



CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT

Research Matters

Issue 35 / Spring 2023



Proud to be part of the University of Cambridge

Cambridge University Press & Assessment unlocks the potential of millions of people worldwide. Our qualifications, assessments, academic publications and original research spread knowledge, spark enquiry and aid understanding.

Citation

Articles in this publication should be cited using the following example for article 1: Suto, I. (2023). Creating Cambridge Learner Profiles: A holistic framework for teacher insights from assessments and evaluations. *Research Matters: A Cambridge University Press & Assessment publication*, 35, 6-26.

Credits

Reviewers: Gill Elliott, Pia Kreijkes, Jo Ireland, Tony Leech and Sylvia Vitello

Editorial and production management: Lisa Bowett

Additional proofreading: Alison French

Cambridge University Press & Assessment is committed to making our documents accessible in accordance with the WCAG 2.1 Standard. We're always looking to improve the accessibility of our documents. If you find any problems or you think we're not meeting accessibility requirements, please contact our team: Research Division, ResearchDivision@cambridge.org

If you need this document in a different format contact us, telling us your name, email address and requirements and we will respond within 15 working days.

All details are correct at the time of publication in March 2023.

Contents

- 4 Foreword:** Tim Oates
- 5 Editorial:** Tom Bramley
- 6 Creating Cambridge Learner Profiles: A holistic framework for teacher insights from assessments and evaluations:** Irenka Suto
- 27 A conceptual approach to validating competence frameworks:** Simon Child and Stuart Shaw
- 41 Teachers' and students' views of access arrangements in high stakes examinations:** Carmen Vidal Rodeiro and Sylwia Macinska
- 60 Who controls what and how? A comparison of regulation and autonomy in the UK nations' education systems:** Pia Kreijkes and Martin Johnson
- 80 Assessment in England at a crossroads: which way should we go?** Tony Leech
- 94 Research News:** Lisa Bowett

Foreword

My friend and colleague David Raffe, Professor of Sociology of Education at Edinburgh, died unexpectedly in February 2015. His quiet wisdom still is greatly missed. I continue to believe that the most important of his many contributions to the study of education systems was the brilliant 1999 paper “The case for ‘home internationals’ in comparative research” (Raffe et al., 1999). This outlined just how rich could be the comparisons between Scotland, Northern Ireland, Wales and England. This remains salient and important. It forms a key part of the methodological background to and motivation for the four-nation study included in this issue of *Research Matters*, and also helps in resolving some of the policy dilemmas and conflicts raised by the “future of education” debates described in the final article here. Why is his paper so important... still? The answer lies in the distinctive character of the British Isles – the existence of similar (though not identical) structures of education, common family structures, common labour markets and economic pressures, and so on. All this means that implementation of different models of education across the four nations comprises a massive, long-term natural experiment. In the 1970s and 80s, the differences in Scottish education were leading to higher equity and attainment rising more rapidly than other parts of the UK. But the 2010 introduction of major curriculum changes in schools has seen the Scottish lead disappear, and not just because of rising performance in England – both have introduced new national curricula in the last decade or so but, crucially, with very different models and principles driving those respective instruments. The fact that the nations are going in such different directions but are comparable in so many other ways gives this profound “crucible of comparison”. It’s what policymakers should pay huge attention to, and what other nations lack – a means of discriminating between the many conflicting voices and approaches in the calls for reform and innovation; a rich source of insights into what has brought about improvement and what has not. In addition, across the UK we enjoy a wealth of data beyond the periodic global transnational surveys (PISA, TIMSS, PIRLS) – although the early 2000s gap in national testing in Wales and the 2017–present cancellation of the Scottish Survey of Literacy and Numeracy removed some of the valuable data for tracing equity and attainment across all systems. And our national data allows robust longitudinal analysis, something which the cross-sectional studies like PISA (which tests successive groups of 15 year olds) cannot really achieve. David was right to point out just how powerful these “home internationals” are; and we keep alive his spirit of rich enquiry and policy support in the work we are continuing to do.

Tim Oates, CBE Group Director, Assessment Research and Development

Raffe, D., Brannen, K., Croxford, L., & Martin, C. (1999). Comparing England, Scotland, Wales and Northern Ireland: The case for ‘home internationals’ in comparative research, *Comparative Education*, 35(1), 9–25.

Editorial

The Covid-19 pandemic and its aftermath have prompted a lot of debate about the purpose of education and the role of assessment. All the articles in this issue touch more or less directly on these big themes.

In our first article Irenka Suto presents a conceptual framework for thinking about what “educational success” looks like and how teachers and school leaders might use different kinds of assessments to gain insights about the complete educational profile of their students. A particular focus for educational reform in many countries in recent times has been on how to assess various kinds of “competence”. Our second article, by Stuart Shaw and Simon Child, suggests a systematic general approach to defining and validating competence frameworks, based on clarifying claims and establishing evidence and arguments to support them.

For many years assessment organisations have provided ways for students with specific needs to access their assessments. Much research has focused on whether these access arrangements succeed in creating a “level playing field”. Less research has been carried out on what schools and students think of access arrangements, and how they use them. Our third article, by Carmen Vidal Rodeiro and Sylwia Macinska, fills this gap by reporting results from an international survey of schools taking Cambridge qualifications.

Our fourth article, by Pia Kreijkes and Martin Johnson, uses a detailed comparison of the four devolved national education systems in the UK as a basis for reflecting on issues of autonomy and control, in particular the role of the “middle tier”: organisations occupying the space between the central government and individual schools.

In the last year or so, educationalists and think tanks have been falling over themselves to pronounce on what the future of education is or should be. In our final article Tony Leech compares and contrasts seven different published reports about this desired or anticipated future in England, focusing on four areas of particular interest: high stakes assessment at age 16; how many subjects should be studied (and at what ages); the use of digital assessments; and the relationship between academic and vocational study.

Tom Bramley Director, Research Division

Creating Cambridge Learner Profiles: A holistic framework for teacher insights from assessments and evaluations

Irenka Suto (Cambridge CEM)

Introduction

The pandemic has led to school closures and online learning in many countries, including in England. Enormous numbers of learners have experienced upheaval and missed learning opportunities. For example, the United Nations (2022) estimates that 147 million children missed over half of their in-person instruction in 2020–21. Some children have lost previously acquired knowledge, and wellbeing and social skills have been affected too.

When schools have re-opened, teachers in England have been eager to make the best use of their time with learners. Some secondary school teachers have favoured a highly focused approach, targeting their teaching around the content of high stakes examinations for General Certificates of Secondary Education¹ (GCSEs) to ensure their learners achieve the best possible grades. Similarly, many primary school teachers have renewed their emphasis on reading (TES, 2020) and on other fundamental academic knowledge, skills and understanding.

Alongside this potential narrowing of the curriculum, the need to take a broader, more holistic approach to understanding and supporting learners has also become more salient to teachers of all age groups. This includes understanding learners' wellbeing and resilience during these times of adversity and uncertainty. In addition to its inclusion in the National Curriculum for England and in other curricula around the world, psychological wellbeing is worthy of attention for its own intrinsic value. Given that children typically spend around 15 000 hours in education (Rutter et al., 1979), teachers and parents want them to feel well, and to feel that they are doing well, during this substantial phase of their lives. Moreover, nurturing all aspects of a learner's growth acknowledges the broader social,

¹ GCSEs are qualifications in traditional school subjects and are obtained by most 16 year olds in England, Wales and Northern Ireland. They are a passport to higher-level study and are valued by many employers.

societal, and economic responsibilities of education.² Overall, it could be argued that Covid-19 has resulted in a reshaping of education, simultaneously narrowing and broadening different elements of it.

The need for a holistic framework for teacher insights into educational success

Periods of change provide an ideal opportunity to reflect upon the bigger questions of *what we want a high-quality education to achieve*, and *what that education could look like in practice*, especially in terms of the insights that teachers can gain along the way from associated assessments. Such questions are central to achieving the United Nations' fourth Sustainable Development Goal, on quality education: "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (United Nations, 2022). Broad, holistic conceptualisations of high-quality education can accommodate the idea that each school around the world is unique, embedded within different local and national cultures and legal frameworks. Each school's senior leaders must ultimately determine the school's own educational approach, answering the big questions (above) for themselves. This is important when thinking about how best to support schools to optimise learning and wellbeing.

Given the potentially overwhelming complexity of what a rich and holistic education entails, a simple conceptual framework which is grounded in evidence about what is really important to learners' educational success could help school leaders and teachers to make informed decisions, whichever educational philosophy or goals they choose to adhere to. In this article we present and explain such a framework, which articulates the insights that teachers can gain from different types of educational assessment and evaluation.

It is important to acknowledge from the start that human performance is highly integrated; that is, it has many different interacting components. Educationalists and psychologists seek to understand it by disaggregating it and by distinguishing different influences upon it. However, performance and contributory factors can be divided up in many different ways and to different levels of granularity. There are alternative ways to slice the cake, to create a series of easy-to-articulate components or "constructs" that will, in reality, interact in a variety of ways. Relationships among components are often complex and opaque. John Hattie's renowned syntheses of over 1850 meta-analyses explore success in schools in great detail, identifying hundreds of influences upon educational outcomes (Hattie, 2012; Visible Learning, 2022). However, it can be challenging to hold all of these in mind simultaneously, to see the bigger picture, and moreover, to unpick the relationships among the influences. To complement previous research, we would assert that a higher-level understanding and structure that focuses upon the roles of assessments would be helpful.

