The Future of Assessment

2025 and Beyond

The Future of Assessment
2025 and Beyond

Ayesha Ahmed
Anton Béguin
Newman Burdett
Prue Huddleston
Brian Lightman
Tom Sherrington
Alison Wood

Edited by
Dale Bassett

Foreword by
Andrew Hall

About the contributors

Dr Ayesha Ahmed is Research and Teaching Associate in the Faculty of Education at the University of Cambridge

Dale Bassett is Director of Public Policy at AQA

Dr Anton Béguin is Director of Research and Innovation at AQA's Centre for Education Research and Practice, and Director of Measurement and Research at Cito (Netherlands)

Dr Newman Burdett is Head of the Centre for International Comparisons at the National Foundation for Educational Research

Andrew Hall is Chief Executive of AQA

Professor Prue Huddleston is Emeritus Professor in the Centre for Education Studies at the University of Warwick

Brian Lightman is General Secretary of the Association of School and College Leaders

Tom Sherrington is Headteacher of Highbury Grove School

Alison Wood is Head of Curriculum and Assessment Practice at AQA

Contents

Foreword

Andrew Hall

This publication is the culmination of two years of debate, discussion and thinking across the education sector about how assessment – particularly, formal assessments taken by 15–19 year olds in England – can and should evolve over the next decade.

We deliberately set out to consider how the system needs to change in the long term, beyond the 'here and now' of the current reforms to qualifications and the day-to-day pressure of central accountability on schools. This isn't about passing judgement on the current government's policies, or indeed those of any previous government. It's about acknowledging that our imperfect system has served us well for 30 years – but it will, at some point, need to change if it is to continue to support our young people's education and safeguard their opportunities to progress.

Our starting point, then, was to ask: if we started from scratch, what would we want our assessment system to achieve? What would it look like, and how could it ensure that students are equipped with the knowledge and skills they need? And how might we start to get from where we are now to where we want to be?

Since we launched The future of assessment: 2025 and beyond, there has been a huge collaborative effort to answer these questions. At the launch event in November 2013, we asked stakeholders across the system what they thought were the 'big questions' facing the assessment system. We spent the next year bringing together experts from across the country and internationally at public debates, roundtable discussions and workshops to dig into the issues. Teachers, assessment specialists, policymakers and students

if we started from scratch, what would we want our assessment system to achieve?

considered how they wanted assessment to evolve and discussed ideas in person and online.

From this work, three key areas emerged that those across the system thought offered the greatest potential for change: a more effective balance between assessment and school accountability; better assessment of vocational and practical learning; and the opportunity of technology to take the validity and reliability of assessment to new heights. The result of this – a vision for the assessment system in 2025 – is set out on the next page. These six long-term objectives are, I believe, the result of genuine collaboration, and as such represent not AQA's views specifically but a powerful consensus from across the system about how assessment should change to meet the needs of the future.

Our final challenge was to consider how to start to move from today's system to this vision of assessment that works better for everyone. The working group that produced this publication has attempted to answer that question, producing a blueprint of ten things that can start to be done now, to lay the foundations for our vision for the future. It's not definitive, of course – but it is, I hope, a collection of things that could realistically be done in the next few years, to make a difference and put the system on track for achieving our vision for assessment in 2025.

Before the public launch of this programme of work, AQA held discussions with politicians and policymakers, school leaders and teachers, academics and employers, to try to understand what those in the system saw as being the major challenges for assessment in England and the opportunities for change. I was struck at the time that the one theme that repeatedly emerged from these conversations was trust: the lack of trust in the system between teachers, school leaders, exam boards, government, the regulator, parents and, crucially, young people.

So I think it is significant that many of the ten actions in our blueprint are largely intended for those within the system – notably,

teaching professionals and assessment experts – rather than government. I think this demonstrates that we have a system that understands how things can change and is ready to lead that change, instead of waiting for central imposition of policy. Better assessment, and better education, for our young people is there for the taking.

In 2025...

Assessment in 2025: a vision for the future

This vision sets out six long-term objectives for the future of assessment – this is how stakeholders across the system would like assessment to be in 2025. The objectives are a representation of a wide range of views, building on the valuable input and views of education professionals and the general public shared through the discussions that formed the basis of this project.

In 2025…

Balancing assessment and accountability
- Professionalised, expert teacher assessment contributes substantially to students' results
- School accountability supports improvement and is based on more than exam grades

Assessment for the real world
- Employers, teachers and assessment experts work together to determine outcomes and standards
- Credible qualifications demonstrate value through exemplification of skills and evidence of progression

21st century assessment
- Technology does not replicate paper exams but facilitates new, more valid types of assessment
- Skills that cannot be assessed summatively are captured 'in action', using technology where possible

Towards the future: a blueprint for the next five years

These ten actions, which are underway already or could be started now, are useful, achievable steps that will begin to move the system towards the long-term vision set out above.

Assessing 'hard to assess' skills

- The assessment community should work to develop valid assessments of complex, interactive tasks such as collaborative problem solving, with the aim of assessing the process as well as the outcomes of these tasks.
- As more valid and reliable assessments of 'hard to assess' skills are developed, government, Ofqual and exam boards should consider how these could be incorporated into qualifications in the future.

Assessing vocational and practical learning

- Exam boards should work with training providers to develop valid, authentic assessments of vocational skills. Using technology where appropriate, assessments should be developed that can capture the 'journey' as well as the final outcome of the candidate's work.
- As more authentic and real-time vocational assessments are developed, government should work with Ofqual and the exam boards to consider how the qualifications and

accountability frameworks can support the use of these more valid assessments.

Improving teacher assessment

- Teachers, school leaders, teacher training providers and the assessment community should work together to equip all teachers with assessment expertise.
- The teaching profession should help to lead the creation of an assessment ethics framework to develop increased trust between teachers, government, Ofqual and the exam boards.

Getting more out of test results

- Exam boards should work to develop rich assessments and reporting tools to inform teaching and learning and school improvement.
- As richer assessment becomes used to support robust school self-evaluation, government should consider how the accountability system could evolve to reflect this and focus on supporting school improvement.

A National Baccalaureate for England

- The National Baccalaureate Trust should support and build a system-wide alliance to establish a National Baccalaureate for England.
- As the system-led National Baccalaureate gains traction, government should consider its implications for the curriculum, qualifications, accountability and funding frameworks, and how schools and colleges could be supported to deliver rich, broad Baccalaureate provision for every young person.

What is good assessment?

Newman Burdett

While this seems, on the surface, to be a simple question it is actually a very complex one – and one that failure to understand and answer at a policy level has led to a raft of problems. To take just one example, the problems with GCSE English in 2012 were caused in part by poorly thought-through assessment policy, and compounded by unintended consequences of accountability measures (House of Commons Education Committee, 2013).

Looking to the future of assessment, the principles of good assessment are unlikely to change as exams or qualifications change. We might be getting better at assessing what we want to test, we might be getting more efficient at how we do it and technology might be allowing us to assess things that previously we would have struggled to assess but the fundamentals of what makes good assessment will not change.

In a large part this is because good assessment cannot be divorced from good education. It is very important to get both right, and to understand the complex interplay between them, because an education system that works well brings benefits at many levels. For individuals it can fundamentally alter life chances; for nations it can mean the difference between boom and bust (South Korea is a well-studied example of improving education leading to economic benefits).

Given the complexity and debate that surrounds education and assessment, nationally and internationally, it is hard to say categorically what 'good assessment' is – values and cultural influences blur the borderlines – but we can state that good educational assessment needs to meet some basic criteria.

Firstly, the assessment needs to have a clearly defined purpose. There is no point assessing needlessly or placing unrealistic or potentially conflicting demands on the assessment. Experience teaches us that there will be counterproductive tensions if a single assessment is used to monitor national standards and act as an accountability measure for teachers. If we do not understand why we

are asking the question, we will struggle to understand the answers to that question.

Secondly, it must be fit for that purpose: it must measure what we want learners to learn. It is good that Ofqual is now focusing on the validity of assessment to ensure that the results are meaningful, useful and appropriate rather than just repeatable (Stacey, 2014). Good assessment needs to reflect everything that we consider important to a good education – it is not a case of *if* important things should be assessed but *how* they are assessed. Good assessment should start from the intended learning and does not mean valuing only what we can measure well, but finding ways to measure what we value.

Most importantly, and often most overlooked, good assessment should borrow from medicine the principle of *primum non nocere* – it should do no *harm*, in this case to the learners.

