

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/360806817>

Rethinking assessment in schools.

Article · May 2022

CITATION

1

READS

683

1 author:



[Bill Lucas](#)

The University of Winchester

126 PUBLICATIONS 1,461 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



The Habits of an Improver - a model for improving healthcare [View project](#)



Learning to be an Engineer [View project](#)

Rethinking assessment in schools: Moving from a deficit to a strengths-based model

Bill Lucas

ACROSS the world assessment in schools is not working. We are not yet evidencing the kinds of dispositions and capabilities society increasingly wants. Assessment is increasingly out of sync with curriculum and pedagogy. Where teachers are becoming evidence-based in their practices, they, and all those who administer tests and set examinations, are failing to keep up with the science of assessment. Educational jurisdictions place too much reliance on high-stakes, standardised testing skewing the focus of assessment towards summative events rather than formative processes. We need ways of evidencing the full range of young people's skills by devising curricula, pedagogies, and assessment systems to cultivate and evidence their strengths.

In the last few years initiatives such as Rethinking Assessment (<https://rethinkin-gassessment.com>) in England, New Metrics for Success (<https://education.unimelb.edu.au/new-metrics-for-success>) in Australia, and the Optimizing Assessment for All initiative of the Brookings Institution (<https://www.brookings.edu>) in the USA are indicative of a global backlash against an often reductionist and deficit-based model of education.

The beginnings of a global curriculum

Across the world there is a growing consensus as to the kinds of knowledge, skills, and dispositions students need to thrive today. There are a dozen or so models of what contemporary curricula should look like of which The World Economic Forum (2015), is

widely cited. The model is titled 'Twenty-first century skills'.

It lists seven 'foundational literacies' (how students apply core skills to everyday life) – Literacy, Numeracy, Scientific literacy, ICT literacy, Financial literacy and Cultural and civic literacy, four 'competencies' (how students approach complex challenges) – Critical thinking/problem-solving, Creativity, Communication and Collaboration, and six 'character qualities' (how students approach the changing environment) – Curiosity, Initiative, Persistence/grit, Adaptability, Leadership and Social and cultural awareness.

The process of curriculum evolution can easily become politicised. Some lament a perceived move away from a focus on subject disciplines while others celebrate the increased prominence of concepts like creative thinking, collaboration, and communication. This tension can be presented as a traditional versus progressive tug of war. I have argued (Lucas, 2019) that this is unhelpful. The World Economic Forum, for example, frames the curriculum as being made up of 'twenty-first century skills'. But to many ears this phrase sounds vague, faddish, almost evangelical. From a researcher's point of view, it is patently silly; if the premise of these skills is that they are changing to meet the needs of a fast evolving world it is a nonsense to pretend that we know which skills will be needed in eighty years' time at the end of the twenty-first century let alone in the shorter time it takes for curricula to evolve.

A misunderstanding of the role of skills in learning

There is much nonsense talked about skills today. By those who see the acquisition of knowledge as the main purpose of education, an emphasis on skills is often portrayed as an attempt to dumb down or distract schools from their core purpose. By those who see dispositions and capabilities as being centrally important there is a temptation to hold fast to bigger concepts such as creativity or collaboration without recognising that they are in reality made up of aspects of knowledge and clusters of skills.

Knowledge and dispositions are not polar opposites, just different ways of categorising what we can learn. The ‘currency’ of both is skills. Skills are what matter in life; skills are the ‘connective tissue’ between knowledge and dispositions. As we practise a skill in different contexts we become more competent, confident, and capable until it becomes a disposition, something we are disposed to do.

Planning an essay. Delivering a speech. Critiquing an argument. Having a good idea when you need one. Tying your shoelace. Trying different approaches when faced with a tricky problem. Using your common sense when your satnav takes you to a cul-de-sac not yet updated in its software. Reading the mood of those with whom you are working. Facilitating a workshop where you are a content expert. Facilitating a workshop when you have only a basic knowledge of the context but can transfer facilitation skills learned in other contexts to the task at hand. And, yes, recalling decontextualised information in a pencil and paper examination is also a skill, but not one that adults need to use much in a digital age.