² Note that holistic conceptualisations of education are not new. As discussed by Suto et al. (2022), Jacques Delors' report to UNESCO, *Learning: The Treasure Within* (Delors et al., 1996), put forward a widely regarded holistic and integrated vision of education for the 21st century. This thinking draws in part from the deep-rooted educational philosophy of "Bildung" which has a long history in Germany and Scandinavia and denotes an educational ideal that can be traced back to Antiquity (Klafki, 1998, 2009).

In developing our framework, our main aim was to provide school leaders and teachers with a useful and memorable organising instrument. It has value in helping teachers to understand the main factors that influence learners' educational success, and hence the areas that a combination of different assessments would ideally cover during a learner's educational journey and why they are useful in a formative sense. We believe teachers could use the framework to combine numerical data from baseline and formative assessments with insights from observations, professional judgements, and learner discussions, to structure actionable learner profiles. They could also then identify complementary teaching and support strategies.

A new holistic framework

Our new holistic framework is presented in Figure 1. It comprises five interacting conceptual areas that contribute to learners' educational success, and into which we believe it is important for teachers to gain insights. These areas are: (i) cognitive skills and capabilities, (ii) cross-curricular knowledge, skills and understanding, (iii) subject domain knowledge (precursor curriculum coverage), (iv) teaching and learning environment, and (v) personal attributes. Learner data collected during the educational journey, both quantitative and qualitative, can be grouped in these five areas. Additionally, data can be collated on measures of educational outcomes and successes, such as qualification grades and progression to higher education institutions and employment.

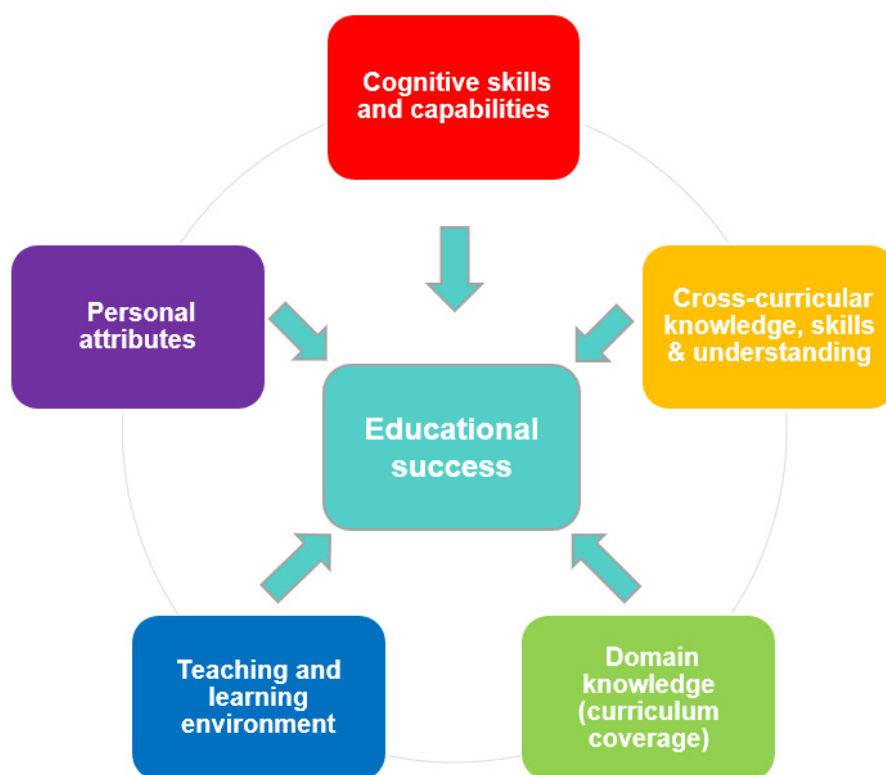


Figure 1: Our holistic framework for teacher insights from assessments and evaluations.

In the following sections of this article we explain each of the five areas of teacher insight in turn, to provide an evidence base for our framework.

Cognitive skills and capabilities

Our first area for teacher insight is that of cognitive skills and capabilities. Constructs in this area are assessed in many baseline and entrance tests, which are often taken at the start of the school year (for example [Cambridge Centre for Evaluation and Monitoring \(CEM\)'s baseline tests](#) and [GL Assessment's CAT4](#)). Cognitive skills are “curriculum free” in the sense that they do not typically feature in English or other school curricula. They include non-verbal reasoning and some types of verbal reasoning. CEM's assessments measure learners' ability in 3-D visualisation, spatial aptitude, pattern recognition and logical thinking.

In our recent analysis of assessments in this area (Suto et al., 2022) the assessment results of Year 6 learners (aged 10/11) in England were compared with their subsequent GCSE results (at age 16) in a range of school subjects. In line with earlier research (Deary et al., 2007), correlation coefficients were found to be high: for example, 0.77 for GCSE Mathematics, 0.72 for English, and 0.71 for Geography. These high levels of predictive validity are reassuring rather than surprising, given what is known about the constructs assessed.

Assessments of cognitive skills and capabilities have been described as measuring “*eductive ability*”, which is “the capacity to forge new knowledge, discern meaning in confusion, perceive, and identify relationships” (Querioz-Garcia et al., 2021, p. 85), and to “*make meaning*” (Raven, 2008). *Eductive ability* is a close relative of analytical thinking. In their renowned taxonomy of educational objectives, Bloom and his colleagues explain that *analysis* “emphasizes the breakdown of material into its constituent parts and the detection of the relationships of the parts and of the way they are organized” (Bloom et al., 1956, p. 144). Both non-verbal and verbal reasoning tasks require candidates to undertake precisely these analytical tasks. In their revision of the cognitive domain of Bloom's taxonomy, Anderson and Krathwohl (2001) present *analysis* as a higher order thinking skill, which is more demanding than retrieving information from memory, understanding information, and applying it. Analytical skills are required in a wide range of secondary school subjects. They are also included in pedagogical taxonomies of critical thinking (Black, 2012) and 21st century skills (Suto, 2013), which would suggest that they are important for success across broader conceptualisations of education such as “*Bildung*” (Klafki, 2009) and that of Delors et al. (1996).

It is worth noting that although cognitive skills and capabilities are very difficult to teach, learner data in this area is of great value to teachers. Insights can be used to assign learners to teaching groups or streams, and to anticipate the levels of support that they will need. Baseline, entrance, and similar assessments can thereby support targeted teaching and encouragement for everyone.

Cross-curricular knowledge, skills and understanding

Like cognitive skills and capabilities, cross-curricular knowledge, skills and understanding (KSU) are relevant to many different school subjects; however, they are much easier to teach. They include core mathematical concepts, vocabulary and language comprehension, and they are also assessed in baseline and entrance tests (Cambridge Centre for Evaluation and Monitoring, 2023a; 2023b). In our recent predictive validity analysis we confirmed them to be good predictors of subsequent academic performance in a range of subjects (Suto et al., 2022).

Why might this be? Mathematical thinking arises in school subjects as diverse as the sciences, computer science, economics, geography, business, and design and technology. Learners of these subjects are often required to perform calculations, interpret graphs, and measure things. Language comprehension is needed in all school subjects, since learners must comprehend what teachers and fellow learners are saying, what they read in textbooks and other teaching resources, and of course, what examination questions are requiring of them. Similarly, it is needed in non-assessed elements of education, such as personal, social and health education, and in elements that contribute to and constitute education in its broadest sense, including extra-curricular activities.

Vocabulary is critical to language comprehension (Quigley, 2018). To ensure comprehension, the reader must know 95 per cent of words in a text, and this percentage is as high as 98 per cent in many texts for older learners (Blachowicz & Fisher, 2015). It could be argued that cross-curricular KSU is a subdomain of domain knowledge, however, we would argue it is a discrete area of teacher insight. While learners acquire some cross-curricular KSU through the taught curriculum, much vocabulary and therefore language comprehension is acquired outside of lessons. This includes during extra-curricular activities and during social and other aspects of school life, as well as at home. Not only are basic everyday words, which are often acquired outside the classroom, important for good communication with all teachers and fellow learners, but they are also needed to understand the background contexts used in questions in subjects such as geography, history, economics, English language, and drama. Even in mathematics, problem-solving questions contextualise mathematical content to check that learners can apply their mathematical understanding in the “real world” (Beck et al., 2013).

Domain knowledge

Arguably, this area of the framework needs little explanation or justification since it is well known and highly intuitive that prior KSU in a given subject domain is also a solid predictor of educational outcomes in that subject. In England, for example, correlations between GCSE grades (typically at age 16) and A level grades (typically at age 18) in the same subject tend to be between 0.5 and 0.65 (Ofqual, 2017; Sutch, 2013). Similarly, the relationships between performance at GCSE and prior attainment in national curriculum assessments at the end of primary school (Standardised Assessment Tests, known as SATs) have also been explored and reported to be strong (Benton & Sutch, 2013). Put simply, the greater a learner’s

coverage of the curriculum and the deeper their understanding of the content, the better their performance in summative assessments is likely to be.

While the studies cited here relate to formal assessments, it is important to note that teachers can gain insight into their learners in this area through both internal and external assessments of varying formality and consequence. While some progress tests are offered by assessment organisations, others are constructed by teachers, and even less formally, day-to-day classroom interactions provide many opportunities to establish learners' domain knowledge.