Many of these issues are quite nuanced in England because, contrary to common misconceptions, I think learning and assessment are actually quite good and our education system staffed by a lot of dedicated and skilled professionals. It is by no means perfect but anyone who wants to argue should first look at how the education system is respected and valued overseas – in 2011, education exports were worth £17.5 billion to the UK (BIS, 2013).

It may be easiest to exemplify what difference 'good' assessment can make by looking at a context where instances of 'bad' teaching and 'bad' assessment have highlighted the impact that a well-designed and well-run assessment system can make. The example of Pakistan throws the contrast between the two into stark relief (Greaney and Hasan, 1998).

A group of local head teachers became disillusioned with preparing their students for a system of rote learning, driven by narrowly-focused exams. They observed that the activity taking place in many classrooms could be characterised as chant-response

Good assessment does not mean valuing only what we can measure well, but finding ways to measure what we value.

lessons where even comprehension exercises were learnt off by heart. The group of head teachers felt this approach did not encourage 'deep learning' and resolved to do something about it. They approached the Aga Khan University and worked together to set up an independent exam board – the Aga Khan University Examination Board (AKU-EB). This new independent body started producing exams that more effectively tested the whole curriculum, with a particular focus on higher order skills and other harder-to-assess skills such as practical work. A fundamental, underlying aim was to encourage better teaching practices.

NFER recently conducted an evaluation of the AKU-EB's first 10 years in schools and found the impact of the relatively simple change of improving the assessment system was transformational – on learning and on the whole school atmosphere. This was the case even in low-resource schools in poor neighbourhoods or remote regions. A quote from a ministry official interviewed as part of the review exemplifies how this visibly improved learning outcomes:

"I was in a small remote area of Pakistan [in Khyber Pakhtunkwa Province]. There the AKU Board provided the examination system over there. And I realised that the people over there are much more intelligent than the people in Karachi who are taught by the [traditional] board. … I tell you it was amazing because when you go to a remote place in your country and you feel that the people would be less privileged, but I found the people, the school children very intelligent. So I think in this place I found the students over there very smart, smarter than the people, the kids and the students I had come across in major cities like Karachi, Islamabad, wherever, different [traditional] examination boards are going on."

During our fieldwork we conducted research in a school where

there were both AKU-EB and traditional examination board students present studying in different streams. All the girls were from similar backgrounds. To the interviewers they were all clearly bright students who wanted to study engineering and medicine at prestigious universities in Pakistan. However, the difference in confidence between the AKU-EB students and the traditional board students was striking and obvious even within the first minutes of meeting the group. The AKU-EB students were able to articulate their thoughts clearly, could engage with the interviewers (and clearly enjoyed doing so), could respond when probed and challenged about their views, and were confident in expressing their opinion. They took the lead during the tasks they were set. In contrast, the traditional board students were polite, reserved and passive. They were unwilling to voice their opinions even when asked directly and, when discussing the tasks, they deferred to the AKU-EB students. The traditional exam board students were conscious of this difference and one explained:

> *"[The AKU-EB students] are different in many respects. They are more confident than us. They are more creative. They can speak in front of anyone."*

We ascribed these differences as being due to the students being taught in a different style, driven by a different examination system. We looked at not only the immediate impacts (the AKU-EB students were more likely to score highly on university entrance tests than students from traditional exam boards with similar grades) but looked at the students who had taken the AKU-EB and were now at universities and found that there was clear evidence that they were better prepared for the demands of higher education and found the transition easier.

This powerfully illustrates that if you get the assessment right, and aligned with the education you want, it can make huge

differences. Good assessment can have long-term positive impacts on learning and on students' life chances.

So what was it that the AKU-EB system did that the traditional examinations system did not do?

Firstly, in designing their curricula and assessments they took an integrated view with assessment as a key part of the learning experience: they had clear purposes for their assessments and made sure schools, and teachers, understood them. This was not by any means perfect and many schools are still struggling to adapt their teaching (and teachers) to what is a more challenging style of teaching and one that requires greater pedagogic skill and assessment literacy. Secondly, AKU-EB ensured the assessments were fit for purpose: they concentrated on trying to validly and reliably assess the skills that they wanted their students to learn. They looked at what children, with a wider view than just passing exams, needed to learn and designed assessments that should support and challenge them to learn. Again, our research in Pakistan showed that this was not perfect and that in some aspects, especially the school-based assessments, teachers were struggling to deliver the change, but it was a big improvement on what had been previously available.

This example is from a relatively extreme case study but these lessons do not just apply to systems in crisis; they apply to all education systems, including England's. Getting the assessment right is so important. Assessment has a huge impact on teaching and what actually happens in the classroom. Good assessment practices can lead to real improvements in learning just as poor assessment can undermine good teaching and learning.

In educational debate it often feels that assessment is something that is *done* to education. It is not – assessment is an *inherent part* of good education and learning. Teachers and students need to be able to accurately assess and understand their learning if they are going to improve. There needs to be good monitoring and formative assessments in addition to certification and accountability,

and all of these are important. It is crucial that all the different types of assessments are compatible and work together. For this to happen there needs to be a set of well-defined learning outcomes which are linked to standards (and vice versa). In this way classroom practice and assessment will be complementary. A lot of excellent work, backed by good research, has gone into how assessments within the classroom can improve learning. The work of the Assessment Reform Group[1] has received international acclaim, but they concluded that a lack of clarity or understanding of fitness for purpose has led to some strange outcomes. Currently school-based assessments are an excellent example of how the false dichotomies between learning and assessment (or validity and reliability) have ended up with a system of controlled assessments and coursework that struggle to be either educationally beneficial or useful assessments.

To come back to our original question – "what is good assessment?" – possibly the most important thing that makes for good assessment is for everyone in the system to understand that assessment is as complicated and complementary to learning as good pedagogy. Good assessment is inextricable from good learning. There are no simple fixes but we have a huge range of good research that will allow us to get it right and we need to make sure that that knowledge and expertise is used.

Reference

1. www.aaia.org.uk/afl/assessment-reform-group/

Assessing 'hard
to assess' skills

Ayesha Ahmed

Working collaboratively in groups is a crucial skill in the 21st century, information-based economy. Education systems must respond to the changing landscape of employment and one of the ways to do this is to ensure that children are taught the skills for effective group discussions and collaborative problem-solving. We know from reports by organisations such as the London Chamber of Commerce that employers and businesses expect children to leave school with these skills (Wright et al, 2010), yet we are not currently emphasising them sufficiently in classrooms – or in assessment.

To improve the teaching and learning of group discussion and problem-solving skills we need to consider how we can monitor and assess these skills. Good assessment has a positive backwash effect on teaching and learning, raising the profile of the skills being assessed. But the imperative to assess these skills is greater than just the possible backwash effect. In order to consider the assessment of these skills we must consider what we mean by these skills: what is the construct of collaborative group work? Understanding the nature of the construct will further our understanding of how to improve learning of these skills.

I will define collaborative group work as *students working in groups to have effective discussions and to solve problems*. There is much work to be done, though, on identifying the important skills that make up this construct.

Three critical questions follow:

1. What sorts of tasks will allow us to observe behaviours that show evidence of effective group discussions and problem solving?
2. How can we ensure that our assessments are as valid as possible, measuring the intended construct in a fair and informative way?
3. How can we use the results of the assessment to inform the teaching and learning of these skills?

The construct as I have defined it relates to group discussion as well as problem solving. This reflects one of the issues with assessing children working in groups. Should we assess the process of the group work (often a discussion) or the products (the solution to a problem) or both? I believe that we should endeavour to assess both.

Another critical question to consider when mapping out the skills that make up a construct is to ask what progression might look like for these skills. How do we decide if a learner has progressed? Sometimes the answer will be that they are showing more evidence of a particular skill but sometimes there can be a qualitative change in the way the skill is manifested in a performance.

Assessing collaborative problem solving

The OECD has recognised the importance of collaborative problem solving and this will be assessed for the first time in the PISA 2015 tests (OECD, 2013). In these tests children work collaboratively to solve problems with a computer agent.

The OECD defines Collaborative Problem Solving (CPS) as:

"… the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution." (OECD, 2013:6)

In these tests the test-taker and a computer agent communicate via a chat box. The test-taker is assessed on his or her actions, communications with the agent and the products of the collaboration (solution to a problem). Using a computer agent is one way to address concerns of reliability and fairness as well as the manageability of such a large scale high-profile assessment. However, it ignores the critical social element of group work.

The Assessment and Teaching of 21st Century Skills Project (ATC21S), based at the University of Melbourne and sponsored by large corporations, involves Australia, Finland, Portugal, Singapore, England and the USA. Griffin, McGraw & Care (2012) and Griffin & Care (2015) report on the progress of this large-scale long-term project.