The deeper your knowledge and the more you practise your skills in a variety of contexts, the more capable you become. Dispositions are clusters of skills which have been practised so well that they have become habitual; you are routinely disposed to deploy them. And skills are the mechanism by which knowledge is applied and dispositions are lived out.

From a Rethinking Assessment perspective, it is clear that, while global curricula are beginning to be rebalanced, such rebalancing varies greatly from country to country. In England, for example, the National Curriculum (Department for Education, 2014) remains doggedly knowledge-focused:

The national curriculum provides pupils with an introduction to the essential knowledge that they need to be educated citizens. It introduces pupils to the best that has been thought and said; and helps engender an appreciation of human creativity and achievement.

(Department for Education, 2014, p.6.)

The mechanisms by which students move from being told about great thinkers and speakers to developing an appreciation of creativity is spectacularly absent in the document. Contrast this with the ambitious ‘four purposes’ of the new National Curriculum in Wales (Welsh Parliament, 2020):

Ultimately, the aim of a school’s curriculum is to support its learners to become:

1. ambitious, capable learners, ready to learn throughout their lives
2. enterprising, creative contributors, ready to play a full part in life and work
3. ethical, informed citizens of Wales and the world
4. healthy, confident individuals, ready to lead fulfilling lives as valued members of society (Welsh Parliament, 2021, p.2).

While thinking about a global curriculum may be gathering pace, the outcomes of such deliberations are not equally distributed among the four home nations of the UK. As educational jurisdictions think about what our children need to learn it will be helpful to explore such questions as:

- What kind of knowledge is it important for all young people to have?
- What kind of dispositions are important for all young people to acquire?

- How can we ensure that young people acquire and apply useful knowledge in a range of settings?
- How can we teach young people to work across subject disciplines as happens in the real world, ensuring that they have the necessary building blocks in place to be able to do this?
- How can we ensure that important dispositions for learning and for life are best cultivated in a range of disciplinary contexts?
- How can we develop strength, breadth, and depth in learning to facilitate its transfer across contexts?
- Which pedagogies work best for promoting deep learning?

The answers we give will help us determine what a broad, balanced, and contemporary curriculum might look like. But, given that what gets measured tends to get taught, it leaves much work to be done in considering how we can align assessment methods to such desired curriculum outcomes.

Assessment systems narrowing and undermining learning and learners

Currently the knowledge that is typically assessed in schools is from a narrow range of subjects, rarely explored in depth and almost never interdisciplinary. Practical knowledge and skill is not much assessed in general education and individuals rather than teams remain the focus. Complex, higher order skills are rarely assessed in ways that recognise the subtleties involved (Darling-Hammond, 2017). Many dispositions or capabilities known to be important in life are not assessed at all.

Assessments frequently require recall of content but rarely demand the kind of deep thinking, problem-solving, or application needed in the real world. Traditional areas, literacy, maths, and science continue to require considerable content to be tested, while newer areas such as citizenship, engineering, sustainable development, and ethical understanding are only briefly

explored. Except in a very few countries there is little or no interdisciplinary assessment. Students' capabilities in planning and undertaking extended investigations are rarely assessed. Although the ability to collaborate with others is widely valued in the workplace it is only acknowledged at school on the sports field or in music and drama performances.

While dispositions or capabilities are becoming more visible in curricula they are rarely assessed; at a global level PISA's innovative domain tests of collaborative problem-solving (OECD, 2017) and creative thinking (OECD, 2019) are exceptions.

Assessment methods are too blunt. Most tests used in schools still rely on paper and pencil. They examine aspects of knowledge and routine skills. They test students' ability to remember and write about something rather than apply or do the thing they have been learning. Concepts and skills are tested in individual subjects and only very rarely across disciplines:

Traditional assessment methods typically fail to measure the high-level skills, knowledge, attributes and characteristics of self-directed and collaborative learning that are increasingly important for our global economy and fast-changing world.

(Griffin et al, 2012, pp.v-vi.)

