



Assessment for Social Justice

and the potential role of new technologies

A workshop report

By Lyndsay Grant

June 2008

The report summarises discussions held during a workshop on ways that assessment could be used to promote social and educational justice, and the potential role of new technologies in achieving this.

We are very grateful to those who participated in the workshop, and contributed their experience, skills and insight to a lively and informative discussion. They are: John Gardner (Queen’s University), Kathryn Stowell (Charlton School), Anita Sawyer (Charlton School), Alan Howe (University of Gloucester), Nina Reeves (University of Gloucester), Ian Sillett (OneFourSeven), Gary Stevens (OneFourSeven), John Galloway (Consultant), Tony Wheeler (Goldsmiths University), Dan Buckley (Cambridge Education), Mike Davies (Consultant), Sean McCusker (Durham University), Terry Waller (Becta), Denise Whitelock (Open University), Andrew Littledate (BBC).

ASSESSMENT AND SOCIAL JUSTICE

Assessment is central to educational practice – and has a huge impact on young people’s lives. High-stakes assessment instrumentally affects further education, employment and life choices available and being judged a failure can lead people to see learning as a closed book for the rest of their lives. When used well, assessment supports learning, but when applied badly, can also hinder it. Because assessment is such a core part of education, it is crucial that it accommodates the diverse needs of all learners, including disadvantaged learners.

The education system discriminates against learners from poorer backgrounds, with learners from low-income households achieving lower GCSE grades than their more privileged peers. The assessment system itself must be implicated in this discrimination. Socio-economic and cultural barriers continue to pervade the education and assessment systems. Further barriers to equitable participation in assessment include physical and intellectual learning difficulties and disabilities, access to resources (including technology), and class and cultural and language differences.

A fairer assessment system would also mean a fairer and more socially just education system. This workshop therefore sought to explore the following questions:

- What might fairer assessment look like?
- How might technology be used to create fairer assessment tools?

WHAT’S WRONG WITH THE CURRENT SYSTEM?

We currently use a variety of assessment tools and systems. An indication of the range of common assessment tools currently used is given below.

1. Evidence of work	2. Multi-part tests	3. Interviews	4. Non-hierarchical assessment	5. Practical skill demonstration
Written exams	Multiple choice quizzes	Viva (oral defence of thesis)	Peer assessment/ review	Competency/proficiency tests (eg driving test)
Coursework	Psychometric tests	Face-to-face interviews	Self-assessment	Observation of practice
Portfolios and exhibitions	Adaptive computerised	Performance review/appraisal		Musical composition
				Sporting trials

	tests			Physical measurements
--	-------	--	--	-----------------------

During workshop discussions, the problems with such current assessment tools were discussed, falling broadly into eight categories described below.

Mistaking high-stakes assessment for learning

Most assessments do not measure 'learning' itself, but use a proxy measurement to infer that learning has taken place. For example, a three-hour written exam cannot possibly measure two years' worth of learning, but it is often used as a proxy. The danger is that, in high-stakes assessments, the proxy measurement comes to be seen as more important than the learning that it is designed to measure. Thus learning becomes confused with its assessment; and assessment, which should serve learning, comes to dominate it. An over-emphasis on the final assessed performance can lead to teaching to the test, and a failure to value and recognise the whole journey of learning.

Mode of communication

Assessments will favour those who are well able to communicate their knowledge and understanding through the format specified by the exam. For example, people who are not confident speakers may perform worse at interview than people with similar, or lower, levels of knowledge and skill, but who express themselves better verbally. Similarly, our current assessment tools cannot accurately measure some things we may value highly, such as the ability to collaborate with others, and we may need to develop new ways to communicate such skills and attributes.



Time-bound

Observing someone's performance in a particular time means that it is subject to the constraints of that moment. Exam nerves and other external factors can all affect the quality of a one-off performance. Such a single performance measure values the final accomplishment of learning, without acknowledging the process of that learning.

Subjective judgements

While the importance of human judgements on quality of work was recognised as giving more nuanced judgements than sticking to rigid and strict criteria, the subjectivity of such judgements can introduce (however unintentionally) bias and prejudice. For example, graffiti art may be judged by some as ground-breaking creativity, while others dismiss its merits. Relationships between the assessor and learner, both good and bad, can influence the judgements made about the quality of the work produced.



Measuring what you can't do

Assessments based on competency frameworks often assume a 'norm' and measure performance against this, looking for evidence of what a learner is unable to do. This type of deficit model of assessment works against learners who demonstrate abilities and knowledge that are different from, but not necessarily less valuable, what the examiner is looking for.