One of the main target areas of the project is the large-scale assessment of collaborative problem solving which is being trialled with 11–15 year old learners. ATC21S considers how to measure the social skills of participation, perspective taking and social regulation, as well as cognitive skills (Care et al, 2015). They are trialling content-free tasks, measuring inductive and deductive thinking skills, as well as content-dependent tasks designed to measure knowledge.

Considering the process of problem solving and the social skills involved is an important part of ensuring the validity of an assessment of such collaborative work.

Assessing group discussions

What are the skills that make effective group discussions? Mercer et al (1999) discuss 'ground rules for talk' and show how group discussions that follow these rules can be beneficial to learning. Resources to help teachers and children to achieve this are provided on the Thinking Together project website.[1] These include 'ground rules for exploratory talk' such as "everyone in the group is encouraged to contribute" and "everyone is prepared to accept challenge". These ground rules give us a useful way to consider how to assess group discussions. We can ask how well students are using these rules and look for evidence of this.

The Oracy Assessment Toolkit,[2] mentioned in a recent Commons Education Committee report (2014), identifies the set of skills that learners should develop to be good at using speech to communicate effectively with others. Many of these skills are relevant for assessing children's contributions to group discussions.

It is important that the difficult issues of the social dynamics of group work are not ignored, but are incorporated in some way into the assessment.

An understanding of the skills that make up the construct we hope to assess is a necessary starting point. However, when assessing children in groups there are a number of other factors to consider:

- Should we assess the process or the product of group work or both?
- Should we only assess individual students' performances?
- Should we also assess the performance of a group as a whole?
- Should peers assess each others' contributions to the group?
- Should students assess their own contributions to the group?
- How do students learn from each other during the assessment?
- What are the effects of the content or context of a task?
- How do we decide on group composition?
- Should children be allocated specific roles in advance?

These are all interesting questions that must be considered during any task design and trialling of assessments of group work. It is important that the difficult issues of the social dynamics of group work are not ignored, but are incorporated in some way into the assessment.

Wilson et al (2012) suggest defining the construct for any group task by focusing only on aspects of performance that operate in a group situation. In this case students would be given scores based on the quality of the outcomes of their particular role (e.g. facilitator, leader). In order to be able to generalise, students must then be given the opportunity to perform in a variety of groups and in a variety of contexts, and these performances should be aggregated. This kind of sampling over a number of groups and contexts can begin to address reliability concerns.

Wilson et al (2012) also suggest that students can be given an individual performance score and a group performance score, and statistical techniques can be used to look at the relationship between these scores. They go on to discuss the possibility of group members giving feedback on each individual's performance. This could add useful information to an assessment of group work but it must be done in a classroom atmosphere of mutual trust.

Blatchford et al (2003) discuss some of the necessary conditions for this, including developing respectful relationships and making the physical environment conducive to group discussion. If group work is seen as a normal classroom activity then assessing this will be easier as students will be used to the idea of working together and supporting each other.

There have been some recent criticisms of group work in classrooms (e.g. Peal, 2014) that argue that it is not effective, but as Mercer (2015) points out, group work can be unproductive when children are not taught how to do it well. If appropriate ground rules are used then it can be highly beneficial to learning.

Assessing the process of group work is likely to be more

difficult than assessing the products. Methods for observing and coding or scoring elements of the process would need to be trialled and concerns of reliability and bias must be kept in mind. By sampling performances in different situations and with trust in the judgements of teachers, some of these concerns can be addressed. Furthermore, when students work collaboratively to solve problems on a computer, a wealth of data can be collected including detailed activity logs (Von Davier & Halpin, 2013). We are only now beginning to understand how such rich data can be used, but this increase in information about group processes can open up new possibilities for assessing these skills in a valid and reliable way.

Moving towards assessing 'hard to assess' skills

It is clear that the question of how to assess children working collaboratively in groups is not an easy one to answer. These are hard skills to assess. If something is hard to assess should we stop trying to assess it and concentrate instead on the well-trodden paths of assessing knowledge and understanding in written examinations? To some extent this has occurred. For example, from summer 2014 the speaking and listening component of GCSE English was removed from grading so that students' performances on this critical element of English do not count towards their final grade. Part of the reason for this change is that these skills are hard to assess in a way that is accepted as valid and reliable. However, there are also examples of successful assessments of hard to assess skills such as the Extended Project Qualification.

My view is that we cannot afford to ignore the importance of the skills required for effective group discussion and problem solving. These are important skills for today's children to leave school with. They are skills that can improve their learning as well as prepare them for life beyond the classroom. One of the ways to ensure that these skills are taught well is by designing good quality methods for assessing them. The assessment community, including

teachers, exam boards and researchers, should work together to try to develop valid assessments of complex, interactive tasks such as collaborative problem solving, with the aim of assessing the process as well as the outcomes of these tasks. As more valid and reliable assessments of 'hard to assess' skills are developed, the government, Ofqual and exam boards should consider how these could be incorporated into qualifications in the future.
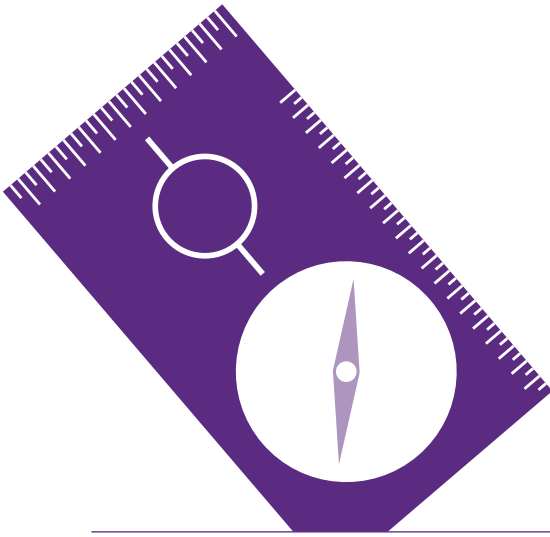
If we assess only what we can easily measure then we end up defining constructs in terms of what can be assessed rather than what is important in a domain. This could have a devastating effect on the richness and breadth of learning in the classroom. If something is hard to assess we should not simply avoid assessing it, but instead we should look for better ways to assess it.

A blueprint for assessing 'hard to assess' skills
- The assessment community should work to develop valid assessments of complex, interactive tasks such as collaborative problem solving, with the aim of assessing the process as well as the outcomes of these tasks.
- As more valid and reliable assessments of 'hard to assess' skills are developed, government, Ofqual and exam boards should consider how these could be incorporated into qualifications in the future.

References
1. https://thinkingtogether.educ.cam.ac.uk/
2. http://www.educ.cam.ac.uk/research/projects/oracytoolkit/

How should we assess vocational and practical learning?

Prue Huddleston

Assessing vocational and practical learning presents a number of challenges since there is a range of actors who have a legitimate interest in the outcomes: employers, teachers and trainers, learners, assessment experts and funders. However, these individuals may have different views on what counts as reliable, valid and credible evidence of achievement. Unlike assessment designed for academic purposes, mainly characterised by written examinations and tests completed under timed conditions, vocational and practical assessment also requires learners to demonstrate their skills and knowledge through performance in the workplace, or within simulated workplace conditions. "It is possible for someone to have a mental grasp of procedural knowledge, but no ability to perform the given task" (Carr, 1999). Clearly such an outcome would be regarded as deficient in terms of learners' employability.

This complexity is compounded by the fact that within the English system of vocational education and training (VET) the term vocational is used in at least three different ways: workplace vocational education, which focuses upon occupational practice and performance (Billett, 2006); general vocational education, which provides a general introduction to a vocational or occupational area, characterised by many full-time programmes offered to 14–19 year olds (Pring, 2007; Pring et al, 2009); and pre-vocational programmes, which seek to develop employability skills in those deemed not yet ready to move into the world of work and include elements such as literacy, numeracy, and work preparation (Acquah and Huddleston, 2014).

This typology provides a continuum through which to view a diverse range of vocational offerings, but raises challenges for those attempting to assess performance. For example, what type of assessment can reliably measure a candidate's 'employability' (predictive validity) since the assessor is being asked to predict future behaviour and performance, in different contexts and under altered conditions? Who is qualified to make assessment judgements about

the taste of a soufflé, the 'tone' of a presentation, the welding of a joint, or the candidate's ability 'to work in a team'? Clearly, a range of assessment techniques is required to give assurance of what learners have achieved and of their competence, often within a workplace context.