Students are tested at set times rather than when they are ready often to meet the needs of the next educational provider or, frequently ineffectively, of employers. These inflexible encounters with assessment ignore the huge variety of student achievement levels, 'in any given year of school, the most advanced learners in areas such as Reading and Mathematics can be as much as five or six years ahead of the least advanced learners' (Masters, 2013, p.3), the fact that 'attainment is only loosely related to age' (Wiliam, 2007, p.248), and the differing levels of maturity found in any cohort on account of birth dates.

In the last two years when Covid-19 has made the experience of life and school particularly challenging for young people, the negative impact of some assessments is especially worrying. Students can all too easily conclude that they are failures (Education Policy Institute, 2019). They can be demotivated to the extent that they may not stay on at school or find employment (Milligan et al., 2020). High stakes testing has a negative impact on young people's wellbeing (Howard, 2020) and can reduce performance through anxiety, especially for students of lower ability (von der Embse et al., 2018).

More fundamentally, most assessments fail to capture the degree to which students have progressed over time. Instead they:

...provide snapshots of achievement at particular points in time, but they do not capture the progression of students' conceptual understanding over time, which is at the heart of learning.

(Pellegrino et al., 2001, p.27.)

Assessments need not be done in this way:

'Measuring progress provides a deliberate counterpoint to the traditional practice of measuring achievement at specific time points'

(Hipkins & Cameron, 2018, p.22).

Rethinking assessment as a driver of student learning

More than a decade ago, in a magisterial overview of the factors affecting student achievement, John Hattie reflected:

I realised that the most powerful single influence enhancing achievement is feedback. This led me on a long journey to better understanding this notion of feedback.

(Hattie, 2009, p.12.)

Hattie describes the kinds of feedback which matter, student to student, teachers

to student, student to teacher and teacher to teacher. In these apparently simple interactions Hattie is essentially articulating a true assessment system at work, one with myriad kinds of feedback in play, one that is focused on improving the learner's progress.

Sadly, over the last few decades we have progressively lost our way with educational assessment. What we assess grows ever further away from what we want young people to be able to know, do, be, and become in the complex world in which they live today. The focus of most systems is on summarising rather than understanding, recalling rather than applying, noticing deficiencies rather than celebrating strengths. We seem happier using numbers rather than narratives, keener on judging rather than prompting improvement.

Twenty years ago, the Committee on the Foundations of Assessment in the USA (National Research Council, 2001) considered the degree to which advances in the cognitive sciences were impacting on educational assessment. It concluded with a vision of assessment which has still not been realised in education:

In the future envisioned by the committee, educational assessments will be viewed as a facilitator of high levels of student achievement. They will help students learn and succeed in school by making as clear as possible to them, their teachers, and other education stakeholders the nature of their accomplishments and the progress of their learning.

(National Research Council, 2001, p.292.)

The Gordon Commission in the USA in 2013 made a number of key recommendations about designing and implementing assessment that supports a more ambitious and expansive vision of education. It is vitally important, the Commission argued, that assessments 'best represent the kind of learning students will need to thrive in the world that awaits them beyond graduation' (Gordon Commission, 2013, p.8).

In Australia, at the same time as the Gordon Commission, the Australian Council for Educational Research (ACER) undertook a review of educational assessment, (Masters, 2013). Geoff Masters reminds us that ‘the fundamental purpose of assessment is to establish where learners are in their learning at the time of assessment’ (pp.5-6).

Using assessment to evidence students’ strengths

Despite the continued existence of high-stakes testing across the world there are many signs of progress as educators seek to evidence their students’ all-round knowledge, skills, and dispositions. The following is a selection of promising examples loosely organised into eight categories:

1. Psychometric tests

In the main these are self-reported online tests or apps often used to evidence an aspect of character, wellbeing, or metacognition:

- Angela Duckworth’s grit scale (<https://angeladuckworth.com/grit-scale/>)
- The VIA Survey of Character Strength (<https://www.viacharacter.org/>)
- The Harvard Human Flourishing app (<https://hfh.fas.harvard.edu/flourishing.app>)
- Carol Dweck’s Growth mindset assessment (<https://blog.mindsetworks.com/what-s-my-mindset>)
- The Metacognition Awareness Inventory (<https://services.viu.ca/sites/default/files/metacognitive-awareness-inventory.pdf>)

These tests as yet have varying degrees of reliability and validity. KIPP schools (<https://www.kipp.org/approach/character/>) in the USA, for example, have adopted approaches to developing and assessing character using resources such as the playbooks provided by CharacterLab (<https://characterlab.org/playbooks/>).