Generalisation from a simple model of human performance

The final two areas of insight in our holistic framework were identified by revisiting some of our previous research on human performance in a different educational assessment context (Suto & Nádas, 2008). We propose that to increase educational success, on the one hand we can reduce “task demands” (see below) and on the other hand we can increase learners' personal expertise in learning.³

The idea of learning expertise is very similar to the multidimensional construct of “learning power”, which stems from the idea that learning is learnable (Deakin et al., 2004) and to which we return subsequently. However, “expertise” or “power” imply agency and are not inclusive enough terms to capture all the personal attributes that come into play in acquiring new knowledge, skills and understanding. For example, psychological wellbeing affects learning and can be affected by factors that are beyond a learner's personal control (McLellan, 2021). Wellbeing can decrease through no agency or fault of the individual. Capacity or readiness to learn could therefore be added to learning expertise to comprise the “personal attributes” that influence educational outcomes (see purple box in Figure 1). Such capacity has been defined by Maddox, Forte and Boozer (2000) as the degree to which learners have prerequisite cognitive, emotive attitudinal, and behavioural attributes, skills and orientations that will prepare them for involvement in learning.

Elements within this model have already been explored extensively in past research, and in the remainder of this article we bring together the two bodies of literature on these two major routes to optimising educational performance: (i) the teaching and learning environment and (ii) personal attributes (learning expertise/power/readiness). They complete our framework of five areas of insight for teachers (Figure 1).

Teaching and learning environment: reducing demand in learning

The term ‘reducing demand’ could easily be misinterpreted. In the present context, it is *not* about reducing curriculum content demand by decreasing the volume, depth, or breadth of domain coverage or by lowering the sophistication of the mental processing required of learners.⁴ In England during the pandemic, the

³ Our original research on marking expertise (Suto & Nádas 2008) showed that marking performance can be improved by decreasing the demand of the marking task and increasing the examiner's personal expertise.

⁴ See Suto, Grotorex, Vitello & Child (2020) for a detailed discussion of what increases and decreases curriculum content demand.

narrowing of secondary school teaching around the content of high stakes GCSE examinations (mentioned previously, and as instructed by the national regulator) is an example of “reducing demand” in this sense. There are clear risks associated with this approach. Arguably, the overall quality of education is impoverished and learners leave school with reduced KSU, which will serve them less well as preparation for employment or further study.

Instead, “reducing demand” in order to improve learners’ performance is broadly equivalent to applying cognitive load theory in education (Sweller, 1988). This is the longstanding idea that when curriculum content is held constant, cognitive load typically increases when unnecessary demands are imposed on a learner, making the task of processing information overly complex. In the context of education such demands include, for example, the unnecessary distractions that can occur in a classroom and inadequate teaching methods. When cognitive load is managed well, learners can learn new things more easily than when high cognitive load interferes with the creation of new memories (Anon., 2022). Sweller (*ibid.*) originally argued that instructional design can be used to reduce cognitive load in learners, and over 30 years later, Hattie’s (2022) reported effect sizes would certainly support that position.

We would argue that in order to understand cognitive load in education, teachers and school leaders need deep insight into what we label the learner’s “environment”; this is a fourth area of insight in our holistic framework. School leaders and teachers will want to adjust what is within their control to make the learner’s environment as conducive to learning as possible.

There are different ways of dissecting and articulating the learner’s environment. For example, McLellan (2021) writes:

“Bronfenbrenner (2005), in talking about human development, puts the individual at the heart of his ecological systems theory, that describes the interacting systems within which we exist in society. So, for example, a student is learning within the microsystem of a classroom interacting with staff and students. In turn this microsystem is nesting within and interacting with other systems outside the classroom – such as the family and local community, government policy and societal beliefs. All these systems evolve and change with time and what happens in one part of the ecosystem affects the rest of it, ultimately impacting the learner” (McLellan, 2021).

Kyriakides et al. (2020) offer a similarly dynamic model, focusing specifically on the educational environment. It includes: (i) system level, (ii) school level and (iii) classroom level factors affecting learners’ academic outcomes, and is underpinned by recent analyses and meta-analyses of studies in the burgeoning field of educational effectiveness and improvement.⁵ At the level of the system, control factors include national and regional policy for education, plus curriculum, pedagogy, and accountability. School-level factors include school leadership

⁵ Kyriakides et al. also include learner-level factors in their dynamic model. We consider these factors later in the paper.

practices, school policy, and the quality of implementation of such policy. Classroom-level factors relate to the quality of teaching and include: orientation, structuring, modelling, application, questioning, assessment, management of time, and the classroom as a learning environment (Creemers & Kyriakides, 2008).

Many of the numerous factors evaluated by Hattie (2022) can be placed within these three levels, and within the school and classroom levels in particular. For example, factors with relatively large effect sizes such as classroom discussion (0.82), scaffolding (0.82), and summarisation (0.79), which can all be elements of instructional design, come into play at the level of the classroom. Overall, Hattie asserts that the key to making a difference is to make teaching and learning visible. He argues that learning becomes visible when teachers are also learners (i.e., evaluators of their own teaching), and help learners to become their own teachers through metacognitive strategies, feedback and reciprocal teaching (Hattie, 2012).

To Kyriakides et al.'s three levels of influential factors, we can also add pastoral-level factors, which can be highly distinct in boarding schools, for example. In addition to the school or educational environment, a learner's home environment is also very important for education (Hattie, 2022). A great many aspects of home environment potentially influence educational outcomes, both physical and social. For example, for some learners, a challenging home environment can be one in which they receive little parental support, interest, and encouragement. In stark contrast, but also potentially just as challenging, other learners may have extremely demanding parents with unrealistic expectations for their academic achievement. This pressure may be coupled with little free time to unwind and relax.

Personal attributes

As discussed previously, in addition to reducing the demands of learning by focusing on educational environments, teachers can also improve educational outcomes by attending to their learners' personal expertise in learning and their learning readiness. We believe every learner possesses a unique combination of personal attributes beyond their cognitive ability that help (or hinder) their learning. "Personal attributes" are essentially emotive, attitudinal and behavioural descriptors of individual learners and we use the term broadly and simply. They are the fifth and final area of teacher insight in our framework.

Wellbeing

Personal attributes range from being relatively stable "traits" to more transient "states", although all can change to some extent. This is true even for personality traits which are at the more stable end of the spectrum (Rantanen et al., 2007). For example, Soto's (2015) analysis of a nationally representative sample of over 16 000 Australians reveals that personality traits and aspects of wellbeing reciprocally influence each other over time. Wellbeing is defined most comprehensively as a transient psychological state which combines both feeling well (hedonic wellbeing) and functioning well (eudaimonic wellbeing). It can change from month to month or even week to week, since it is influenced by life events (McLellan & Steward, 2015).

There is increasing evidence that wellbeing is linked to academic performance. For example, researchers in the UK (Gutman & Vorhaus, 2012) and the USA (Suldo et al., 2011) identified correlations between wellbeing and educational performance (including at GCSE) when they were assessed at the same point of time, for multiple age groups (from age 10 to 16 years); moreover, they identified similar associations when academic performance was assessed two years after wellbeing was assessed. More recently, an international literature review (Lindorff, 2020) concluded there is evidence of links between wellbeing and attainment and between whole-school approaches to wellbeing and attainment, but that the latter is heavily dependent upon implementation. The author also concluded that these relationships hold true across different contexts and countries, albeit with some variation.

Recent research on mechanisms to explain the link between wellbeing and readiness to learn, and therefore educational outcomes, has shown that better wellbeing is associated with more adaptive forms of motivation, such as wanting to learn and progress, rather than focusing on performance relative to others or avoiding learning situations (McLellan (2021) citing Wormington & Linnenbrink-Garcia (2017)). Additionally, wellbeing is associated with engagement. An influential and substantial annual survey of young people's wellbeing in the UK revealed that those with lower levels of wellbeing were more likely to truant (The Children's Society, 2018).

So-called 21st century skills and emotional intelligence

Looking beyond wellbeing, a large and important group of personal attributes are those that are often known as "21st century skills". This term stems from the view that the skills needed to compete in today's global economy are quite different from those upon which 19th and 20th century education systems have traditionally focused. According to Silva (2009), there are hundreds of descriptors of the skills set, including life skills, workforce skills, interpersonal skills, applied skills, and non-cognitive skills. We have argued previously that many skills of this kind, such as creativity, innovation, critical thinking, problem-solving, decision-making, learning to learn, metacognition, life and career skills, citizenship, and information literacy skills, are in fact not very new (Suto, 2013). They fall within Anderson and Krathwohl's (2001) revision of the cognitive domain of Bloom et al.'s (1956) taxonomy of educational objectives (Suto, *ibid*; Suto & Eccles, 2014). According to Silva (2009), creative, critical and analytical thinking skills have been articulated and valued by many philosophers and educators from Socrates 2400 years ago, to John Dewey in the 20th century.

Others have argued that many 21st century skills, including creativity, problem-solving, decision-making, communication, collaboration, citizenship, and personal and social responsibility, are linked inextricably to personality characteristics and so-called "emotional intelligence". Petrides (2001) and Petrides and Furnham (2003) have defined emotional intelligence as a constellation of behavioural dispositions and self-perceptions concerning one's ability to recognise, process, and utilise emotion-laden information. Emotional intelligence is further

conceptualised by the authors as an aspect of personality, which is malleable and still developing well into a person's twenties.

It is hard to deny that 21st century skills and emotional intelligence are a good thing to have and contribute to an individual's personal learning expertise and readiness. They are frequently mentioned in job advertisements, and some would argue that curricula and pedagogy should be structured around such attributes (e.g., RSA, 2022). Taking each attribute separately, a literature review will quickly engender a convincing argument for each one's worthiness of teacher attention. However, when it comes to determining the most worthwhile insights for teachers to obtain, we identify two distinct challenges.

Firstly, there are hundreds of attributes that could be considered 21st century skills and emotional intelligence, but there is little in the way of strong rationale as to why any particular combination of attributes is more relevant to education than any other. In addition to the skills and attributes mentioned so far, countless others are all valued by teachers, employers, and in society at large. These include being mentally fluent (Partington 2011), self-reflective (Shaw et al., 2018), articulate, resilient, responsible, confident, flexible, honest, motivated, hard-working, tolerant, and pragmatic, as well as having linguistic aptitude, common sense, integrity, and perseverance, and this list is far from exhaustive.