Designing robust assessment for vocational and practical learning involves a consideration of three important and inter-related dimensions: 'learning that', 'learning how' and 'learning where'. It involves the development of knowledge, skills and behaviours that require a range of assessment techniques suited to the inter-connected purposes of that learning. This cannot be achieved simply by testing one dimension, but all three: knowledge, practical skills and their application in the workplace or 'real life' contexts. What is required is an integrative approach that permits a combination of methods – theoretical and practical – to build confidence in qualification users. Instruction, learning and assessment need to be compatible, because the three elements interact with one another rather than operate in isolation. In vocational and practical learning 'making sense' of theory often involves applying that theory in practice, reflecting on it and confirming, or changing, behaviour as a result (Cooper and Ord, 2014).

The recent predominance of a target-driven climate within the English education system has resulted in an assessment-dominated culture in which teachers and learners can become more interested in the assessment and its outcomes than the learning. In the worst cases assessment has become so atomised, de-contextualised and summative as to make an understanding of the whole almost impossible. This is a particular danger for vocational and practical learning when there is insufficient linkage between theoretical aspects and their practical application; the practical realities of workplaces are messier than theory suggests. This is exacerbated by the fact that this learning often takes place in different settings: schools, colleges and workplaces, each with their own cultures,

communities of practice, and preferences.

Therefore, effective assessment requires engagement from a range of actors, each with distinctive contributions and expertise in order to give credibility to what is offered and to provide "an integrated account of knowing, acting and being" (Barnacle, 2009).

What would good assessment look like?

Much of the recent debate concerning the need to align vocational qualifications, and by inference their assessment, more closely to the needs of the labour market has focused upon the perceived gap between what is taught, learnt and assessed in school or college, and what is needed at work. Contemporary workplaces, and the nature of work itself, are changing so rapidly that different ways of learning and assessing are necessary. Skills require regular updating; their development cannot be curtailed by the pace of the academic year and written tests. Learning is highly context-dependent; even within the same sector different ways of working may be required, different tools and materials used, and learners need to be able to apply their knowledge across a range of different jobs and contexts. Assessment has to be sufficiently flexible to manage the 'here and now', not what pertained last year. This makes a strong case for integrating assessment within the learning process, whether that learning happens in the workshop, office, laboratory or classroom.

Wiggins (1989) has suggested that authentic assessment tasks should truly reflect the performance expected within the field (face validity); permit learners to reflect upon and make changes to performance as their work develops (formative, continuous and self-assessment); and allow students to present and defend their work, not just write about it, to ensure that mastery is genuine. Eisner (1993) extends this list of requirements to include, inter alia, the need for authentic assessment to reflect the norms and values of the communities from which tasks are derived. That is, it is not only important to be able to accomplish the task, but to perform it in a

manner acceptable to the professional community.

The work of Gulikers et al (2004, 2009) provides a particularly helpful contribution to our understanding of authentic assessment. They characterise it as comprising five inter-related features: the five-dimensional framework. The basic principles on which the framework rests may be described thus: "an assessment that requires students to use the same competencies, or combinations of knowledge, skills and attitudes that they need to apply in the criterion situation in professional life" (Gulikers, 2006).

When planning for authentic assessment these five inter-related 'dimensions' need to be considered: the assessment task (is it a real workplace activity?); the physical context in which it takes place (is it an actual workshop, studio, laboratory?); the social context in which it occurs (does it mirror the social conditions of professional practice, including supervisors, other workers?); the expected output (for example, an artefact, report, a process undertaken to professionally acceptable standards); and the criteria against which these four aspects should be judged must relate to realistic professional practice in professional contexts.

How can this be achieved?

Given this diversity, it is clear that assessment for (not just of) vocational and practical learning must include a range of activities suited to the specific needs of the learner and of the context in which they operate. A 'one hit and you're out' mentality does not reflect the way in which knowledge, skills and competence are built up over time. It is inconceivable that a crafts person would achieve competence on the first try. A range of assessment methods is required to cover the array of learning events and activities that contribute to vocational and practical learning and to its ultimate accreditation.

Recent, and continuing, developments in technology provide novel and engaging ways of capturing 'learning in action'; the use

A 'one hit and you're out' mentality does not reflect the way in which knowledge, skills and competence are built up over time.

of electronic logs can chart progress over time; practical skills tests under controlled, but realistic, workshop conditions can demonstrate competence. The quality of artefacts made, or performances given, by learners can be assessed by vocational experts. Often these experts will not be vocational teachers in the traditional sense, and nor should this matter. The boundaries between classrooms and workplaces should be permeable.

Assessment 'on demand' for certain types of test is a legitimate, and achievable, request for vocational learners and employers, given the range of technological tools available. Research has shown it to be highly motivational for those learners who have been 'frozen out' by traditional annual examinations, and employers do not want to wait a year to know if potential employees are qualified (Acquah and Huddleston, op.cit.). Accreditation should be available as soon as possible after the completion of the assessment activity. Written tests are normally administered in a restricted and dedicated period of time, for example one hour. In the workplace, professional activities may spread over days, or may require fast and immediate reaction. An agricultural instructor recently remarked: "isn't it strange how the tractor's wheel always seems to come loose in a dark and muddy ditch and never in the controlled conditions of the college workshop?"

Assessment within vocational and practical learning can be viewed as a continuum in terms of its levels of complexity. At one end of this continuum are fairly straightforward and routine activities, which might be covered by trade tests for example, drawn from a bank of items agreed by sector representatives and providing an accessible form of on-demand assessment. At the other end of this continuum are those complex activities that involve substantial tasks requiring learners to demonstrate a combination of knowledge, skills, professional - and sometimes personal - attitudes and behaviours within very specific contexts, for example a product design 'live project' lasting several months.

In the latter case a range of assessment techniques is required, and should be accommodated, to allow for the interplay of 'learning that', 'learning how' and learning where' to be properly recognised. There is room for far more creativity in the design of assessment for vocational and practical learning; technology can be harnessed to help achieve this, as it is already in many workplaces. Multi-national companies have replaced international meetings involving expensive foreign travel with video conferencing and electronic transmission of data. The emphasis on the regulation of assessment within the English VET system has stifled innovation because too often alternative approaches have been viewed as tantamount to cheating. Assessment of practical activity has been regarded as too subjective and therefore unreliable and, ipso facto, of little educational merit.

Along this continuum is a whole range of programmes and courses, as described in the opening section of this chapter, which are barely vocational and certainly not very practical. These will require specific approaches to assessment that should be well matched to the purposes for which the programmes are designed – pre-vocational or general vocational.

Who says so, anyway?

It has been recognised that the assessment of vocational and practical learning is a complex, sometimes messy, and rapidly changing undertaking. Many actors have a legitimate stake in its outcomes at micro (individual), meso (organisational or institutional) and macro (national) levels. Sometimes the interests of stakeholders are in conflict since their expectations of what achievement should look like can vary, for example: 'Will I get a job?'; 'Does this apprentice have the right practical skills to operate in my workplace?'; 'Will this candidate be able to progress to further study in this vocational area at a higher level?'; and 'Who says so, anyway?'

What is distinctive about assessment within vocational and practical learning is that it should afford opportunities for learners to

bring together theoretical knowledge, practical skills, attributes and attitudes that demonstrate competence within a professional context and to professional standards. Assessment needs to capture the journey taken to reach that achievement and not rely solely on the final outcome.

A blueprint for assessing vocational and practical learning
- Exam boards should work with training providers to develop valid, authentic assessments of vocational skills. Using technology where appropriate, assessments should be developed that can capture the 'journey' as well as the final outcome of the candidate's work.
- As more authentic and real-time vocational assessments are developed, government should work with Ofqual and the exam boards to consider how the qualifications and accountability frameworks can support the use of these more valid assessments.

Assessment in a self-improving system

Brian Lightman

The desks in the classroom have been rearranged into a square in preparation for a meeting chaired by the headteacher. The purpose of the meeting is to moderate the final examination essays and award grades for the coveted Abitur qualification in English, in a German secondary school. All of the English teachers, a representative of the education authority and a rather bemused young foreign-language assistant are present. For a couple of hours we scrutinise the students' work against the mark scheme and assessment criteria designed and set by the teachers and approved by the external education authority. As foreign-language assistant, my advice is sought on the English expression used in the essays. Finally grades are awarded taking into account the large amount of assessment data gathered by the teachers in the course of their work with students over the last year or two.

The closest comparison to this experience is of a medical case conference. In serious, detailed discussions professionals use their expertise to reach an evidence-based conclusion underpinned by a very clear methodology and an ethical framework.

At the time, long before I trained as a teacher, I naïvely assumed that this is the way teachers operate. Although my teacher training did briefly explore the topic of assessment and the design of appropriate tasks in the classroom it certainly did not include the master's level study of the theory and practice of assessment that my German counterparts had completed. That had to wait until much later in my career and be sought out on my own initiative.