2. Smart multiple choice

While some multiple choice tests can be reductionist and focus on recall or simple computa-

tions, others can be a way of evidencing aspects of dispositions such as critical thinking:

- Raven’s Progressive Matrices Test (<https://www.pearsonclinical.co.uk/Psychology/AdultCognitionNeuropsychologyandLanguage/AdultGeneralAbilities/Ravens-Progressive-Matrices/Ravens-Progressive-Matrices.aspx>)
- California Critical Thinking Skills Test (<https://www.insightassessment.com/article/california-critical-thinking-skills-test-family>).

The Mission Skills Assessment (https://bbk12e1-cdn.myschoolcdn.com/ftpimages/721/misc/misc_178206.pdf) is an interesting example of a multiple choice test looking to evidence a wider set of constructs - creativity, curiosity, ethics, resilience, teamwork, and time management.

3. Performance based assessment

Performance based assessment is a broad field encompassing traditional approaches the Viva to AI simulations. Other examples include tests of proficiency in the arts and sports. Increasingly it is being used to enable students to perform tasks or activities that are meaningful and engaging. Performance also includes exhibitions, presentations, and debates, sometimes associated with pedagogies such as project- and problem-based learning.

Alelo’s Oral Language Simulation (<https://www.alelo.com/>) is a recent example of a computer-based simulation designed to measure not only proficiency in a foreign language. The program allows a student to interact directly with an avatar in a variety of languages.

Both the PISA test of collaborative problem-solving and of creative thinking are performance based assessments and the Victorian Curriculum and Assessment Authority’s online tests of Critical and Creative Thinking are examples of scenario-based tests of performance (<https://www.vcaa.vic.edu.au/assessment/f-10assessment/edstateap/Pages/cct-assessments.aspx>).

4. *Extended investigations*

Extended investigations are increasingly seen by researchers (Conley & Darling-Hammond, 2013; Hipkins & Cameron, 2017; Soland et al., 2013) as the kinds of assessments needed to best evidence higher-order thinking skills present in many dispositions. Many extended investigations contain elements of performance-based assessment.

In addition to a growing number of schools using such approaches across the world, there are externally validated examples to consider:

- The Extended Project Qualification (EPQ, <https://qipsucas.com/qip/extended-project-qualification-epq#:~:text=The%20EPQ%20is%20a%20single,%2C%20report%2C%20dissertation%20or%20artefact>) in England and Wales. The EPQ enables students to undertake an investigation in the context of a project topic they have selected. Outcomes can be a design, performance, report, dissertation, or artefact. An EPQ 'counts' for half the points value of an A level for university entrance.
- The New South Wales Personal Interest Projects (<https://educationstandards.nsw.edu.au/wps/portal/nesa/11-12/stage-6-learning-areas/hsie/society-culture/personal-interest-project>), as their name suggests offer the potential for extended investigation but, like many such investigations are assessable only in a written essay format.

Alongside the four methods illustrated in the last pages there are two other dimensions worth noting to do with the availability and timeliness of assessments, micro-credentialing and on demand tests.

5. *Micro-credentialing*

Micro-credentials, sometimes referred to as digital badges, take an idea long used by scouts, guides and other informal youth organisations and bring it into the digital age. Aspects of a larger concept such as

a disposition are reduced to a small number of skills and 'badged' up to enable students to acquire credential in bite sizes. Badges have the advantage, too, of providing a visual image of a student's progress at a glance. Badges are, in the jargon of assessment, 'stackable'. That is to say they can be combined together to evidence many different assessment outcomes. In Europe there is an attempt to build micro-credentials into a reliable system of credit transfer between schools and universities, (Futures et al., 2020).