Cheating

While some less restrictive assessments might allow knowledge and abilities to be demonstrated more flexibly, there is the possibility that they can open the way for cheating. For example, coursework, originally seen as allowing learners to demonstrate their abilities free of exam pressure and to develop their thought over a longer period, is now criticised because of the possibility of cheating. Cheating is perhaps almost inevitable when passing the test is seen as more important than the learning that it attempts to capture in the first place. Cheating, as it is hidden from other learners and the assessor, means that it is not possible to tell what forms of support and help different learners have had access to; and therefore inequalities in support and resources can transpire into inequalities in learners' assessed performance.

Cost and time

In order for assessments to be scalable, they must be able to be implemented at a reasonable cost and within a reasonable time scale. This often works against more in-depth, nuanced assessments that would generate a more complete picture of someone's learning process. This tension between scalability and standardisation on the one hand, and flexibility and depth on the other often comes down to questions of time and cost.

Hierarchy of assessment

Peer assessment is seen as potentially democratising the assessment process, by bringing the judgements of peers as well as 'experts' into the process. Self-assessment similarly places the responsibility for making judgements on the learner themselves. While many teachers might value peer and self-assessment, they are sometimes perceived to be unreliable, and are rarely used for high-stakes summative assessments.

WHAT WOULD WE WANT TO BE DIFFERENT FOR CHILDREN BORN TODAY?

Children born today in 2008 will be in 18 in 2026. Workshop participants discussed how they would like to see assessments change by 2026, to be fairer and more equitable.

These fell into six broad categories:

1. Assessment to be learner-driven

The role of school would dramatically change, with schools acting as educational 'brokers', connecting learners to experts and resources of their choice. Schools would become learning 'hubs' connecting learners of all ages and all stages, with no year groups. Learners would make real choices about the tools, activities and time of their learning and assessment. Each learner would have a personalised model tailored to their needs and interests, deciding how they wish to present their learning for assessment. Assessment



becomes a dialogue between learner and assessor, rather than something that's 'done to' the learner.

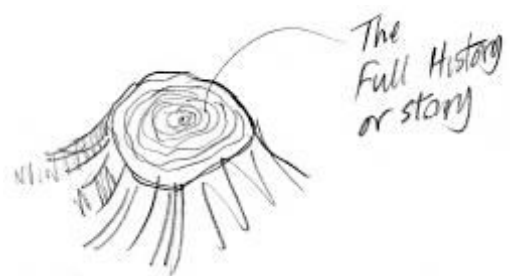
Assessment would be built on a success model rather than a failure model, sympathetic to the learner's current stage of development, building on what they have achieved rather than focusing on their deficits. Assessment would be closer to the learner, with peer and self-assessment having a greater role.

2. Teacher's role

Teachers would have more time to be involved in each learner's journey, understanding their learning in more depth and able to facilitate their progression. The teacher's role could be distributed amongst lots of people in the community, valuing the learning that happens both in and out of school. The teachers' role in assessment may be a 'staging post' in the process of transferring greater control to the learner.

3. Focus on the whole story of the learning process

The process of learning would receive more recognition and validation, with the question of *how* you got there as or more important as *where* you got to. Making a distinction between the 'learning' and the 'assessment' is in many ways a false distinction; assessment would be part and parcel of the whole process, rather than an add-on at the end.



4. Redefine what gets assessed

Assessment would cover a broader range of skills, knowledge and competencies. Rather than focusing on specific content, assessment would include skills and competencies such as collaboration, creativity and problem-solving. Rather than subject knowledge, assessment could be extended to include an understanding of what it means to think 'scientifically' or 'geographically', etc. It was also recognised that current subject and content-focused assessment was not meeting the needs of employers, and that assessment needs to focus on the kinds of skills likely to be needed for employment in the 21st century. However, there remains a question as to how we really can predict what those skills will be.

5. Reduce focus on high-stakes assessment

There was agreement that over-emphasising high-stakes assessment was damaging to schools and learners alike. A future assessment system would remove league tables and feature less formal assessment throughout a child's school career, allowing more space to be creative. Assessment would also be more about *describing* learning rather than *measuring* it, providing detailed evidence of a learner's achievements and abilities without the need for a definitive comparative measurement.



6. More inclusive

Questions of when and how assessment takes place, and the purpose of assessment, need to be addressed with an explicit consideration for social inclusion and educational justice. Competition between learners, parents and schools misses the point and needs to be replaced with a focus on meeting the different needs of all our learners at an individual level.

HOW COULD WE USE TECHNOLOGY TO CREATE FAIRER ASSESSMENT TOOLS IN THE FUTURE?

Groups explored ideas for new assessment tools that could become available in the future, to make assessment fairer. They were asked to think of changes to technologies and to practice that might be possible in 18 years' time.