During those early years, however, the curriculum was drawn up by the school. As a modern languages teacher I became heavily involved in the graded objectives movement. Teachers from a number of different schools worked together meeting after school in classrooms and teachers centres. We developed programmes of

study and assessments based on short-term objectives leading to certificates at a range of levels the highest of which were externally accredited by awarding bodies. With the help and support of specialist advisers and experienced colleagues we implemented a motivating curriculum into which assessment of all four skills – speaking, listening, reading and writing – was fully integrated. Many aspects of that good practice were incorporated into the then new GCSE examinations which were largely welcomed by teachers.

It seems to me that the two experiences I have described had one thing in common: trust. Professionals were trusted to develop and implement programmes of study and assessments which were in the best interests of their students and were indeed expected to do so responsibly and ethically.

The other common characteristic was that it was the curriculum rather than accountability which drove the assessment practice. But this was certainly not a golden age. In England at that time there was too little accountability and far too many students were completing their education with little to show for it.

Since then of course the world has changed almost beyond recognition and I certainly do not want to depict all of those changes as bad. It was vitally important that the progress of pupils was systematically recorded and it was equally important that all students wherever they were in the country had access to a similarly broad range of curriculum options. This is what the National Curriculum set out to do and, in comparison with what had happened before its introduction, it was largely successful. Teachers spent a great deal of time learning the language of levels which were designed to make summative assessments at the end of each key stage. However, in my view, two things went badly wrong. First, the original purpose of levels was derailed in many cases into a tick box process dominating teaching in ways in which it was never designed to do. Second, the skill of curriculum planning was lost by many teachers as the curriculum was externally defined.

The largest change, however, was the introduction of performance tables and high-stakes accountability. Many of the indicators used brought with them perverse incentives and unintended consequences. Rather than being the servants of the curriculum, assessments in too many cases became the master. The stakes attached to examination results and the coveted C grade were simply too high for schools to ignore. This led to the destructive discourse of 'gaming', which blamed the teaching profession for responding to the incentives imposed by successive governments. Most forms of classroom-based assessment have now been excluded from externally-awarded qualifications.

Recently there has been a greater recognition of this problem with moves towards more intelligent accountability measures such as Progress 8. But the new secondary National Curriculum is still dominated by the content of external examinations, and there is a real risk that it will remain impoverished by an unhealthy emphasis on the test. The teaching profession is a long way from knowing exactly what learning outcomes will lead to a specific grade in the new GCSEs. The assessment tail really is wagging the curriculum dog.

But in spite of this there is real light at the end of the tunnel. As expressed in ASCL's blueprint for a self-improving system,[1] which has been widely welcomed throughout the education community and across the political spectrum, we are on the cusp of a golden opportunity for the profession to step up and drive our education system into a better place. There is a growing recognition from politicians that they need to step back and take a more strategic role as we rise to this challenge. In order to make this a reality there are three things we must do.

First, we must ensure that the starting point for curriculum planning is the agreed vision for the education to be provided in that school. In our blueprint we have proposed a broad, nationally-defined core curriculum framework. While politicians will rightly always have the final say, ASCL has argued that this framework

an over-reliance on external testing and a low-trust culture risks driving the curriculum into the straitjacket of what can be assessed in a written test

should be determined by an independent advisory commission for curriculum review with representatives of school leaders, governors, teachers, parents, employers and politicians. The framework would be reviewed once, and once only, every five years, and would draw on objective evidence drawn from the best research analysed and brought together by an independent national evidence centre of education. We propose that this body could be funded by endowment, feeding national and international evidence of best practice into the policymaking process at national level, and the professional practice of teachers and school leaders. Such a body would be independent of both government and the profession. Beyond that, schools would build their own curriculum, bringing creativity, dynamism and relevance into curriculum development.

Second, we must equip all teachers with the skills and a profound understanding of the theory and practice of assessment. I was very pleased that Andrew Carter's report on initial teacher training[2] called for a national framework to ensure that all teachers are educated in the body of knowledge that underpins professional practice. The science and methodology of assessment must be part of this so that our profession is not further deskilled by an over-reliance on external testing and a low-trust culture that risks driving the curriculum into the straitjacket of what can be assessed in a written test. Part of this solution is giving all teachers access to professional knowledge and skill in assessment. Chartered assessors – teachers skilled to work across schools to challenge, support and develop classroom practice with assessment at the heart – are an example of how this can be achieved, and it is a longstanding ASCL policy to support this initiative.

Third, we also need a profession-led assessment ethics framework, which would help to end the destructive discourse of 'gaming'. ASCL is very pleased to be involved in work Ofqual is undertaking on such a framework.

The importance of this has nowhere been demonstrated more

vividly than in the recent heated debate about practical science. Schools do not need to blindly follow decisions that have been made about testing. If our vision for the curriculum says that science practicals are important then we must incorporate them into our teaching, rather than blindly narrowing our curriculum to exclude practicals simply because they do not contribute to the exam grade.

We need a deep-seated culture change in which we stop asking, 'what can be done by others in order to ensure that our curriculum meets the needs of our students?' and instead ask, 'what can *we* as school and college leaders do to address this?'

Assessment in a self-improving system will refocus attention on the benefits of strong formative assessment and effective feedback, which we know can boost learning by an extra nine months in an academic year. It will be based on fewer things in greater depth, and in understanding how a curriculum and assessment model based around mastery and 'going deeper' sits alongside an accountability model that prioritises progress. It will be developed by schools working together to develop robust approaches to progress-tracking.
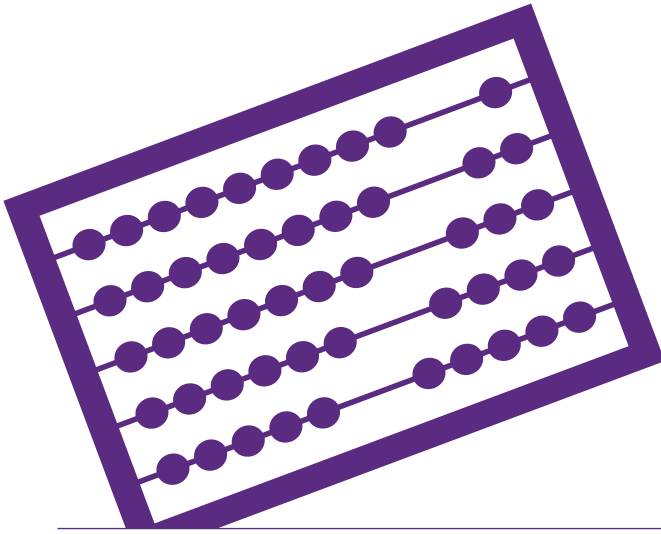
Our profession must be in the driving seat. As Joel Klein says, "You can mandate adequacy; you can't mandate greatness. It has to be unleashed." It is up to us to drive forward assessment practice in ways that help us to take our education system to new heights.

A blueprint for improving teacher assessment
- Teachers, school leaders, teacher training providers and the assessment community should work together to equip all teachers with assessment expertise.
- The teaching profession should help to lead the creation of an assessment ethics framework to develop increased trust between teachers, government, Ofqual and the exam boards.

References

1. http://www.ascl.org.uk/news-and-views/blueprint-for-selfimproving-system/blueprint-for-selfimproving-system.html

2. https://www.gov.uk/government/publications/carter-review-of-initial-teacher-training

Using exam results to inform teaching and accountability

Anton Béguin & Alison Wood

The current approach to the use of test or examination results at school level is strongly externally focused. As a consequence, they cannot clearly be seen to support school improvement and enhance the professional development of teachers.

Test results are, at present, used as a key indicator of how well schools are performing. In many cases, they are clearly seen as the most important one. They are the central piece of information upon which accountability measures are based.

As indicators of the quality of education, however, this type of measure is necessarily limited. A summative test will try to assess the relevant content of a subject, but the need to keep testing time-manageable and the limitations of test logistics mean that there are restrictions on how much content can be assessed. A single test can only ever take a sample from the content of the subject. Due to this limitation, a test might well have a certain degree of predictability ("Well, there is always a question on x and another on y … and they can't really ask about z, so we won't worry about that too much …"). As a consequence, it might reasonably be argued that there is a lot more to education than can be measured in a test.

When used for accountability purposes, test results are often analysed at an aggregate level, say as average scores or grades for a school. School performance is then evaluated in comparison with other schools. Sometimes comparisons are made between schools which are similar in terms of background variables, such as the prior attainment or socio-economic background of their intakes.