6. *On demand and online*

In our learning lives outside school the idea of only being able to take a driving test, for example, on a set date would be laughed at. In most countries such tests require some theory (typically an online multiple choice test) and a practical demonstration on the road. The on the road test is, unsurprisingly attempted when we and our driving instructor think that we are capable. In similar vein only being able to use our memory rather than our ability to search and apply knowledge from the Internet or from notes or materials we have made while studying is increasingly perverse in a digital age.

On demand testing is already widely used in music, for example. While still at school many students take grade exams to assess the quality of their playing classical, jazz or rock music when they have reached a standard that they and their teachers deem them to be ready. On demand testing more generally in education makes practical sense and a number of educational jurisdictions are experimenting with it, for example the VCAA in Australia (<https://www.vcaa.vic.edu.au/assessment/f10assessment/ondemandtesting/Pages/benefits.aspx>) and Aberdeen University in Scotland (<https://on.abdn.ac.uk/courses/access-maths-5/>).

Sugata Mitra suggests that, in the future, assessment should precede teaching and learning should actively encourage critical

thinking and consensus-building using the Internet. He imagines a post-Pandemic classroom:

Sessions usually start with a set of questions. In the pre-pandemic times, this would have been called a test. Tests were usually given after the 'teaching and learning' were over. Not so anymore. Sessions can start with tests. The children have no idea what the answers might be, they haven't been 'taught'. But they can look up things on the Internet and talk to each other. When the answers come in, the teacher begins a discussion. She encourages the children to talk about their answers, sometimes, very occasionally, she adds a bit. They arrive at a consensus by the end of the session about what the answers are and why.

(Mitra, 2020, p. 287.)

Mitra's open style of teaching has an equivalent in assessment, open book tests. Here students are able to have key texts with them as they answer a question. Open book tests do not just test a student's ability to recall information but require a more critical and analytical and applied approach to answering questions.

In a recent blog, *The searching questions that will allow us to rethink assessment* (<https://rethinkingassessment.com/rethinking-blogs/the-searching-questions-that-will-allow-us-to-rethink-assessment/>), Guy Claxton suggests that we need not spend time worrying about end of school exams but instead could leave summative assessment – he calls this process MOECs (Methods of Evidencing Capability) – to a college or university or to an employer rather than undertaking it at the exit point from school.

7. Games-based assessment

Games-based assessment is in its infancy in schools. Nevertheless, there are examples from which we can learn.

Keenville (<https://www.fablevision-studios.com/keenville>) is a formative game-based assessment initiative for 1st and 2nd graders in the state of Georgia,

a collaboration between the Georgia education department, the Georgia Center for Assessment at the University of Georgia, and FableVision Studios.

Posterlet is an educational game-based assessment that measures two design thinking choices: students' choices to seek critical (i.e. negative) feedback and to revise their work while they learn graphic design principles through creating posters. Posterlet is a game created by Stanford's Graduate School of Education to measure students' choices in seeking feedback and revising work while at the same time learning graphic design principles.

Critical and Creative Tests developed by VCAA in partnership initially with ACER and now with NFER in England use a range of engaging scenarios to test students critical and creative thinking skills.

8. Comparative judgment

Comparative judgement uses the principle that people are better at making comparisons *between* pieces of work than at making *absolute* judgements about their quality. First identified as a useful principle of assessment nearly a century ago (Thurstone, 1927), the addition of digital technology enables a kind of crowd sourcing of teachers' judgments of particular a usefulness when reviewing written work. A number of approaches are being developed by No More Marking in the UK (<https://www.nomoremarking.com/>).

An urgent need for a Visible Assessment Toolkit

Over the last two decades there has been a step change in the way in which evidence about learning is used by schools and within school systems; the science of learning is a widely accepted concept in education. This development owes much to John Hattie's *Visible Learning* (Hattie, 2009) and to the *Education Endowment Foundation Teaching and Learning Toolkit* (<https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit>) which has made research clear and practically useful to teachers.

By analogy with *Visible Learning* we now need to enhance the assessment literacy of all teachers, drawing on the learning sciences, to create *Visible Assessment*, a practical toolkit of multi-modal assessment methods carefully curated to show their purpose, benefits, validity, and reliability.