Secondly, valid and reliable measurements of many personal attributes of this kind are hard to come by. Self-assessment is difficult due to the Dunning-Kruger effect. That is, people tend to hold overly favourable views of their abilities in many social and intellectual domains. Not only do they reach erroneous conclusions and make unfortunate choices, but they also lack the metacognitive ability to realise it and they cannot evaluate themselves accurately (Dunning & Kruger, 2000). Although teacher assessment can be a better option, this relies upon teachers themselves having sufficient personal attributes to recognise and evaluate them in their learners.

Personal attributes with predictive validity

A constructive way up and out in this potential quagmire is for teachers to focus upon those attributes that researchers have found, so far, to predict educational outcomes. Not only are they of most (known) formative value, but evidence of predictive validity assumes a degree of robustness to assessment. Note that predictive validity here could be measured not only in terms of examination results, but also outcomes associated with the broader conceptualisations of education discussed previously, including employment.

The empirical and theoretical work of Deakin Crick and her colleagues centres around the well-validated, multidimensional construct of "learning power". It is derived from the idea that learning is learnable (Deakin Crick, 2007), and as mentioned previously, it is conceptually similar to the idea of personal expertise in learning. Learning power comprises seven basic dimensions, or *learning dispositions*, which Buckingham Shum and Deakin Crick (2012) describe "as a key requirement for life in the 21st century" (p. 2). The dispositions are: (i) resilience, (ii)

strategic awareness, (iii) learning relationships, (iv) creativity, (v) critical curiosity, (vi) making meaning, and (vii) changing and learning. The research group developed ELLI, the Effective Lifelong Learning Inventory, to assess learning power via the learning dispositions (ELLI Global, 2022). This self-report measure has been validated among 100 000 people globally, and the learning dispositions have been found to be correlated positively with standardised assessment outcomes (Buckingham Shum & Deakin Crick, 2012).

In line with these findings, Hattie (2022) has reported effect sizes of 0.35 for creativity, and of 0.48 for motivation as well as concentration, persistence and engagement, which are closely linked to Deakin Crick's construct of resilience. Similarly, Kyriakides et al. (2020) include perseverance and motivation as learner-level factors in their evidence-based model of influences on educational outcomes.

There is considerable overlap between learning dispositions and the constructs assessed in the [Cambridge Personal Styles Questionnaire \(CPSQ\)](#)⁶ and the "Big Five" personality model that underpins it. This five-factor model is well researched (e.g., Norman, 1963; Goldberg, 1992) and comprises (i) openness to experience, (ii) conscientiousness, (iii) emotional resources, (iv) extraversion, and (v) agreeableness. Its traits predict real-world behaviours that are important to participation in school and the workplace, such as productive study habits (Credé & Kuncel, 2008). A major meta-analysis by Poropat (2009), in which cumulative sample sizes extended to over 70 000, concluded that academic performance correlates significantly with openness, conscientiousness, and agreeableness, in primary, secondary and tertiary education. Research to explore the predictive validity of CPSQ directly has recently been conducted at Imperial College London. Among undergraduates, CPSQ's conscientiousness measure was found to be highly and significantly correlated with better exam results (Dale, personal communication).

Development of 21st century skills

Personality traits and dispositions tend to be quite stable over time (Rantanen et al., 2007). However, many teachers are interested in how the more malleable personal attributes develop and the extent to which they can be taught. We have argued previously that such development commonly occurs through lessons in traditional school subjects. Puntis (2011) explained how academic subjects such as mathematics and the sciences can also be reconceptualised in terms of the 21st century skills they engender, which include critical thinking, problem-solving, and creativity – a position shared by the Advisory Committee on Mathematics Education (2011). Many vocational courses can also nurture 21st century skills. Rose (2011), for example, has articulated some of the highly sophisticated analytical, problem-solving and creative skills developed on electricians' courses. Another longstanding and complementary perspective is that personal attributes of this

⁶ This online self-report assessment measures: (i) intellectual curiosity and open thinking, (ii) motivation to achieve and self-management, (iii) resilience and adaptability to demands, (iv) communication, and (v) collaboration.

kind can be nurtured successfully outside of lessons in extra-curricular activities (Haensly et al., 1985).

We share these views and would add that learners may *initially* develop each attribute in a small number of specific contexts. These could be either within the taught curriculum or external to it, including in social interactions during breaktimes and at home. This position stems from our review of Marzano and Kendall's (2007, 2008) respected taxonomy of educational objectives (Suto et al., 2020). The authors describe various metacognitive skills as levels of mental processing that are initially taught within specific domains of knowledge, building cumulatively upon less sophisticated mental processing. These metacognitive skills include monitoring: (i) goals, (ii) progress towards them, (iii) clarity (the degree to which you are free from ambiguity about the knowledge you are attempting to acquire), and (iv) accuracy (the degree to which you understand the given knowledge). Essentially, this means making learning an object of learning and reflection. At an even higher level of mental processing (which the authors term "self-system") these skills cover examining one's overall motivation, emotional response to learning, efficacy, and whether knowledge is important or meets a need or personal goal (Marzano & Kendall, *ibid.*).

Marzano and Kendall's taxonomy coheres with the findings of a review by Watkins (2001) which suggested that teachers can promote learning about learning or "meta-learning" by using classroom activities which make learning an object of attention, conversation, reflection, and learning. Watkins went on to argue that if meta-learning is to develop in classrooms, then two principles must apply: (i) the monitoring must engage the agency of the learners, and (ii) the language used must be owned by the learners themselves. These principles can be advanced through classroom practices such as noticing, narrating and navigating.

As learners' education progresses, we hold the view that their personal skills gradually generalise across contexts as learners make connections across subjects and test out and apply what they have learnt in new settings, consciously or otherwise. Generalisation of this kind is well known to occur in the development of vocabulary (Beck et al., 2018) and other aspects of language development (Tamminen et al., 2015). This would provide an important explanatory mechanism for generalisation for several metacognitive and other personal skills, given their reliance upon language, though there could well be many others.

We would cautiously suggest that by the time we reach adulthood, 21st century skills are often well generalised but continue to develop further. For example, many courses offered in the workplace support the further development of these skills. Also, many teachers, as working professionals, appreciate on a personal level the value of 21st century skills in the workplace and in life in general. They perceive them as valuable "goods" of education, and their own experiences and needs for professional development in adulthood could explain enthusiasm in some teaching communities to structure curricula and pedagogy around such attributes (e.g., RSA, 2022).

Tentatively relating these ideas to our framework of five areas of teacher insight in Figure 1, we suggest that during the course of childhood education and well into adulthood, metacognitive skills and other malleable personal attributes move from the green “domain knowledge” area, through the yellow “cross-curricular knowledge, skills and understanding” area, to the purple “personal attributes” area of insight. Different attributes are likely to transition at different paces.

Creating learner profiles

In Figure 2 we show how some assessments developed within our organisation offer teachers insight into their learners in the five areas in our holistic framework. A teacher could collate data from these assessments to create a “Cambridge Learner Profile” for each individual in their class as they progress along their educational journey. Using simple statistics, teachers and senior leaders within schools could also profile whole classes and year groups. Profiles of individuals and groups could also include data on educational outcomes (shown in turquoise in Figure 2) such as results in general and vocational qualifications and, more holistically, evaluations of the [Cambridge Learner Attributes](#).



Our framework of five areas of insight into learners

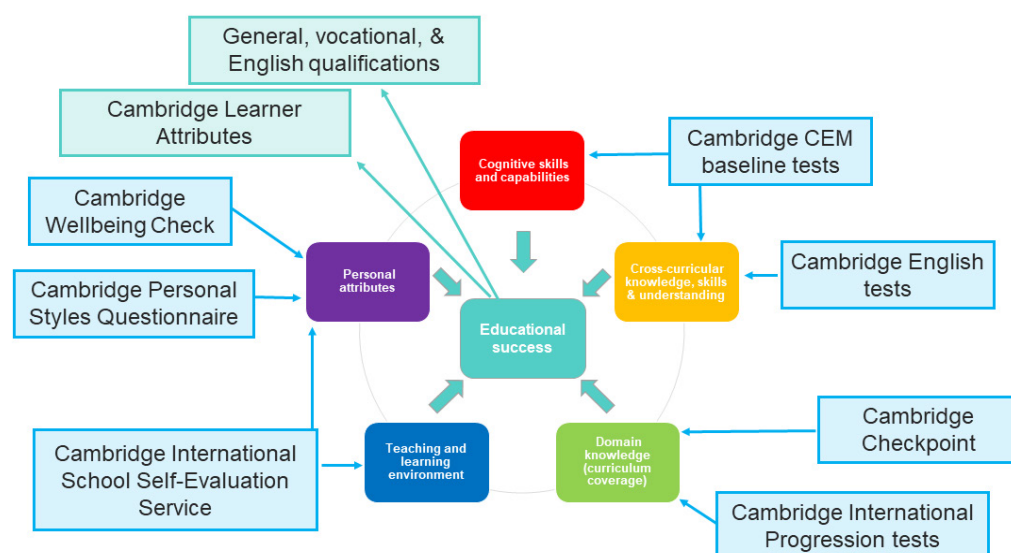


Figure 2: Mapping of Cambridge’s assessments to our framework for teacher insights.