This approach to evaluating a school has value as a high-level indicator of quality, but it is too shallow to be used to support improvements in teaching and learning. In a worst-case scenario, test results are given so much emphasis that they can lead to unintended and undesirable consequences, such as teaching to the test, narrowing the curriculum and focusing on those students whose performance has the greatest impact on the headline accountability measures. All this potentially places risk on the validity

of the accountability system and might even undermine trust in schools and teachers.

This chapter gives a brief overview of different types of assessment and their suitability for accountability purposes, and explores the potential for a somewhat different approach to using test results in accountability. This approach is not a radical departure from current practice, but an adaptation of the current approach, which builds on its strengths and begins to address its more major weaknesses. For this reason, we will refer to our approach as the 'adapted approach'.

The adapted approach builds on the current approach, because it retains test results as the core of the accountability system. The difference lies in the way in which the results are used. The adapted approach aims for a more process-based accountability system, in which test results are used for school self-evaluation, evidence-based improvement and school-driven accountability.

Different types of assessment and different uses of results

It is helpful to begin by distinguishing between:

1. the locus of assessment in the educational process
2. the type of assessment carried out
3. the entity being assessed.

With respect to the *locus* of assessment, a distinction is made between a final test or examination, which is at the end of a programme of study, and day-to-day teacher assessment, including progress testing, which takes place on an ongoing basis during the course of study.

Type of assessment distinguishes between a summative assessment, where the aim is to determine if predetermined standards have been reached, and formative assessment, where the primary aim is to support student learning. These two types of assessment are, of course, linked, because formative assessment can lead to interventions or adaptations to teaching and

learning, better to enable a student to meet the standards of the summative assessment.

Finally, the *entity being assessed* can be the student, the teacher, the school, an organisation (such as a local authority or academy chain) or an educational system.

Where the entity being assessed is the student and where the type of assessment is a final, summative test, at the end of a course of study, the information that the test must provide is an accurate and reliable result. It must accurately measure or rank the student, in comparison with a pre-set standard and/or with his or her peers, and it must do so consistently, test after test. This is because, in many cases, the result is used for high-stakes decisions, such as entry to university.

Where the entity being assessed is the student and where the type of assessment is a progress test, which is formative in nature and takes place during a course of study, the requirement is for information which is as detailed and informative as possible. Clearly, the information provided cannot be manifestly inaccurate, but very high levels of accuracy are not as essential as they are in the case of a summative assessment.

Discussions of formative and summative assessment, in the literature, are interesting. Summative assessment and testing have generally been seen to be of major interest at teacher and school level, informing evaluation of teaching and learning. Formative assessment is generally seen as more educationally valuable than summative assessment and testing, because it can focus on validity, using a wide range of assessment methods and ranging across the whole of the subject content. It can support learning, rather than just reporting final outcomes, as summative assessment is generally held to do.

It has generally been thought that formative assessment, using progress testing, cannot be used for summative purposes, largely because it lacks the controls that are put in place around

It is essential that teachers owners of the educational

summative assessments. Final tests or examinations take place in highly controlled conditions, with standardised marking and grading, to ensure that all students are treated equally and fairly. It has also been argued that trying to use formative assessments for summative purposes risks undermining the strengths of those kinds of assessments. If the results become high stakes, it is argued, then the focus will cease to be on identifying strengths *and* weaknesses, and the ability of the assessments to support student progression will be lost.

The adapted approach

Relatively little consideration has been given to the ways in which summative test results can best be used by teachers and schools formatively, to improve teaching and learning and to inform self-appraisal and evaluation.

At present, test results are usually reported to schools in terms of grades or marks. In the adapted approach, summative test outcomes could be reported in more detail and these more detailed reports could serve as the basis upon which teachers could evaluate the approaches and methods they use in the classroom. They could provide opportunities for systematic, evidence-centred evaluation of education. Different approaches to teaching and learning could be evaluated, and the effectiveness and efficiency of these approaches given evidence-based consideration.

can see themselves as the process and as professionals

Requirements for the adapted approach

For final tests to serve a formative function, it is essential that a clear link can be made, by teachers themselves, between the approaches they take in the classroom and the outcomes from the assessments. It is essential that teachers can see themselves as the owners of the educational process and as professionals, whose reasoned interventions can have an impact on student outcomes. To support this, we need to move away from reporting in terms of marks or grades to reporting which:

1. is given in terms of small-grained outcomes:
   - sub-domains (small areas of the curriculum) and pre-specified standards
   - levels of understanding (based on taxonomies)
   - types of errors
2. allows for analysis of student performance taking into account background variables and, for example, specifying how minority groups perform on specific domains
3. allows for the identification of types of error such as common misconceptions
4. can be based on indicators of growth in which it is specified how results on the final test relate to previous performance. This allows for reports showing performance on specific domains (curriculum areas) in relation to the previous performance. So, for example, for the Mathematics GCSE

in a particular school we might see that students perform less well at manipulating algebraic expressions than in other schools, but also that this was the case in that school for a test taken in Year 9 over the previous three years.

With richer reporting, teachers and schools can deepen their own understanding of their results and be better placed to explain them to key stakeholders, such as governors, parents and Ofsted. Richer reporting would support engagement with partner schools, to support raising attainment and, with comparator schools, to inform robust self-evaluation.

Properly understood and used, richer reporting can be used to set educationally meaningful goals for future performance. Teachers and schools can agree their priorities, focusing, say, on particular domains of the curriculum in a coherent way, or on at-risk groups of students. They can meaningfully set tangible, realistic targets for improvement and monitor their progression towards them. With this type of school evaluation, it is schools themselves who are in charge of and in control of their own development. The inspection and accountability frameworks could, over time, evolve to focus on the robustness of schools' own self-evaluation and improvement systems, and there could be meaningful engagement about the goals between schools and Ofsted.

Infrastructure and information-sharing

The adapted approach has IT infrastructure implications and consideration will need to be given about how data is collected, validated and can be used appropriately by stakeholders. The database underpinning the adapted approach would need to include background variables and student results and enable teachers to:

- select sub-groups of students
- select sub-sets of items (questions)
- aggregate results from a student, over a range of subjects, based on overarching taxonomies

- link test results with prior achievement
- store test results of multiple years and allow for comparison of outcomes over years.

This functionality would allow teachers to construct their own outcome indicators and engage meaningfully with the results. They could, for example, select particular sub-groups of students for detailed analysis. They could exclude particular students, or sub-groups, from the analysis, to compensate for year-to-year variations in the cohort and so give themselves an indicator of the underlying stability of their underlying results. They could investigate the performance of particular sub-groups of students over time. Teachers could test hypotheses about the comparability of their own students with the national cohort, or with those in similar schools.

Teachers could select particular test items for analysis, such as those which they judge to be essential to the mastery of a subject. They could analyse performance on particular sub-sets of items with particular sub-sets of students, enabling them to make evidence-based judgements about the particular performance of, say, their highest achieving students, or those with particular needs.

The ability to aggregate student results across a range of subjects could be a powerful driver for whole-school initiatives. Teachers might, for example, focus on literacy across the curriculum, with the identification of relevant items in tests across a range of subjects breaking down barriers between departments and allowing good practice to be disseminated between subjects. The analysis of student performance on those selected items could provide opportunities for discussion between departments about how work in one curriculum area could support that in others and how similarities and differences between subjects can be flagged to students.

The linking of tests to prior achievement allows the construction of growth measures. The approach can be nuanced, with the outcomes from external assessments (such as National Curriculum

tests) linked with those from school-based assessments. Where comparisons need to be made with other schools, school-based assessments could, just for these analyses, be excluded, to ensure the robustness of the outcomes. Teachers might, of course, also be interested in analyses taking account of school-based assessment outcomes, especially if they work in groups of schools who collaborate on assessment.

So, the adapted approach is still based on test data, but uses that data in a far more nuanced way, giving teachers the ability to mine and manipulate that data in ways that can directly inform and support teaching and learning.

Supporting the implementation of the adapted approach would provide government with the opportunity to change the way in which it engages with education professionals and parents. Policymakers would be advocating and supporting the development of rich, summative assessments which would facilitate the use of new reporting tools. This, in turn, would support a move to a genuinely evidence-based approach to teaching and school improvement. In time, government could drive the evolution of the accountability system to one based on robust school self-evaluation and evidence-based approaches for school improvement.

Getting more out of exam results

In this chapter, we propose an adapted approach to school accountability. It is based on providing more granular information from summative assessments and then incorporating that information with other evidence in a database which can be used, on a day-to-day basis, by teachers. That data, with appropriate safeguards, could be shared and used between schools and across the education system, to drive teacher-led and school-based improvement.

