are being expected to expose children to high-pressure learning in the interest of 'social mobility', rather than providing rich learning experiences. The Government would do better to remove the negative pressure and alleviate poverty.

Baseline assessment: a waste of public money

The contract for Baseline for the first two years will hand £9.8M to a private firm (TES 2017b), while other costs at the DfE will multiply that total. Baseline tests are an example of England being used as a laboratory for new school tests.

Using commercial companies enables the DfE to get controversial policy 'done'. In the previous (2015) trial, niche marketing strategies designed to appeal to the early years community made Baseline palatable and acceptable in a way that the DfE could never have done on its own.

Within the context of high stakes testing, many headteachers thought it would be 'irresponsible' not to do the 2015 baseline trial even though it was voluntary. They became customers of a new branch of the testing industry. In the period 2015-17, government spending on Baseline assessment totaled £4.6M, in addition to large hidden costs incurred by schools for training, staff cover, etc. (Clark 2017):

I don't think I have ever come across a situation where Heads feel so at sea.... the Local Authority has been so cut back I think it creates even more scope for people to actually make a buck out of this whole thing if I'm honest. (Primary School Head)

What has happened is we have been opened up to a completely free market and we are being bombarded with sales pitches. I think that the companies at the

moment can really capitalise on the fear factor in schools. Especially when the stakes are so high and you are looking at how we are judged and all the Ofsted... (Primary School Head)

(Bradbury and Roberts-Holmes, 2016b)

The headteachers were concerned that simplistic use of data enables commercial companies to move in, and make money from a 'datafied' early childhood (Lupton and Williamson, 2017). These companies will harvest masses of marketable data. The ethics of where such data will be kept, processed and analysed has not been explored or clarified by Government.

The commercial company that wins the lucrative baseline contract will be able to market solutions for the problems it throws up – a win-win commercial situation. Additionally it will be able to produce and sell commercial spin-offs from its Baseline products such as test practice websites, books and extension activities. Some of the Baseline providers in 2015 even advertised baseline tests for use in nurseries.

At a time of austerity, with education budgets under extreme pressure, handing public money to commercial companies for processes that are not of benefit to children or schools cannot be justified. When far too many children in the UK live in poverty, we are going to spend fortunes on testing rather than feeding them

A better way

More Than a Score believes that Baseline testing as the foundation for high-stakes judgements about schools is deeply flawed.

It will damage children and schools. It will bring the youngest children into a system which results in teaching to the test.

Baseline will narrow the curriculum to very limited versions of literacy and numeracy, rather than the broad range of learning which will help children flourish. Pressure to improve test scores places harmful stress on children and teachers, while depriving children of the rich, creative curriculum to which they are entitled.

Schools should be responsive to the wider society, and to children, their families and communities. **But there are better ways of ensuring that schools are responsible for the quality of education they provide.**

Assessment for learning is a vital part of teaching. Ongoing formative assessment supports children while they are learning. It can engage children more fully and enrich their development.

Assessment in the reception year should not be a one-off test, but an authentic professional judgement based on sensitive and observant assessment over time. This should also pay attention to the insights of parents and information from previous centres. Most schools already make tentative judgements within a term of a child starting school, based on the age-stage bands of *Development Matters in the Early Years Foundation Stage* (Early Education 2012) which leads into the end-of-phase EYFS Profile. These informed, professional assessments can be used to the end of the year and beyond. Policy-makers should do more to develop teacher expertise in this area rather than spend money on baseline tests.

There are various ways of ensuring that educators are responsible to parents and society, and of improving the quality of children's education.

- Schools can compare their assessments in meetings with other schools. This will involve looking closely at children's work and recordings of their activities, then adjusting their initial judgements.
- Children's progress can be discussed in comparison with national expectations, but paying attention to the individual characteristics and family situation. This can be supported by experienced staff from other schools and consultancy support from local authorities.
- Schools can self-evaluate with the help of 'critical friends' from other schools, as well as inspection visits which are supportive and consider local needs.
- Schools should use information from intelligent assessment to work out improvement plans which are in tune with the school's values and responsive to the local community.
- The Department for Education should evaluate national standards of achievement across the whole curriculum through detailed study of a sample of pupils, rather than blanket tests affecting all children. This will enable it to analyse strengths and areas for improvement.

It is time to lift the fear that dominates our schools, and reduce the pressure which is causing children distress.

Children are *more than a score*. They will learn successfully when we stop measuring their every step, and develop more rounded ways to ensure they receive the quality of education they deserve.