Topics within such a toolkit might explore:

Purpose and consequence

- The importance of understanding the purpose any assessment is intended to serve
- The tensions which exist between summative and formative approaches
- The many unhelpful consequences of high-stakes assessment

Depth and Breadth

- Ways of evidencing high-order thinking skills reliably
- The need for better definitions of dispositions and their associated learning progressions
- The desirability of assessments being pedagogically sensitive and educationally valuable
- The complexity of designing ways of fairly evidencing student progress within dispositions
- A growing interest in the concept of mastery
- The need for flexibility to ensure that the full range of abilities can be fairly assessed
- A focus on collaborative rather than just individual performance

Authenticity

- Increasing interest in strengths-based approaches, especially from employers
- The need to design better performance-based assessments
- A move towards scenario-based, authentic assessment
- A move towards assessments of investigations over longer time periods
- Approaches to assessment on demand
- Increased opportunities for student involvement and agency in the process

Progression and Improvement

- The benefits of assessment for and as learning
- The need for multi-modal approaches to assessment, incorporating data from a number of sources

Quality infrastructure

- A better understanding of when to use assessment of, for and as learning
- Enhanced teacher capacity in assessment literacy and moderation
- The desirability of international bench-marking.

In a small way, drawing on earlier research in England and Australia (Lucas, 2016), I have started the process of making assessment processes more visible and building teacher assessment literacy by developing a *Field guide for assessing creative thinking* (Lucas, 2022).

A global movement towards strengths-based assessment

Just as earlier it was clear that within the UK there are very different approaches to national curricula, so, across the world, there is huge variation in terms of assessment. Nevertheless, it is possible to detect some clear trends, see Figure 1.

Each of these ten trends has potentially positive implications for the experiences of young people at school. But each remains a matter ultimately for educational jurisdictions to respond to. Were they all to be adopted, alongside changes to curriculum and pedagogy, they offer a chance for educators to cultivate the full range of young people's strengths and better prepare them for a lifetime of learning.

Shifting from certificates to profiles as a record of a learner's all-round achievements

From conversations with researchers and policy-makers across the world it is becoming clear that the single most important next step we need to take is the development of a credible, web-based learner profile for all school leavers to showcase the full range of

Figure 1: Global trends in assessment (Lucas, 2021).