1. Holistic assessment using Artificial Intelligence

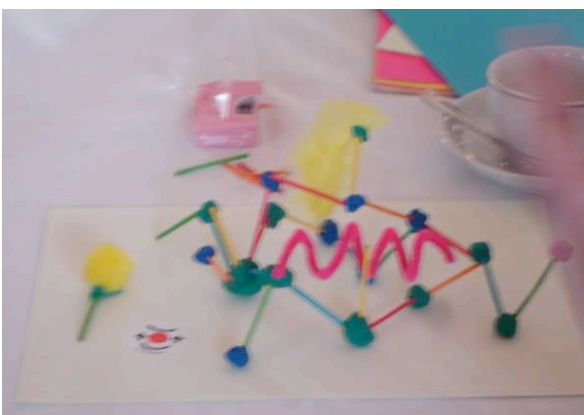


Whereas the current assessment system attempts to fit every unique individual through a one-size-fits-all assessment system, one group proposed using Artificial Intelligence (AI) to assess a far wider range of dimensions, traits and attributes, including emotional and spiritual, as well as cognitive development. This would allow teachers to support and value young people's development as people, rather than simply focusing on narrow standards of academic work. Such a system may be able to show not just what people can do, but what people are like. Teachers and others make such judgements all the time, but by using AI based on pattern recognition, many

different characteristics could be aggregated to enable these judgements to be scaled up so that they would not depend on particular individuals. However, the final judgement would still need to be made by humans, not AI.

This idea raised some interesting questions about the nature and purposes of assessment. Some felt that such an in-depth assessment could potentially be invasive of an individual's privacy, and that some aspects of an individual's life should not be subject to assessment at all. The question of who 'owned' such a tool was also seen as crucial. If the tool was primarily owned by teachers, or external agencies, and used to make judgements about the learner, then it could be experienced as an invasive tool. However, if the learners themselves controlled both the process and the outputs and its main purpose was intended to help learners to take decisions about their own progression – then it might be a very powerful tool indeed.

2. Inspiring and connecting people



This group re-imagined the learning process connecting to inspirational individuals. Seeing everyone as potentially inspirational in some sense, people would learn in different ways and at different times, forming a community of people who both inspired, and were inspired by, one another. This community would not be hierarchical in the traditional sense of school, but instead would use technology, and a system of mentors (who were also learners themselves), to enable learners to recognise and develop their own inspirational qualities, and connect to others who would inspire them further. In this model,

assessment is used to support the individual's own learning journey, rather than as an external reporting mechanism.

3. Following the whole learning journey



This group wanted to recognise and capture the whole learning journey. Rather than seeing this journey as a linear path with a pre-determined goal, their model represents a branching and occasionally circular path of learning. Currently, learning that does not lead to linear progression towards a pre-set goal is often ignored, or seen as failure. This model moved away from rigid notions of success and failure, instead recognising and valuing the learning that can come from an experience of failure.

This idea linked together the learning that happens in 'life' with more traditional learning, such as an apprenticeship or other accredited

course. It would look for links between learning that happens informally, particularly when a learner has to deal with difficult circumstances which might put pressure on their ability to follow a traditional course, and use that informal learning to support the more traditional route. While learning that happens in 'life' is non-linear, academic progression tends to follow a straight route, so support and links would vary depending on the needs of the learner at different times.

4. Connecting diversity

This group represented how schools might more explicitly recognise and connect the different characteristics of learners from diverse backgrounds. Focusing on the moral purpose of schooling as much as the need to develop skills to participate in the labour market, the model here sees developing 'social cohesion' as one of the roles of schools.



HOW MIGHT TECHNOLOGY HELP TO CREATE MORE SOCIALLY JUST ASSESSMENT TOOLS?

While no single technological tool will usher in a more socially just assessment system, technologies may support fairer practices. This section describes a number of such possibilities.

Technologies can be used to capture evidence of learning in a variety of forms. Students can use technologies in creative ways to present their learning in the ways in which they are best able to express themselves.

Storing, organising, aggregating and synchronising large amounts of data is made easier by new technologies, which opens up possibilities for capturing far more evidence and data. This may allow for new assessments that track change over a longer time period, or that take into account a much wider range of factors than is possible with traditional methods.

Artificial Intelligence, and other technologies that respond and adapt to the learner's actions, may be able to capture evidence of learning in entirely new ways that are more accurate and comprehensive. Adaptable technologies may be able to move beyond the one-size-fits-all computerised assessments, to respond more fairly to individual differences and needs.