References

- Andrews, R. (2009). Review of Research in English as an Additional Language (EAL). London: Institute of Education.
- Blair C, Razza RP. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*. 2007;78(2):647-663.
- Bradbury, A., & Roberts-Holmes, G. (2016a). Creating an Ofsted story: the role of early years assessment data in schools' narratives of progress. *British Journal of Sociology of Education*, 38.7 (2017): 943-955. DOI: 10.1080/01425692.2016.1202748.
- Bradbury, A., & Roberts-Holmes, G. (2016b). 'They are children, not robots': The Introduction of Baseline Assessment. London: ATL/NUT.
- Chambers, B., Cheung, A., Slavin, R.E., Smith, D. and Laurenzano, M. (2010) Effective early childhood education programmes: a best-evidence synthesis, CfBT.
- DfE (2017) Invitation to Tender (ITT) for Reception Baseline Assessment.
- Education Datalab. (2015). Seven things you might not know about our schools. Available at: <http://www.educationdatalab.org.uk/wp-content/uploads/2016/02/EduDataLab-7things.pdf>
- End Child Poverty Now (2017) <http://www.endchildpoverty.org.uk/poverty-in-your-area-2018/>
- Jarvis, P. (2017) Is Baseline Missing the Bigger Picture, TES, 21st November.
- Jarvis, P. (2018) Child Development: the invisible man. 'The Psychological Historian' <http://histpsych.blogspot.co.uk/2018/01/child-development-invisible-man.html> 1st January.
- Jensen, E. (2009) Teaching with Poverty in Mind: what being poor does to kids' brains. Association for Supervision & Curriculum Development
- McClelland, M.M., Acock, A.C., Piccinin, A., Rhea, S.A. & Stallings, M.C. (2013). Relations between Preschool Attention Span-Persistence and Age 25 Educational Outcomes. *Early Childhood Research Quarterly* 28, 2, 314-24.
- Moss, Dahlberg, Grieshaber et al. 2016.
- Moss, P., Dahlberg, G., Grieshaber, S., Mantovani, S., May, H., Pence, A., Rayna, S., Swadener, B, Vandenbroeck, M. (2016) The Organisation for Economic Co-operation and Development's International Early Learning Study: Opening for debate and contestation.
- Contemporary Issues in Early Childhood, Vol 17, Issue 3, pp. 343-351.
- Moyles, J. (2015) *The Excellence of Play* (4edn). Maidenhead: Open University Press/McGraw Hill.
- Ofsted (2017) *Bold Beginnings: the Reception curriculum in a sample of good and outstanding primary schools* London: Ofsted.
- Patalay P & Fitzsimons E. Mental ill-health among children of the new century: trends across childhood with a focus on age 14. September 2017. Centre for Longitudinal Studies: London.
- Steele, H., Bate, J. Steele Shanta, M., Dube, R., Danskin, K., Knafo, H., Nikitiades, A. Bonuck, K., Meissner, P. and Murphy. A. (2016) Adverse Childhood Experiences, Poverty, and Parenting Stress. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* Vol. 48:1, pp.32-38.
- Stobart, G. (2008). *Testing times: the uses and abuses of assessment*. London: Routledge.
- Wrigley, T. (2015). *Predicting children's potential: baseline tests*.
- Strand, S. (2016) Educational outcomes among children with English as an additional language (EAL) (Oxford, Migration Observatory, University of Oxford).
- TACTYC (2017), *Bald Beginnings: TACTYC's Response to Bold Beginnings*, <http://tactyc.org.uk/wp-content/uploads/2017/12/Bold-Beginnings-TACTYC-response-FINAL-09.12.17.pdf>
- Tymms, P., Merrell, C., Henderson, B., Albone, S. & Jones, P. (2012). Learning Difficulties in the Primary School Years: Predictability from On-Entry Baseline Assessment. *Online Educational Research Journal* June 2012.
- Tymms, P., Merrell, C. & Bailey, K. (2017). The long-term impact of effective teaching. *School Effectiveness and School Improvement*. Online. <https://doi.org/10.1080/09243453.2017.1404478>
- Whitebread, D. (2014). The importance of self-regulation for learning from birth. In H. Moylett (Ed.). *Characteristics of Effective Learning: helping young children become learners for life* (pp. 15-35). Maidenhead: Open University Press.
- Whitebread, D., Kuvalja, M. & O'Connor, A. (2015). *Quality in Early Childhood Education – an International Review and Guide for Policy Makers*. Report for the World Innovation Summit for Education. Dohar: WISE.

more than a score

The 2015 baseline assessment providers on the government's new baseline proposals:

"Self-contradictory, incoherent, unworkable and ultimately inaccurate, invalid and unusable." – Early Excellence

"If teachers aren't given helpful and beneficial results that can help children, it's verging on the immoral." – Katharine Bailey, Centre for Educational Measurement, Durham University