In turn, external accountability could evolve to focus on the way individual schools use this information to improve teaching and

learning. Targets could be set intelligently, based on this information. They could be based on comparisons with similar schools and so be challenging, but realistic.

This approach is efficient, because it uses data which already exists within the system both to set and monitor targets and to give schools a steer on where they could look for support. This allows for a system where proper monitoring of the educational approach is possible. It gives a significant role for test results, but far more freedom to schools and clear opportunities for the professional development of teachers.

Most importantly, however, this approach will encourage teachers to own assessment outcomes and have proper responsibility for the outcomes education they work so hard to provide for our young people.

A blueprint for getting more out of test results
- Exam boards should work to develop rich assessments and reporting tools to inform teaching and learning and school improvement.
- As richer assessment becomes used to support robust school self-evaluation, government should consider how the accountability system could evolve to reflect this and focus on supporting school improvement.
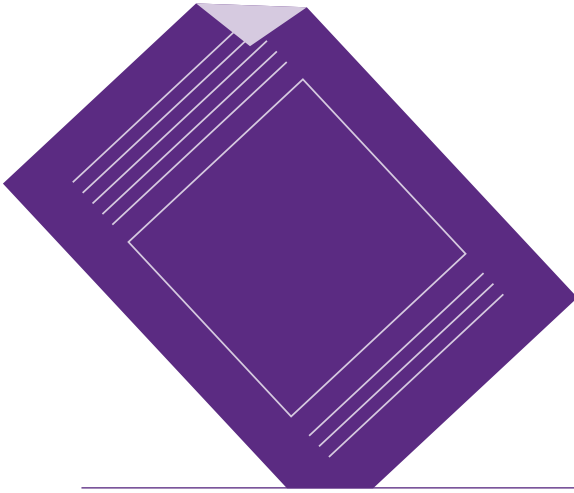
A National Baccalaureate for England

Tom Sherrington

Together with a group of colleagues across the education sector, I am currently involved in the process of building a broad alliance of organisations and individuals to establish a system-led National Baccalaureate for England. We're operating under the auspices of an embryonic organisation called the National Baccalaureate Trust. It is our view that there has never been a better time to push this idea forward and there is a very real opportunity for us to make this happen over the next few years.

Our current framework for recognising the achievements of young people has numerous inherent flaws many of which would be resolved by the introduction of a Baccalaureate framework, built around existing qualifications. This is our chance to change things so that, by 2025, the way we view success and achievement will have changed significantly.

At the heart of the concept is a belief in the value of a broad education; the kind that would allow all young people to develop as rounded individuals with the knowledge, skills and personal attributes required to have productive, fulfilling lives as citizens of the world. Across the political spectrum there is a growing understanding that 'character' and 'values' matter; organisations such as the CBI repeatedly call for a broader view of what young people should study at school.

And yet our current system is so narrowly focused on examination outcomes that major areas of learning and achievement are systematically and disastrously undervalued, and the examination system is placed under unsustainable strain. Ever increasing expectations of rigour, validity and reliability are heaped upon it by the culture of hyper-accountability, leading schools and colleges into perverse behaviours that compromise the integrity of long-standing subject disciplines. Worst of all, success at school is effectively reserved for just over 50% of the population; if many more students gained success at GCSE, the exams would be deemed 'too easy' and standards would be regarded as having fallen, not risen. That's a

bad place to be for any system.

The features of a Baccalaureate are really quite simple:

- *Core learning:* A-levels, Tech Levels, Pre-Us or any other major examinations – at the appropriate level
- *Personal project:* an accredited project or essay such as the Extended Project Qualification or the Pre-U Global Perspectives and Research course
- *Personal development programme:* a programme of learning experiences which might include requirements to feature physical, creative, cultural and community service elements, amounting to 100–150 hours of activity over two years.

Our vision is for a suite of Baccalaureate models to emerge, providing for students at every level from Entry Level, to Foundation (Level 1), Intermediate (Level 2) and Advanced (Level 3). Students would have until the age of 25 to complete their Baccalaureate at the most advanced level they can reach; there would be minimum requirements but no upper limit to the range of qualifications that they could include in their portfolio.

The final element is a transcript that would capture all of this in one common format so that universities and employers could build a picture of young people that is more complete than the standard array of examination grades provide by themselves.

Clearly, there are no panaceas but I firmly believe that the introduction of a National Baccalaureate framework such as this would make a significant contribution to tackling some of these issues.

Bridging the academic–technical divide

All learners, whether in an elite academic sixth form or undertaking a vocational apprenticeship with part-time study, would be included in the same system. Long-standing 'gold standards' such as taking three or four A-levels or even the IB Diploma, as well as new

developments such as the proposed Tech Bacc, would fit inside the National Bacc framework.

The language of success – of 'completing my National Bacc' – would be universal. The transcript would allow for detailed evaluation of strengths of different candidates, depending on what is required, but there would be no sheep and goats; no artificial equivalences through points measures and no barriers to students mixing academic and technical learning. This is a profound shift – a Baccalaureate that is truly inclusive instead of persisting with a system that is dismissive of so many people's achievements.

Creating pathways to success for all learners

The tiered end-point certification (from Entry Level to Advanced Level) is significant. The phrase 'all must win prizes' is so often used as a pejorative to denigrate examination standards. However, the system we envisage could literally mean that, while maintaining very high standards. It would be possible for, say, a young person with learning difficulties to strive for their National Bacc at Entry Level, finally succeeding aged 23 to gain what would be a meaningful and significant achievement: a record of success. Within the same framework, someone else would fly through to gain an Advanced Bacc with five A* grade A-levels. They would both have completed a personal project; they would both have logged a range of personal achievements in their personal development programme; they would both be rounded individuals with successes for all to see on their National Baccalaureate transcript. So, yes, all can win prizes. If the prizes have value, why would we want it any other way?

Reducing the emphasis on examinations at 16

With increasing participation in post-16 education and higher education, it is anomalous that we continued to give so much weight to GCSEs at 16 in our accountability measures. The introduction of a National Baccalaureate would help to shift the emphasis to terminal

outcomes at 18 as the default point at which to measure success. In time this could lead to diffusion of the institutional intensity that prepares students for GCSEs, creating a genuine 14–18 curriculum. 11–16 schools would have additional incentives to form partnerships with post-16 providers to ensure continuity and progression, and all learners would complete Key Stage 4 with the bigger picture of their Baccalaureate firmly in mind.

Giving value to personal development within the curriculum

It is no use talking about building character or developing resilient, articulate young people if we do not build in the opportunities for these attributes to be developed. A personal development programme (PDP) may, in many cases, be the sum of the range of activities that many young people already engage in. But, without question, this is far from being universal across the system. Some students will be undertaking an 18-hour per week single-subject BTEC and do absolutely nothing else. We know that exams are not the be-all and end-all of learning and achievement so we need a way to give value to all of the rest; a Bacc that requires students to complete a PDP does just that. Free from spurious points and equivalences, the PDP simply has to be completed and the content logged for all to see.

Locating issues of examination reliability and validity in a broader context

Much-needed recent moves to tighten the validity and reliability of our national examinations have led to concerns about a narrowing of the focus of our curriculum towards that which can be measured through examination. A Baccalaureate with a strong PDP element allows for a wide range of learning experiences to be taken out of the domain of rigid examinations while still giving them value. For example, a PDP might include the requirement to engage in a real public speaking activity; arguably this is more authentic, more

appropriate and more sustainable than creating the machinery required to moderate national exams in speaking and listening for English GCSE. The Bacc gives space for exams to do their job, releasing them from the requirement to represent everything that matters.

Giving value to extended personal study

It is significant that most Bacc-style qualifications such as the Pre-U and the IB Diploma all include an extended study component. The pursuit of a personal project of some substance is very powerful in terms of knowledge acquisition and the development of a range of skills and dispositions, as well as providing learners with space to explore an area of genuine personal interest. It is too time-consuming and cumbersome to include this type of experience within every type of qualification – death by coursework is all too familiar! However, the requirement to complete a Personal Project as part of the National Bacc would ensure that all learners have the opportunity, regardless of the content of their core learning. It could work at every level and in every learning environment, providing a powerful unifying learning experience across the diverse and fragmented landscape of post-16 education.

