

Data in the early years and the problems of assessing young children

Alice Bradbury and Guy Roberts-Holmes

Our research over the last few years in early years settings and Nursery and Reception classes in schools has found that the early years sector is currently obsessed with data. The collection and analysis of data have become key features of early years practitioners' and teachers' work, motivated by the need to track children's progress and fulfil government policy such as the introduction of Baseline Assessment in Reception, and the EYFS Profile. But the important question is often overlooked: what does this data tell us? Children are complex, complicated beings, so that the lived experiences of those who work with them contrast with the simplicity of assessments that require simple yes/no answers. Our research suggests that the early years is undergoing a process of 'datafication', that is the shift towards prioritizing data collection, and that this can be a reductive process, which reduces the complexity of children's learning to numbers. This research is motivated by continued policy decisions – such as the re-introduction of Baseline recently reported – which force practitioners to engage in more data collection.

Our research in the last few years has involved case studies of Children's Centres, Nursery Schools, and Nursery and Reception classes in schools, as well as detailed analysis of the introduction of Baseline Assessment in 2015. We have found that there is a tension between the idea that data can be produced that is definitive and accurate, and the views of practitioners that assessment with young children is very complicated. In particular, teachers objected to having to make simplistic decisions about what a child could do or not do; as one teacher we interviewed commented, 'It is a binary yes/no! And children aren't like that, children are more complicated than that'.

Reducing children's learning to numbers is in stark contrast to the views of many early years practitioners and researchers, who argue that assessing young children's learning includes being able to sensitively 'tune' into their multiple and diverse 'voices', often observed during children's play, to understand their meaning-making. On-going observational assessment as found within the EYFS Profile 'listens' to what children are interested in during children's playful and sociable learning activities. Such complex assessment demonstrates what children can do, often in play contexts, and is open to children's unpredictable complex and diverse potential. This is important because the positive development of young children's self-regulation and growth mindsets have greater longterm benefits than narrow cognitive skills, which can be developed more effectively later, as in most European countries.

So what do we know about assessing young children? Research suggests that many other factors can affect children's scores. When children first arrive they may not be able or willing to show what they can do in an unfamiliar environment or with a new adult. Scores are affected by how tired or hungry the children are, according to the teachers who conducted Baseline, as well as factors such as familiarity with the tools of assessment, such as a tablet. Similarly Margaret Clark's research on Baseline found that various aspects of context affected the results, such as the teacher's level of training, previous attendance at nursery and the child's home language. This is the 'messy' lived social reality and context in which data are generated; early years settings are not scientific laboratories in which all the many changing variables of four-year-olds' fast-moving lives can be statistically controlled for. So, the problem here is that even 'objective' computer tests are in fact distorted, inaccurate and invalid due to the inherent multiplicity of testing variables in individual school contexts. Formal one-to-one assessment, even when it uses toys and interactive tablet-based characters, depends on a child's willingness to engage at that moment and their relationship with the adults, which damages the reliability of any assessment.

Furthermore, early years research suggests that a number of social factors have an impact on attainment, and crucially on readiness to engage with assessment tasks, at this age. We know from the EYFS Profile data that attainment at the end of Reception reveals disparities by gender, ethnic group, EAL status and levels of disadvantage. Potentially, using a one-off assessment such as Baseline can lead to low expectations for a child's whole educational career.

So, is any 'accurate' assessment at this age possible? Teachers have been using observational assessments on entry to school or nursery for a number of years. The key difference has been that this information has been used to plan, and to compare with assessments a short time later. It has been part of the picture that the class teacher has built up with the child, rather than information decided at one snapshot point in time and solidified into a number used for comparison. Children at this age change rapidly almost from day to day, so a one off test cannot capture this fast paced changeability. So for example, this means that if a child defies expectations, for example by counting cups in the water tray when playing, but they refused to count the bears on the ipad in the assessment the day before, the teacher can revise their assessment, because it is not formalised. It does not need to be 'accurate' in the sense of validity and reliability, because it is not fixed. This form of flexible assessment is more appropriate with young children and useful to the teacher, but of course does not serve the purpose of setting a standard for later comparison. Although the EYFS Profile is not a perfect tool, it is very popular among teachers because the process involves building a picture of what a young child can do often during children's playful and sociable learning activities.

Observing, assessing and sharing what children *can do* helps to build young children's self-confidence and positive learner identities.

We know that trying to assess children who are not yet sufficiently emotionally competent is highly problematic. Young children are remarkably sensitive to context so are likely to be aware of being tested and will probably pick up on their teachers' and parents' anxiety too. In our research, teachers commented that some children found Baseline Assessment so difficult that they cried, even though it was designed to be responsive to which questions children answered correctly

Given all of these problems with assessing young children and reducing their complexity to data, well documented in ours and others' research, the re-introduction of Baseline is a backwards step. Some problems raised after the 2015 Baseline fiasco have been considered – such as the use of different providers - but there remain some fundamental questions about the appropriateness of an assessment at four being used to predict attainment. The Government claim that Baseline will 'not be used to 'predict' the Key Stage 2 results of individual pupils – although the entire premise of measuring value added works on the basis of an expected level of progress - and there is no suggestion that children should be explicitly allocated an expected score for Year 6. However, Baseline data will inevitably leak into the system over the seven years between Reception and Year 6; moreover, parents (and schools) will want to know what children scored in Reception (and ethically and legally probably have a right to know).

Therefore, finally, we raise another question about the wisdom of using assessment of young children to *predict* later attainment. Analysis of existing assessment data reveals some interesting points on prediction: Education Datalab found in their analysis of Key Stage 1 and 2 results that only 55% of pupils 'make the anticipated linear progress to reach the Key Stage Two Level that is predicted for them from their Key Stage One score'. So, in the space of four years, nearly half of all children either do not make the necessary progress or exceed expectations. Moreover, we know that there is a temptation to 'game' the system: Education Datalab research suggests that where Key Stage One teacher assessments are used as a baseline for measuring progress, teachers deflate scores to improve progress measures. This is shown through the differences in scores from nursery schools, where teachers have no incentive to deflate scores.

Through our work on 'datafication' in early years we have researched the use of both one-off assessments like Baseline and longer-term assessments; both play a part in reducing children to numbers, but the simplistic form of assessment used for Baseline is particularly problematic in that it overlooks the complexity of children, instead telling pseudo-scientific, deficit-based and potentially damaging story about young children at the very start of their schooling journey.