In a sense, the five areas in the framework are “multi-purpose”: not only can they be used to articulate the main influences upon educational success, but they can also be used to categorise areas of education and the outcomes or goods of that education. That is, summative assessments, as well as baseline and formative assessments and associated curricula and teaching resources,⁷ can be positioned in each area. Through its inclusion of personal attributes in particular, the framework accommodates broad, holistic conceptualisations of education such as “Bildung” (Klafki, 2009) and that of Jacques Delors et al. (1996).

⁷ This is with the possible exception of teaching resources for cognitive skills and capabilities since they are very difficult to teach.

Teachers' own assessments and evaluations of their learners and schools can also be added to all areas in the framework, as could local and national assessments, plus any others that are used in a particular school. Indicators of educational success other than qualifications and the Cambridge Learner Attributes can also be mapped. A simple illustrative example is given in Figure 3.

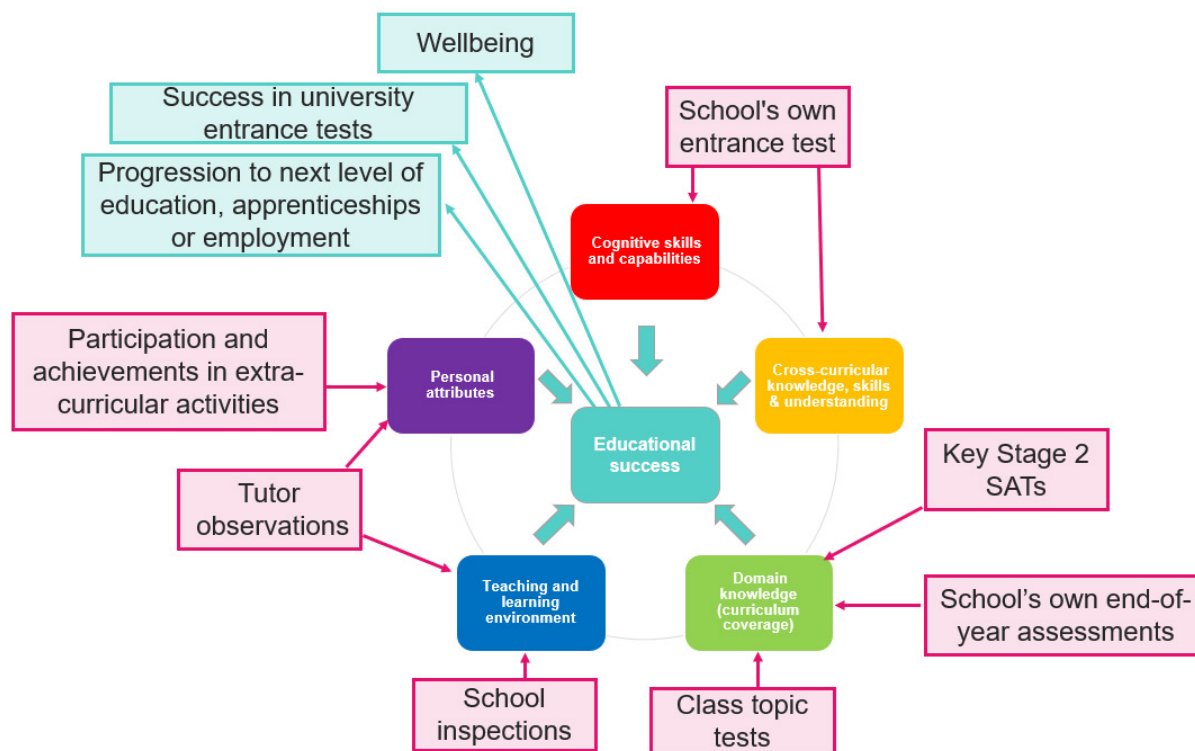


Figure 3: A simple mapping of some non-Cambridge assessments and evaluative approaches to our framework.

It is important to emphasise that the framework's main benefit lies in steering teachers' and school leaders' attention towards considering information in all five areas, rather than advising them on what to focus upon or which assessments to use. As mentioned previously, since every school is unique, its own senior leaders must ultimately determine its own approach to educational assessment. Returning to the bigger questions mentioned near the start of this paper, of what we want a high-quality education to achieve, and what that education could look like in practice, we stress that they are for each school leadership team to answer for itself.

To be successful, staff teams need members with the quantitative and data-handling skills to understand numerical assessment results (Schildkamp, 2019). As Figures 2 and 3 illustrate, and as mentioned previously, data from across the five areas of teacher insight could be brought together to create a profile for every learner, which could then be used formatively, to guide next steps in teaching, learning and pastoral support. Data from groups of learners could be used to shape class teaching plans and even whole school policies, around numeracy, literacy and wellbeing, for example. This is, in essence, the idea of data-based decision-making for school improvement, and teachers may also wish to integrate

their data with insights drawn directly from research findings, parents, learners' previous schools, and other evidence sources (Schildkamp, *ibid.*).

For many years, some schools have used spreadsheets to collate, link and analyse their data. More recently, others have begun to develop or purchase user-friendly dashboard software to draw together data from multiple sources, including learners' demographic data. In the future, we anticipate many further developments in this area. However, we acknowledge the importance of robust legislation to ensure data protection and privacy within an ethically defensible framework, and also the risk of drawing the wrong conclusions because of over-reliance on indicators that are easily quantified. Although it has been beyond the scope of this article to explore the numerous interactions that exist across constructs in our five areas of insight, software of this kind will undoubtedly facilitate such data analysis in schools and educational research in the future.

Just as educationalists such as Watkins (2001) and Marzano and Kendall (2007, 2008) advocate making learning an object of learning and reflection for learners, we believe this principle is also critical at the level of the teacher community. Arguably, it should be a whole-school mentality. However, approaches that depend heavily upon various forms of assessment and evaluation can, if done badly, distract from learner agency and responsibility. Rather than thinking of learners as objects and in terms of the numbers associated with them, there is more to be gained all round from viewing them as independent agents who need more skilled and reflective self-agency (Watkins, 2001), and by actively including them on the journey to improvement.

The professional judgements that contribute to teachers' wider insights and decision-making include those on academic achievement but also those around progress and outcomes in other aspects of school life. These include speaking up in class, performance in school debates, sportsmanship on the playing field, personal organisation, handling disappointment, and so on. These judgements can be made through careful observation but importantly, also through deep engagement with the learners themselves via discussions and other interactions. To be successful, school staff teams need self-confidence in their personal expertise and ultimately, in their ability to draw upon both quantitative and qualitative data, both of which are critical to the approach we have advocated.

Conclusion

To summarise, in this article we have presented an evidence-based, holistic framework of five interacting areas of insight into educational success. These areas are: (i) cognitive skills and capabilities, (ii) cross-curricular knowledge, skills and understanding, (iii) domain knowledge, (iv) teaching and learning environment, and (v) personal attributes. We believe the framework will be useful for teachers and school leaders in making sense of the vast number of assessments and evaluation tools that are available to them from our organisation and many others. There is a risk that data and information on learners is so plentiful that, at times, it can be difficult to see the wood for

the trees. We hope that our framework helps teachers to perceive the wood by organising data, combining it into learner profiles, and using insights appropriately to inform teaching. It also has the potential to highlight gaps in insight, where further assessment and evaluation could be beneficial.

Acknowledgements

Colleagues from across Cambridge University Press & Assessment have contributed ideas and information to this article and I would like to thank them all for their input. They include Kate Bailey, Dan Bray, Elizabeth Cater, Lee Davies, Catherine McKenna, Tim Oates, Rod Smith, and Tristian Stobie.

References

- Advisory Committee on Mathematics Education. (2011). *Mathematical needs: Mathematics in the workplace and in Higher Education*. The Royal Society.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Anon. (2022). *Cognitive load theory: How the cognitive load of a learning task affects a person's ability to memorize it*. *Psychologist World*.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. (2nd ed.). Guildford Press.
- Benton, T., & Sutch, T. (2013). *Exploring the value of GCSE prediction matrices based upon attainment at Key Stage 2*. Cambridge Assessment Research Report. Cambridge Assessment.
- Blachowicz, C., & Fisher, P. J. (2015). *Teaching vocabulary in all classrooms* (5th ed.). Pearson.
- Black, B. (2012). *An A to Z of critical thinking*. Continuum.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals; Handbook I: Cognitive Domain*. Longman.
- Bronfenbrenner, U. (2005). *Making human beings human: Bioecological perspectives on human development*. Sage Publications Inc.
- Buckingham Shum, S., & Deakin Crick, R. (2012). Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics. Proc. 2nd International Conference on Learning Analytics & Knowledge (29 Apr–2 May, Vancouver, BC). ACM Press: New York
- Cambridge Centre for Evaluation and Monitoring. (2023a). *Cambridge Primary Insight for ages 5-11*.
- Cambridge Centre for Evaluation and Monitoring. (2023b). *MidYIS for ages 11-14*.
- Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, 3(6), 425–453. doi.org/10.1111/j.1745-6924.2008.00089.x
- Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge.
- Deakin Crick, R. (2007). Learning how to learn: The dynamic assessment of learning power. *The Curriculum Journal*, 18(2), 135–153. doi.org/10.1080/09585170701445947
- Deakin Crick, R., Broadfoot, P., & Claxton, G. (2004). Developing an effective lifelong learning inventory: The ELLI project, *Assessment in Education*, 11, 248–272. doi.org/10.1080/0969594042000304582

Deary, I., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35(1), 13–21. doi.org/10.1016/j.intell.2006.02.001

Delors, J. et al. (1996). *Learning: the treasure within*. Report to UNESCO of the International Commission on Education for the Twenty-first Century. UNESCO Publishing.

Department for Education. (2014). *National curriculum in England: framework for key stages 1 to 4*.