Providing incentives to develop post–GCSE Maths and English qualifications

One problem with the current range of qualifications is that post-16 study of Maths and English is limited to re-sits of GCSEs or full Level 3 qualifications. If we want more learners to continue to study Maths and English to 18, we need new ways of accrediting their achievement. The Bacc development would allow space for students to take new slim-volume qualifications, possibly assessed online on a 'when ready' basis, as part of their broader package of qualifications and activities. If these elements became requirements for Bacc completion, the market for new qualifications of this kind would

Lots of students get hatfuls of A*s. So what? Does that tell employers all they need to know?

stimulate much needed development in this area. Currently it is very difficult to for schools and colleges to mandate post-16 study of Maths and English because the range of accreditation routes needed simply doesn't exist.

Detailing achievements to universities and employers

It is already part of school folklore that universities find it difficult to select candidates. We hear that everyone looks the same, that personal statements are too manufactured. We also hear from employers that they find it difficult to know what a candidate is *really* like based on their qualifications. Lots of students get hatfuls of A* grades. So what? Does that tell them all they need to know? Well, obviously not.

A central feature of the National Baccalaureate concept is that all students would carry their transcript with them. Using very simple QR code technology,[1] a full record of their achievements could be made available to anyone. This could include exam grades, component scores, their personal project and the details of their personal development programme. Every young person's transcript would follow a standard format to facilitate the process of comparing candidates at the required level of depth for different purposes.

Furthermore, schools and colleges would be able to use the transcripts formatively: "What does your transcript say about you?" That has the potential to be a very powerful lever to motivate students – not only to aspire for academic and technical excellence, but also to develop ever more impressive portfolios of personal experiences; to become truly rounded people with something special to offer.

It will be a long haul. We're deliberately starting out slowly; purposefully seeking to build a consensus around the National Baccalaureate from the grassroots, free from the political imperatives inherent in a government-led system. The prize of a fully inclusive, challenging and holistic framework that captures the outcomes of

each young person's education in a fully rounded sense is achievable – and has to be worth pursuing.

A blueprint for a National Baccalaureate

- The National Baccalaureate Trust should support and build a system-wide alliance to establish a National Baccalaureate for England.
- As the system-led National Baccalaureate gains traction, government should consider its implications for the curriculum, qualifications, accountability and funding frameworks, and how schools and colleges could be supported to deliver rich, broad Baccalaureate provision for every young person.

Reference

1. A QR code is a type of barcode that can be read by any smartphone. A unique QR code can be easily generated for each student, allowing anyone to scan their National Bacc transcript and link through to supporting information.

References

Acquah, D. and Huddleston, P. (2014), "Challenges and opportunities for vocational education and training in the light of Raising the Participation Age". *Journal of Research in Post-Compulsory Education*, 19(1):1–17.

Barnacle, R. (2009), "Gut instinct: The body and learning". *Educational Philosophy and Theory,* 41(1): 22–23.

Billett, S. (2009), "Constituting the workplace curriculum". *Journal of Curriculum Studies*, 38(1):31–48.

Blatchford, P., Kutnick, P., Baines, E. & Galton, M. (2003), "Toward a social pedagogy of classroom group work". *International Journal of Educational Research*, 39:153–172.

Care, E., Griffin, P., Scoular, C., Awwal, N. & Zoanetti, N. (2015), "Collaborative problem solving tasks" in Griffin, P. & Care, E. (eds.), *Assessment and Teaching of 21st Century Skills: Methods and Approach*. London: Springer.

Carr, D. (1999), "Further Reflections on Practical Knowledge and Dance a Decade on", in McFee, G. (ed.) *Dance, Education and Philosophy*. Oxford: Meyer and Meyer.

Cooper, S. and Ord, K. (2014), "Developing 'know how': a participatory approach to assessment of placement learning". *Journal of Vocational Education and Training*, 66(4):518–536.

Department for Business, Innovation and Skills (2013), *International Education: Global Growth and Prosperity*. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/340600/bis-13-1081-

international-education-global-growth-and-prosperity-revised.pdf [28 May 2015].

Department for International Development (2013), *Summary of DFID's Work in Pakistan 2011–2015*. Available: https://www.gov.uk/government/uploads/ system/uploads/attachment_data/file/279819/Pakistan-summary2.pdf [6 May 2015].

Eisner, E. (1993), "Reshaping assessment in education: Some criteria in search of practice". *Journal of Curriculum Studies*, 25 (3); 219–233.

Greaney, V. and Hasan, P. (1998), "Public examinations in Pakistan: a system in need of reform" in Hoodbhoy, P. (ed.), *Education and the State: 50 Years of Pakistan*. Karachi: Oxford University Press.

Griffin, P. & Care, E. (eds.) (2015), *Assessment and Teaching of 21st Century Skills: Methods and Approach*. London: Springer.

Griffin, P., McGraw, B. & Care, E. (eds.) (2012), *Assessment and Teaching of 21st Century Skills*. London: Springer.

Gulikers, J. (2006), *Authenticity is in the eye of the Beholder: Beliefs and Perceptions of Authentic Assessment and the Influence on Student Learning*. Maastricht: Interuniversity Center for Educational Research, Open-Universiteit-Nederland.

Gulikers, J. Biemans, H. and Mulder, M. (2009), "Developer, teacher student and employer evaluations of competence-based assessment quality". *Studies in Educational Evaluation*, 35 (2–3): 110–119.

Gulikers, J., Bastiaens, Th. J., & Kirschner, P. (2004), "A five-dimensional framework for authentic assessment". *Educational Technology Research & Development*, 52 (3): 67–85.

House of Commons Education Committee (2013), *2012 GCSE English Results*. Available: http://www.parliament.uk/business/committees/committees-a-z/commons-select/education-committee/news/2012-gcse-english-results-substantive/ [6 May 2015].

House of Commons Education Committee (2014), *Underachievement in Education by White Working Class Children*. Available: http://www.publications.parliament.uk/pa/cm201415/cmselect/cmeduc/647/64704.htm [1 June 2015].

Mercer, N. (2015), "Why oracy must be in the curriculum and group work in the classroom". *Forum for promoting 3–19 comprehensive education*, 57(1).

Mercer, N., Wegerif, R. & Dawes, L. (1999), "Children's talk and the development of reasoning in the classroom". *British Educational Research Journal*, 25(1):95–111.

OECD PISA (2013), *Draft Collaborative Problem Solving Framework*. Available: http://www.oecd.org/pisa/pisaproducts/DraftPISA2015CollaborativeProblemSolvingFramework.pdf [1 June 2015].

Peal, R. (2014), "Lies my teacher training taught me #3: most great learning happens in groups". Available: https://goodbyemisterhunter.wordpress.com/2012/10/24/lies-my-teacher-training-taught-me-3-most-great-learning-happens-in-groups/ [1 June 2015].

Pring, R. (2007), "14–19 and lifelong learning: Distinguishing between academic and vocational learning" in Clarke, L. and Winch, C. (eds.), *Vocational education: International approaches developments and systems*. London: Routledge.

Pring, R., Hayward, G., Hodgson, A., Johnson, J., Keep, E., Oancea, A., Spours, K. and Wilde, S. (2009), *Education for all. The future of education and training for 14–19 year olds*. London: Routledge.

Stacey, G. (2014). "Putting validity at the heart of what we do". Speech to the Federation of Awarding Bodies, Leicester, 14 October. Available: https://www.gov.uk/government/speeches/putting-validity-at-the-heart-of-what-we-do [6 May 2015].

Von Davier, A. & Halpin, P. (2013), *Collaborative problem solving and the assessment of cognitive skills: psychometric considerations*. ETS Research Report RR-13-41. Available: https://www.ets.org/Media/Research/pdf/RR-13-41.pdf [1 June 2015].

Wiggins. G. (1989), "Teaching to the (authentic) test". *Educational Leadership*, 46(7):41–47.

Wilson, M., Neja, I., Scalise, K., Templin, J., Wiliam, D. & Torres Irribarra, D. (2012), "Perspectives on Methodological Issues" in Griffin, P., McGaw, B. & Care, E. (eds.), *Assessment and Teaching of 21st Century Skills*. London: Springer.

Wright, J., Brinkley, I. & Clayton, N. (2010), *Employability and skills in the UK: redefining the debate*. London: The Work Foundation.

Contact
Kimberley Trewhitt
Head of Policy
ktrewhitt@aqa.org.uk

What should assessment look like in 2025?

The future of assessment: 2025 and beyond is a conversation to consider the big questions and issues that will shape how assessment evolves over the next decade and beyond. The project is a collaboration of UK and international experts including teachers, academics, employers and policymakers.

We've gathered a working group of selected senior education experts who will help develop policies that the next government could implement, moving us towards our longer-term objectives on balancing assessment and accountability, assessment for the real world, and 21st century assessment.

AQA

aqa.org.uk/assessment2025
#assessment2025