Because technologies can be used to capture and store information easily, they can be used to record evidence of the entire learning journey, rather than simply the end point. Knowing that the process is being captured can allow learners to take diversions from the straight path towards a single end goal, safe in the knowledge that that learning can be captured and accredited even if they end up somewhere different from their original goal.

Many innovative assessment tools are unwieldy because they are difficult to manage on a large scale. Technology can help 'scale up' new approaches, for example, facilitating large scale peer assessment by breaking down work into smaller chunks that are then anonymously sent by text message or e-mail to peer markers and then later re-combined to give meaningful feedback to the learner.

However, with the exception of AI, most of the ideas and discussion centred not so much around the role of new technologies in creating more educationally just assessment tools and systems, but around the policy and practices of assessment. Technologically, there is very little standing in the way of implementing most of the ideas developed in this workshop. It is instead a question of having the will and the support at a national and local level to start the process of creating assessment practices and policies that support learning in a socially and educationally just manner.

QUESTIONS AND ISSUES RAISED

Assessment *for* learning and assessment *of* learning

While there are potentially problems with many assessment tools, whether or not they offer fair and just assessments also depends on the purpose of the assessment being carried out. Many of the current assessment tools described above are primarily summative – that is, they are assessments *of* learning that measure the learner's achievement at the end of a period of learning. Others, such as peer and self-assessment, are primarily formative; that is they are assessments *for* learning that aim to give an indication of the learner's strengths and weakness to enable both learner and teacher to decide on the most appropriate next steps.

Formative assessment exists primarily to support the learning process through dialogues between teachers and learners and between learners and learners. As such, there is considerable flexibility to tailor it to the needs of the learner, and to give the learner some control over the process. In the sphere of formative assessment *for* learning, there appears to be substantial scope for improving assessment such that it offers a fairer, more educationally and socially just system that supports, rather than stunts, learning.

However, summative assessment is a different matter. It is the key that opens doors to employment and further education opportunities. It attempts to separate out the differences between learners' achievements, allowing others to make judgements and comparisons between individuals. Summative assessment must be reliable, valid, manageable and economic. It may be possible to have two out of these three but not three or four; that is, if an assessment tool is reliable and valid, it can't be manageable or economic. When we add a further condition of social and educational justice, how many of these conditions will it be possible to satisfy?

The question of who owns and controls the assessment process was seen as crucial; with a broad consensus that learners themselves should have far more ownership of the process, and that assessment should primarily be used to support learning rather than making external judgements. However, rather than seeing formative and summative assessment as opposing sides, the answer may lie in finding ways to forge a dialogue between these two purposes, with learners themselves having a say in this process.

Scope of assessment

An interesting debate emerged around the scope of assessment. Because assessment, to some degree, determines what is valued in education, one group proposed an idea that would use Artificial Intelligence to take a more holistic approach to assessment, therefore valuing the whole person.

Rather than measure a narrow set of skills or domains of knowledge, the whole person could be assessed. However, this raises the question of which parts of a person's life should be free from the eye of assessment? As with many other areas, the prospect of AI raises many questions about our values in assessment and education more generally.

Similarly, if we value informal learning, should we attempt to assess it? The very act of assessment would transform the activity to be more formal, so assessing any activity should be undertaken with caution. If we broaden the scope of assessment to include activities such as learning to look after a younger brother or sister and organising social events, should these be compared against traditional academic assessments, and if so, how? Do we even need a system of assessments that can be compared against one another?

Unpredictability of future skills required

Many of the arguments around broadening the scope of assessment to include a wider range of skills and competencies suggest that workers and learners in the 21st century will need a different range of skills from those currently studied. However, there is little evidence that we can reliably predict what those skills will be. If learners are to have a greater say in their learning and assessment, how can we ensure that they are well equipped for future life?

Diversity and standardisation

A fairer assessment system would value and validate a diverse range of achievements, allowing learners to choose what was assessed, and when and how they were assessed. However, in summative assessment, this would make comparison between different individuals more difficult, and subject to the biases of employers and admissions tutors. In some ways, a system that truly was reliable and standardised would allow learners from all backgrounds to be treated equally without discrimination. How can assessment be flexible enough to allow for diversity, and still be reliable and manageable? Should comparison between individuals even continue to be a purpose of assessment?

Assessment in a wider context

Assessment should not be seen as separate from learning and education in a broader perspective; it is part of the whole process. Assessment raises questions about our deepest educational values, and these values need to be seen in the widest global, political and social contexts. The question, then, is whose values are embodied in our education and assessment systems? If the universities had more of a say, would we have a fairer system that valued learning for its own sake? If private enterprise had more of a say, would our education be more tailored to the needs of the labour market? Ultimately, these are political questions that call into question our beliefs about the purposes of education.