Dunning, D., & Kruger, J. (2000). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–34. doi.org/10.1037//0022-3514.77.6.1121

ELLI Global. (2022). *ELLI reveals Learning's 7 Dimensions*. <https://elli.global/>

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26. doi.org/10.1037/1040-3590.4.1.26

Gutman, L. M., & Vorhaus, J. (2012). *The Impact of Pupil Behaviour and Wellbeing on Educational Outcomes*. Department for Education Research Report DFE-RR253.

Haensly, P. A., Lupkowski, A. E., & Edlind, E. P. (1985). The role of extracurricular activities in education. *The High School Journal*, 69(2), 110–119.

Hattie, J. (2012). *Visible learning for teachers: Maximising impact on learning*. Routledge.

Klafki, W. (1998). Characteristics of critical-constructive Didaktik. In B. B. Gundem & S. Hopmann (Eds.), *Didaktik and/or curriculum: An international dialogue* (pp. 307–330). Peter Lang.

Klafki, W. (2009). Didaktik analysis as the core of preparation of instruction. In I. Westbury, S. Hopmann, & K. Riquarts (Eds.), *Teaching as a reflective practice: The German Didaktik tradition* (pp. 197–206). Lawrence Erlbaum Associates.

Kyriakides, L., Creemers, B. P. M., Panayiotou, A., & Charalambous, E. (2020). *Quality and equity in education: Revisiting theory and research on educational effectiveness and improvement*. Routledge.

Lindorff, A. (2020). *The impact of promoting student wellbeing on student academic and non-academic outcomes: An analysis of the evidence*. Oxford University Press.

Maddox, N., Forte, M., & Boozer, R. W. (2000). *Learning readiness: an underappreciated yet vital dimension in experiential learning*. *Developments in Business Simulation & Experiential Learning*, 27.

Marzano, R. J., & Kendall, J. S. (2007). *The New Taxonomy of Educational Objectives* (2nd ed.). Corwin Press Inc.

Marzano, R. J., & Kendall, J. S. (2008). *Designing & assessing educational objectives; applying the new taxonomy*. Corwin Press Inc.

McLellan, R. (2021, July 19). [Supporting wellbeing in school: enabling young people to fulfil their potential](#). *Cambridge Assessment International Education blog*.

McLellan, R. & Steward, S. (2015). Measuring children and young people's wellbeing in the school context, *Cambridge Journal of Education*, 45(3), 307–332. doi.org/10.1080/0305764X.2014.889659

Microsoft. (2022). [What is Power BI?](#)

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality rulings. *Journal of Abnormal and Social Psychology*, 66, 574–583. doi.org/10.1037/h0040291

Ofqual, (2017). [Progression from GCSE to A level: Comparative Progression Analysis as a new approach to investigating inter-subject comparability](#).

O'Toole, C., & Simovska, V. (2022). Wellbeing and Education: Connecting Mind, Body and World. In R. McLellan, C. Faucher, & V. Simovska (Eds.), *Wellbeing and schooling: Cross cultural and cross disciplinary perspectives* (pp. 21–33). Springer: Switzerland. doi.org/10.1080/03323315.2022.2135572

Partington, R. (2011, March). Cambridge University and education in an interconnected world [Paper presentation]. Cambridge Assessment conference 'What kind of education enables us to cope with an interconnected world?' London.

Petrides, K. V. (2001). A psychometric investigation into the construct of emotional intelligence. Doctoral dissertation. (London, University College London).

Petrides, K. V., & Furnham, A. (2003). Trait emotional intelligence: Behavioural validation in two studies of emotion recognition and reactivity to mood induction. *European Journal of Personality*, 17, 39–57. doi.org/10.1002/per.466

Poropat, A. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338. doi.org/10.1037/a0014996

Puntis, A. (2011, March). An Interconnected World [Paper presentation]. Cambridge Assessment conference 'What kind of education enables us to cope with an interconnected world?' London.

Querioz-Garcia, I., Espirito-Santo H., & Pires, C. (2021). [Psychometric properties of the Raven's Standard Progressive Matrices in a Portuguese sample](#). *Portuguese Journal of Behavioral and Social Research*, 7(1), 84–101.

Quigley, A. (2018). *Closing the vocabulary gap*. Routledge.

Rantanen, J., Metsapelto, R.-L., Feldt, T., Pulkkinen, L., & Kokko, K. (2007). Long-term stability in the Big Five personality traits in adulthood. *Scandinavian Journal of Psychology*, 48, 511–518. doi.org/10.1111/j.1467-9450.2007.00609.x

Raven, J. (2008). [Introduction to the Raven Progressive Matrices Tests: Conceptual Basis, Measurement Model, and a Few Findings](#). In J. Raven & J. Raven (Eds.), *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics* (pp. 16–68). Royal Fireworks Press.

Rose, M. (2011). *Rethinking remedial education and the academic-vocational divide: Lessons to learn about language, cognition, and social class*, [Paper presentation] Annual conference of the American Educational Research Association, USA, Louisiana, New Orleans.

RSA. (2022). *What is RSA Opening Minds?*

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours*. Harvard University Press.

Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, 6(3), 257–273. doi.org/10.1080/00131881.2019.1625716

Shaw, S., Kuvalja, M., & Suto, I. (2018). *An exploration of the nature and assessment of student reflection*. *Research Matters: A Cambridge Assessment Publication*, 25, 2–8.

Silva, E. (2009). Measuring skills for 21st-century learning. *Phi Delta Kappa*, 90(9), 630–634. doi.org/10.1177/003172170909000905

Soto, C. J. (2015). Is happiness good for your personality? Concurrent and prospective relations of the big five with subjective well-being. *Journal of Personality Psychology*, 83(1), 45–55. doi.org/10.1111/jopy.12081

Suldo, S. M., Thalji, A., & Ferron, J. (2011). Longitudinal academic outcomes predicted by early adolescents' subjective well-being, psychopathology, and mental health status yielded from a dual factor model. *The Journal of Positive Psychology*, 6(1), 17–30. doi.org/10.1080/17439760.2010.536774

Sutch, T. (2013). *Progression from GCSE to AS and A level, 2010*. Statistics Report Series No. 69.

Suto, I. (2013). *21st Century Skills: Ancient, ubiquitous, enigmatic?* *Research Matters: A Cambridge Assessment publication*, 15, 2–8.

Suto, I., Benton, T. & Copestake, G. (2022). *What influences success in education? A predictive validity analysis and a holistic model of teacher insights from assessments*, [Paper presentation] Annual conference of the Association for Educational Assessment – Europe, Dublin, Republic of Ireland.

Suto, I. & Eccles, H. (2014). *The Cambridge approach to 21st Century skills: definitions, development and dilemmas for assessment*, [Paper presentation] Annual conference of the International Association for Educational Assessment, Singapore.

Suto, I., & Nádas, R. (2008). What determines GCSE marking accuracy? An exploration of expertise among maths and physics markers. *Research Papers in Education*, 23(4), 477–497. doi.org/10.1080/02671520701755499

Suto, I., Grotorex, J., Vitello, S., & Child, S. (2020). *A way of using taxonomies to demonstrate that applied qualifications and curricula cover multiple domains of knowledge*. *Research Matters: A Cambridge Assessment Publication*, 30, 26–34.

- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2) 257–285. doi.org/10.1207/s15516709cog1202_4
- Tamminen, J., Davis, M. H., & Rastle, K. (2015). From specific examples to general knowledge in language learning. *Cognitive Psychology*, 70, 1–39. doi.org/10.1016/j.cogpsych.2015.03.003
- TES. (2020, June 2). Reading ‘should be top priority’ for returning pupils. *TES Magazine*.
- The Children’s Society. (2018). *The good childhood report 2018*.
- Thorndike, R. L. (1947). *Research problems and techniques* (No. 3). US Government Printing Office.
- United Nations. (2022). *Do you know all 17 SDGs?*
- Visible Learning. (2022). *Hattie ranking: 252 Influences and effect sizes related to student achievement*.
- Watkins, C. (2001). *Learning about Learning enhances Performance*. Institute of Education School Improvement Network (*NSIN Research Matters* series No 13).
- Watkins, C. (2015). Meta-learning in classrooms. In D. Scott & E. Hargreaves (Eds.), *SAGE Handbook of Learning* (pp. 321–330). Sage Reference.
- Wormington, S. V., & Linnenbrink-Garcia, L. (2017). A new look at multiple goal pursuit: The promise of a person-centered approach. *Educational Psychology Review*, 29(3), 407–445. doi.org/10.1007/s10648-016-9358-2

A conceptual approach to validating competence frameworks

Simon Child (Cambridge Assessment Network) and Stuart Shaw (formerly Cambridge Assessment International Education¹)

Introduction

The development of competencies is regarded as an important focus in education and in the workplace. The responsibility for developing competencies in learners at all levels of education is increasingly being laid onto educational institutions (European Commission, 2018). This presents a series of significant conceptual and empirical challenges, including the need to achieve a consensus on what is meant by the term “competence” and related terms such as “competency” and “competencies”. A recent review by Vitello et al. (2021) recommended use of Hyland’s (1994) definition of “competence” as broad qualities in relation to a defined standard (e.g., a competent medical doctor). On the other hand “competencies” are narrower atomistic components that are linked to overall competence (e.g., appropriate completion of a medical procedure). We follow that definition in this article and hence talk about “competence frameworks” rather than “competency frameworks”.

Because competence in any domain is a complex concept (construct), it is useful to have a systematic definition of the overall competence and the competencies that comprise it, along with a statement of rationale (why the definitions are the way they are), and how the definitional system can be used in practice. This combination of definitional system, rationale and proposed use is what we call a “framework”. This concept of a framework is quite general. In this article we focus on competence frameworks, but many of the ideas and arguments would apply to validating other similar frameworks.