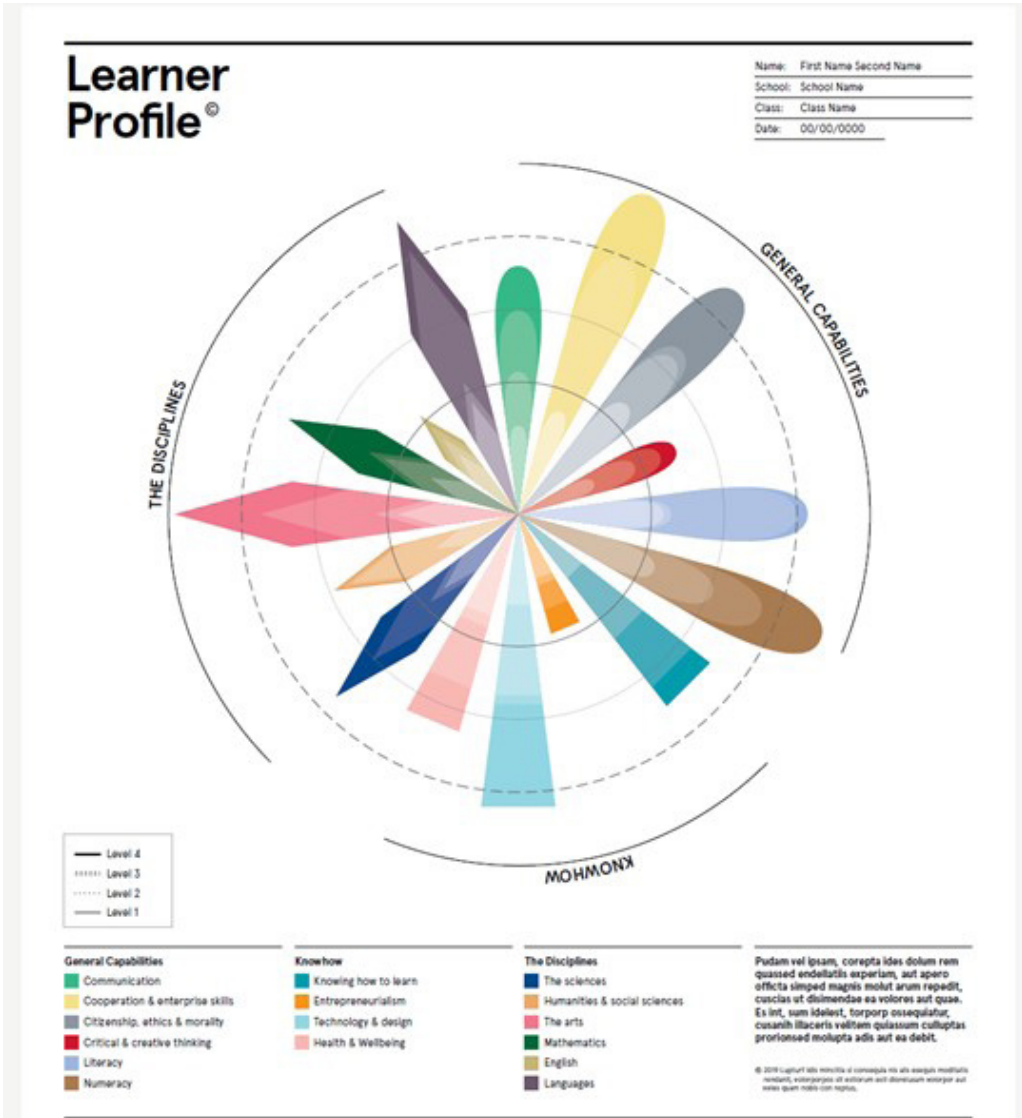
their formally and informally acquired skills, rather than this transition being marked by a set of scores or grades. There is a global move towards using learner profiles, transcripts, records of achievement and portfolios alongside whatever standardised data is felt to be important, led by our colleagues at New Metrics for Success in Australia and by Rethinking Assessment in England. Figure 2 is illustrative of the kind of prototype being developed:

In Australia, *The review of senior Secondary pathways into work* (Education Council, 2020) has already proposed that a national profile should be created for all school leavers. The Mastery Transcript Consortium (MTC, <https://pilot-transcript.mastery.org/>) is introducing extended transcripts, presented visually, using an international consortium of schools and universities. Drawing on work at High Tech High School in California, The Xperiential Learner Profile is a school-based example from England (<https://www.heritagexperiential.org/about-us/learner-profile/>). Meanwhile at Rethinking Assessment we are seeking to bring together data from digital badges, student portfolios and external credit awarders, explicitly building in a powerful element of student formative learning.

The idea of a National Record of Achievement (RofA), a portfolio of documents showcasing a student's academic and non-academic achievements is not new; in the 1990s, for example, the UK government encouraged all secondary students to develop such a Record. These RofAs did not work because the non-standardised elements were often too long-winded to be assimilated by others, it was not clear who was warranting the data and because there was not buy-in from employers. They were ideas before their time and before digital technology was freely available.

Fast forward to today and the world is a different place. Curricula are changing. Pedagogy is evolving. Higher and further education along with employers are looking beyond mere grades. And multi-modal assessment, carefully curated, should begin to enable students and teachers, with external warranting from appropriate bodies, to evidence the breadth of their talents. Such a profile, a personal url address probably, can then travel with its owner throughout her or his learning life. With enhanced teacher confidence in using the full repertoire of assessment methods, better ways of evidencing student capabilities and dispositions and a more effective way of visualising talents we can take a small stride towards truly strengths-based education.

Figure 2: Prototype Learner Profile (Milligan et al., 2020, p. 24).



Acknowledgements

This article draws on work being undertaken by Rethinking Assessment in the UK and in particular on thinking in an earlier paper by the author commissioned by the Centre for Strategic Education in Melbourne, Australia.