Competence frameworks have been developed in a range of educational, vocational and workplace contexts. A primary aim of such frameworks is to provide a structure that articulates an overall competence as well as individual competencies (Chartered Institute of Professional Development, 2021).

Competence frameworks typically offer a description of a range of competencies and the relations between them. They offer a “construct definition”: a statement of the knowledge, skills, and understanding specific to a context such as a workplace. This definition can then be put to a range of uses including to define assessment criteria, to act as a tool for personal reflection about development needs and

1 Undertaken while working for Cambridge Assessment

opportunities, or to offer a set of criteria to support in-role accountability. Competence frameworks reflect values held by their developers (Batt et al., 2019). Thus, they hold the potential to direct and influence models of curriculum, learning and assessment at all stages of education. Competence frameworks, with their emphasis on the application of knowledge in real-world situations, help to ensure that learners who have met the assessment objectives of assessments constructed from them are ready to “function effectively in society” (Mulder et al., 2007, p. 68), equipped with all the necessary knowledge, skills and attitudes required for personal fulfilment and social inclusion. (See European Parliament & Council of the European Union, 2006.)

This article focuses on the important question of how practitioners should validate competence frameworks. In the broadest sense of the term, validation is the process of using a range of methods to check the validity or accuracy of something. In the first section of this article, we explore the question of what it means for a competence framework to be deemed “valid”, drawing upon literature from educational measurement. We argue that a competence framework’s validity relates to the alignment between defined purposes, its structural elements and the overall credibility of associated claims made by framework developers.

This article then explores four methodological issues that permeate the validation of competence frameworks. On-going validation of competence frameworks is important, for example, because it acts as a defence against redundancy of frameworks over time. This is particularly relevant in workplace contexts, where new innovations, technologies, best practice, regulatory frameworks and so on can quickly render existing competence frameworks out of date. Developing sound and replicable validation methods to support the initial design, review and adaptation of competence frameworks can help users find areas of divergence between the framework and best practice.

We conclude this article by suggesting a practical template of key questions to consider when designing a competence framework to support its initial and on-going validation.

What is a valid competence framework?

Establishing a definition of validity that supports the investigation of competence frameworks is important because it can guide subsequent validation practice. In this section, we draw on the educational, psychological measurement, and assessment literature to propose reformulations of the concepts of “validity” and “validation” such that they are relevant to competence frameworks.

Validity as exploring credibility of stated claims linked to the framework

The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014) state that validity resides in the claims made about assessment outcomes and the strength of the arguments and evidence

made to support those claims. In this view, validity is based on the purpose(s) of an assessment and how well the interpretations and uses of the assessment outcomes (derived scores) serve each of the intended purposes (AERA et al., 2014). For competence frameworks, understanding their intended purposes is important because it underpins the subsequent development of claims that are the focus of validation inquiry. We explore the different types of claims made by competence frameworks in the next section.

When considering the validity of a competence framework, a useful concept is that of *credibility*, that is, the quality of being trustworthy or believable (House 1980, 2014). Credibility is achieved by understanding the requirements of the stakeholders who will be using the framework and providing them with enough evidence that the framework, in their view, is “sound” (AERA et al, 2014). Whether a competence framework argument can be “sound” or not depends on the sufficiency and relevance of evidence in support of the stated claims (see below for a list of types of claims made by competence frameworks). Cronbach (1988) outlined three criteria that link to overall credibility – *clarity*, *persuasion*, and *plausibility* – arguing that a validity argument must reflect the “prevailing beliefs and values” (p. 5) of all relevant stakeholders for it to be a just presentation of the validation evidence (see also Kane, 2013).

We have suggested above that the strength of a validity argument and the evidence in support of its claims resides in its credibility (not in its certainty) and its ability to persuade relevant stakeholders of the “soundness” of the overall claim(s). We therefore define the validity of a competence framework as:

An interpretive judgement as to the degree to which the claims regarding the use or uses (either declared or implicit) inferred from a competence framework are credible.

And a validity argument as:

A clear, comprehensible and persuasive defence of the extent to which relevant evidence as well as underlying theory support the purposes and intended uses declared for the competence framework. The argument is subject to ongoing scrutiny and challenge and can be overturned in certain specific circumstances. As such, any validity claim is at best tentative.

These definitions highlight the principal elements that describe validity and validation in relation to the purposes and uses of competence frameworks.

What claims are made by competence frameworks?

An important step in validating a competence framework is understanding the claims made related to its (potentially many) interrelated purposes. A *claim* in this context is a statement or assertion that something is the case, which creates a position that requires validation. Crucially, however, a claim is typically presented initially without providing evidence or proof – in this sense it is initially unsubstantiated. The stated claims thus become the subject of evidence collection, scrutiny, and challenge. Possible claims could relate to the application of a competence framework to educational stages and cultural contexts, the

effect that embedding the competence framework might have on practice, or the quality of the judgements made when utilising the framework (e.g., for assessment, Newton, 2017; Batt et al., 2019). Below are some illustrative examples of overarching claims made concerning competence frameworks (Table 1).

Table 1: Illustrative claims made in relation to competence frameworks.

Organisation	Competence framework	Example competence framework claim
Partnership for 21st Century Skills (P21) (http://www.p21.org/)	A framework for twenty-first century skills	“Embedding the 3Rs [reading, writing and arithmetic] and the 4Cs [collaboration, critical thinking, communication and creativity] makes teaching and learning more relevant, engaging and rigorous”
Chartered Insurance Institute (2015)	CII Insurance competence framework	“... [The CII Insurance competence framework] can support a wide range of business operations providing insurance services, both in the UK and globally”.
European Commission (2007)	Key competencies for life-long learning	“The key competences [competencies] are all considered equally important, because each of them can contribute to a successful life in a knowledge society”

The challenge for practitioners is to find evidence to substantiate the claim or claims that have been made by the competence framework developers. In essence, the validation of a competence framework is an attempt to establish its “fitness-for-purpose”.

From an analysis of the competence framework literature (e.g., Baczynska et al., 2016; Patterson et al., 2013), we have identified claims that fall into four main categories, which are summarised in Figure 1 (while acknowledging that other categories may exist). The claims listed in bold represent the main claim categories that we have identified, and the bulleted statements represent the claim sub-categories that are related to the main claim categories. Note that these categories of claims are not made by all competence frameworks, nor are they mutually exclusive to one another.

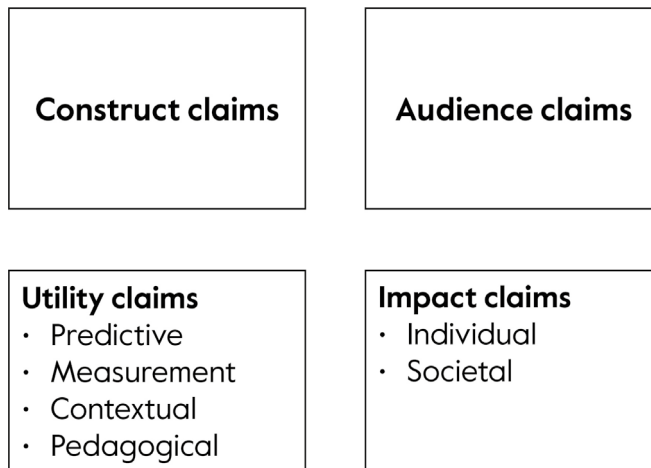


Figure 1: Competence claim categories and sub-categories.

Construct claims

Construct claims are about the definition of the competencies: how the elements are defined and labelled; how they combine to constitute overall competence; and the rationale for defining and organising things in this way rather than some other way. They relate to how well the definitions reflect the stakeholders’ general view of the knowledge, skills, understanding etc. required in their context.

Audience claims

Audience claims refer to the groups of people who should find utility in the competence framework as described. This refers to *users* of the framework – for example those designing assessments based on the framework (such as teachers, exam boards), those who are responsible for making educational decisions based on the competencies within the framework (e.g., employers), and those that are working towards competence within the framework.

Utility claims

Utility claims refer to the potential *uses* of the competence framework. There are four main types of utility that we identify here. *Predictive* claims concern how acquiring the overall competence (and individual competencies) defined in the framework prepares learners for future educational stages or employment. *Measurement* claims concern the utility of the competence framework for the purposes of educational assessment, for example the development and design of assessment linked to the competencies within the framework. Assessment could serve a variety of purposes, including for future learning (known as formative assessment) or as a summation of a learning stage (known as summative assessment). *Contextual* claims are statements concerning the range of contexts in which a competence framework is meant to apply. For example, competence frameworks may make claims about which stages of education the framework would be useful for, or in which countries or contexts the framework could be used effectively. *Pedagogical* claims refer to the use of the competence framework to inform teaching.

Impact claims

Impact claims describe the influence that engagement with the described competence framework would have in terms of its broader impact. *Individual* claims refer to the life outcomes of learners who acquire the overall competence and competencies described in the framework, for example in terms of their earnings and social mobility. *Societal* claims link to the broader economic and social effects resulting from adoption of the competence framework (e.g., economic output or productivity).

At least one *construct* or *audience* claim was found in most of the competence frameworks we analysed in the literature. It is likely that these claims underpin the initial development of a competence framework and its structure, and so are a primary target for any validation exercises. *Utility* and *impact* claims were less commonly observed, and we speculate that this might be due to the methodological challenges related to collating data relevant to the validation of these claims. We explore some of these challenges in the next section.

What are the methodological challenges concerning the validation of competence frameworks?