Correspondence

Professor Bill Lucas

Director, Centre for Real-World Learning
 The University of Winchester
 The Masters' Lodge
 Romsey Road
 Winchester
 SO22 5HT
 Co-Founder
 Rethinking Assessment
<https://rethinkingassessment.com>
 Email: bill.lucas@winchester.ac.uk

References

- Conley, D. & Darling-Hammond, L. (2013). *Creating systems of assessment for deeper learning*. Stanford, CA: Stanford Center for Opportunity Policy in Education.
- Darling-Hammond, L. (2017). *Developing and measuring higher-order skills: Models for State performance assessment systems*. Washington, DC: Council of Chief State School Officers.
- Department for Education (2014). *The national curriculum in England Framework document*. London: Department for Education.
- Education Council (2020). *Looking to the Future: The Report of the Review of Senior Secondary Pathways into Work, Further Education and Training*. Canberra: Department of Education, Skills and Employment.
- Education Policy Institute (2019). *'The forgotten third': a rapid review of the evidence*. Leicester: Association of School and College Leaders.
- Futures, H., Andersen, T. & Larsen, K. (2020). *A European approach to micro-credentials: Output of the micro-credentials Higher Education Consultation Group*. Luxembourg: Publications Office of the European Union.
- Gordon Commission on the Future of Assessment in education (2013). *To assess, to teach, to learn. A vision for the future of assessment*. Princeton, NJ: Educational Testing Service.
- Griffin, P., McGaw, B. & Care, E. (Eds.). (2012). *Assessment and teaching of 21st century skills*. New York: Springer.
- Hattie, J. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge.
- Hipkins, R. & Cameron, M. (2018). *Trends in assessment: an overview of themes in the literature*. Wellington: New Zealand Council for Educational Research.
- Howard, E. (2020). *A review of the literature concerning anxiety for educational assessments*. Coventry: Ofqual.
- Lucas, B. (2022). *A field guide for assessing creative thinking*. Perth: FORM.
- Lucas, B. (2021). *Rethinking assessment: The case for change*. Melbourne: Centre for Strategic Education.
- Lucas, B. (2019). *Why we need to stop talking about twenty-first century skills*. Melbourne: Centre for Strategic Education.
- Lucas, B. (2016). A five-dimensional model of creativity and its assessment in schools. *Applied Measurement in Education*, 29(4), 278–290.
- Masters, G. (2013). *Reforming Educational Assessment: Imperatives, principles and challenges*. Melbourne: Australian Council for Educational Research.
- Milligan, S., Luo, R., Kamai, T., Rice, S. & Keang, T. (2020). *Recognition of learning success for all: Ensuring trust and utility in a new approach to recognition of learning in senior secondary education in Australia*. Melbourne: Learning Creates Australia.
- Milligan, S., Luo, R., Hassion, E. & Johnson, J. (2020). *Future-proofing students: What they need to know and how to assess and credential them*. Melbourne: University of Melbourne.
- Mitra, S. (2020). Children and the internet: Learning in the times to come. *Journal of Learning for Development*, 73(3), 286–305.
- National Research Council (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: The National Academies Press.
- OECD (2017). "PISA 2015 collaborative problem-solving framework", in *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics, Financial Literacy and Collaborative Problem Solving*. Paris: OECD Publishing.
- OECD (2019). *PISA Creative Thinking Framework, Third draft*. Paris: OECD Publishing.
- Pellegrino, J., Chudowsky, N. & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Soland, J., Hamilton, L. & Stecher, B. (2013). *Measuring 21st Century Competencies: Guidance for Educators*. New York: Asia Society.
- Thurstone, L. (1927). A law of comparative judgment. *Psychology Review*, 34, 273–286.
- von der Embse, N., Jester, D., Roy, D. & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, 227, 483–493.
- Welsh Parliament (2021). *Curriculum and Assessment (Wales) Act 2021*. Cardiff: Welsh Parliament.
- William, D. (2007). Once you know what they've learned, what do you do next? Designing curriculum and assessment for growth. In R. Lissitz (Ed.), *Assessing and modeling cognitive development in schools*. Maple Grove, MN: JAM Press.
- World Economic Forum (2015). *New Vision for Education: Unlocking the Potential of Technology*. Geneva: World Economic Forum.