We have argued above that validation of a competence framework is the collection of theoretical arguments or evidence to provide credibility to the one or more claims made by the developers or users of a competence framework. An important element in exploring the validation of competence frameworks is understanding the common methods adopted and some of the methodological challenges that can potentially undermine validation inquiry. In this section, we briefly describe four main methodological problems common to competence framework validation.

Validation methods that focus only on construct claims

The first issue arises when the focus of the validation exercise is limited to the definitional aspects of the competence framework, without considering any audience, utility or impact claims. While construct claims are important to validate, there is always a “real-world” application to a competence framework that needs to be acknowledged and investigated (Priestley & Sinnema, 2014).

The limited nature of validation exercises is perhaps in part due to the selection of methods that engage stakeholders in activities that seek to establish agreement of terminology used with the frameworks. For example, a common method to validate competence frameworks is the Delphi method which has been used in a variety of validation studies in diverse areas such as veterinary science (Bok et al., 2011), nursing (Miranda et al., 2018) and mathematics teaching (Muniz-Rodriguez et al., 2017). While the specifics of Delphi methods vary, broadly they require a panel of experts to arrive at a consensus opinion about the overall validity of the competence framework. This might be through panel discussions, questionnaires, or structured interviewing. As Delphi methods are often used at an initial stage of competence framework development, evidence is rarely collected regarding how stakeholders use the framework. Validation evidence is thus limited to whether the definitions, labels and rationale of the framework make intuitive sense.

Other methods include surveys of large groups of practitioners with the intention of checking internal consistency using statistical techniques such as factor analysis. For example, Sastre-Fullana et al. (2017) sent 54 competence framework items to 600 nursing practitioners. Using factor analysis, they identified eight competency domains, reduced from their original conceptualisation of nine. Although this study utilised a three-stage validation approach, Sastre-Fullana et al. (2017, p. 8) concluded that “further criterion and evidence validity is a needed step forward from the actual position”. In our view, this statement is an acknowledgement that both the *pedagogical* and the *predictive* claims made regarding the competence framework have not yet been validated, specifically the framework’s suitability to be used to support teaching interventions and to predict advanced nursing competence. This highlights a weakness of validation exercises that focus exclusively on the construct claims rather than how they are used in real-world contexts.

Influence of group interactions on validation outcomes

The Delphi method is regarded as a “group consensus” method for validation (Bok et al., 2011). Groups of experts are asked to rank, rate, and discuss the relevance and/or appropriateness of the developed competencies and reach an agreed position. Other methods that utilise a group consensus approach to validation include the Group Sort Method (Ling et al., 2017) in which participants are asked to sort indicators or statements into predetermined categories representing levels of importance, and focus group methods (e.g., Rissi & Gelmon, 2014) in which experts are consulted on the form and anticipated utility of the competence framework, with the scope to introduce changes.

One of the main issues with validation approaches that aim to reach a group consensus is that the outcomes are influenced by the interactions among group members. While group consensus methods are by their nature collaborative and benefit from a defined outcome, problems can emerge when interactions within the group become non-optimal. For example, dominant individuals can have a pervasive influence on the conclusions reached by the group that are not a fair reflection of the group members’ opinions. While some methods attempt to counter this issue by trying to ensure that individual comments are anonymised it is difficult to guarantee, particularly in cases where there are few experts to draw upon (Miranda et al., 2018).

Anchoring effects of validation exercises

The validation approaches described above also introduce the potential for anchoring effects. For example, in the Delphi method approach, the panel of experts are typically recruited with the expectation that they will comment on, approve, and reject specific competency statements. This post-hoc approach to validation reduces the possibility for the panel of experts to *create* relevant statements based on their experience. There are also validation studies that have largely been confirmatory in nature (e.g., Soh et al., 2012) which limits the potential for experts to influence the final form of the competence framework.

Collection of validation evidence from users of the framework

A final methodological issue of competence framework validation relates to the range of users that are engaged to substantiate claims made by framework developers. It is usual for only a sub-set of types of user of the framework to be targeted in the collection of validation evidence. For example, Patterson et al. (2013) developed a competence framework for trainee general medical practitioners in the UK. They used a three-stage method in the development of their framework that comprised stakeholder interviews, a validation questionnaire, and a final expert panel review with general practitioners. What is interesting to note about this study is that, although patient representative groups were consulted in the initial development of the competence framework, they were not asked to contribute to the final validation as part of the questionnaire or expert review. Patterson et al. (2013) note that this omission was due to constraints in accessing patients. Given the intended use of the framework was to “explore the optimal construction of the education, training and career pathway to support trainees” (p. 337), this creates a requirement to engage with new groups including medical schoolteachers, and patients to establish the framework’s overall predictive impact.

Similarly, Baczynska et al. (2016) used a self-report method to validate the construct claims made by their framework which aimed to describe general job role competence and employability that are “similar in most organisations” (p. 10). While this study aimed to validate a *construct* claim, they themselves admitted that this was not validated sufficiently in their study. Importantly, there were additional *predictive* claims that were not explored, namely the claim that engagement with the framework will lead to better employability outcomes for learners. They noted issues with the sampling range that their validation method had access to in terms of culture and organisation type, and access to higher-level management staff.

Agenda for determining the validity of competence frameworks

As we have argued above, judgements about the validity of competence frameworks rely on the collation and interpretation of evidence both in favour and against claims made by framework developers. The analysis of some of the methodological challenges observed in the previous section highlight the difficulties in judging what kind of validation evidence is necessary and how much evidence collation and scrutiny is sufficient. In the final section, we ask three key questions that competence framework developers need to ask when considering framework validation, and offer a practical template to support the consideration of initial and ongoing validation approaches at the design stage of competence frameworks.

How much evidence should validations of competence frameworks rely upon?

The quantity of evidence and the rigour of the validation methodology will be dependent on the resources available. The type of evidence required is influenced by views on the degree of rigour that is appropriate given the contexts of use and the framework claims. If a competence framework is making ambitious claims about causal relationships or high stakes decisions (e.g., educational funding) then more evidence might be required compared to frameworks that make more modest or directly verifiable claims (Kane, 2013).

For example, Sastre-Fullana et al. (2017) claim that their competence framework is “useful for application in healthcare policy programmes for APN [advanced nursing] competency assessment in Spain” (p. 1). The ambitiousness of this claim is in part determined by the interpretation of the phrase “application in healthcare policy programmes”. A more modest claim could concern the successful embedding of the framework into policy decisions and into advanced nursing training or assessment. This utility claim may require a relatively small amount of resource to evaluate, such as desk-based analysis of how the framework has been used in informing policy documentation. A more ambitious claim, on the other hand, would relate to whether the competence framework, if considered as an intervention, resulted in improved medical outcomes, or career progress for patients and nurses respectively. Validation of this claim would require an extensive set of empirical studies and analyses, which may include both the collection of empirical data and the analysis of secondary data. In other words, an extended research programme for its validation would be required.

What validation evidence should we rely upon?

If, as argued earlier, validity is an interpretive judgement as to the degree to which the claims regarding the use or uses inferred from a competence framework are credible, then it is important for the framework developer (together with other stakeholders) to identify and judge contextually relevant *kinds* of evidence and analysis that can be employed. Moreover, any source of evidence or analysis that helps to establish a case for or against the overarching claim should be considered a legitimate source (Newton, 2017). Importantly, it is essential within the remit of a validation exercise to prioritise the kinds of evidence that are most powerful in directly evaluating the claim or claims made by the competence framework. This prioritisation will allow the resources available for validation to be used as effectively as possible.

How can we support the evaluation of the validity of competence frameworks at the design stage?

We have found that there is often a mismatch between the claims made by competence frameworks and the methods used to validate them. It is therefore important that at the design stage there is a consideration of the claims that might emerge from a competence framework and subsequently how evidence can be collected.

In Table 2, we suggest a template of questions for competence framework developers to consider in determining the range of potential claims to be made concerning their framework, in addition to understanding competence framework users and contexts. This is intended to be a template checklist for achieving an effective validation of competence frameworks at the design stage.

Table 2: A practical template for evaluating the validity of competence frameworks at the design stage.

<p><i>Initial consultation and planning of the competence framework</i></p> <p>Have the relevant practitioners/stakeholders been consulted for contextual information? What is the identified need or demand for the framework? What is/are the purpose(s) of the framework? Have the intended outcomes or uses for an intended framework been stated? Do the claimed purpose(s) of the framework relate to proposed uses?</p> <p><i>Identifying the construct(s) the competence framework is designed to measure</i></p> <p>Have the key competencies been identified? How have the key competencies been identified? Have the relations between the key competencies been established? To what extent are the construct definitions and rationale clear and explicit i.e., understandable to key stakeholders? How well aligned is the framework purpose to the relevant competencies?</p> <p><i>Constructing an appropriate approach to validate the competence framework</i></p> <p>Has an appropriate validation approach been identified? Does the validation approach need to be adapted to the appropriate context? Has a validation argument been constructed for each declared framework claim? Does the argument seek to support (or rebut) the stated competency claim? Does the argument draw on a variety of relevant sources to test the framework claims? Have the sources of evidence been prioritised based on consultation with practitioners and stakeholders? Has a decision been taken on how to prioritise framework claims in terms of their ambition? Does the validation approach allow for each source of evidence to be critically evaluated, rather than taken for granted? Does the argument allow each of the framework claims to be tested to evaluate its overall strength? Have criteria for evaluating the strength of the validity argument been identified in advance? Will the final decision on the validity of the competence framework involve relevant practitioners/stakeholders?</p> <p><i>Post-validation activity</i></p> <p>Has consideration been given to short-term, mid-term, and long-term impact monitoring studies of the efficacy of initial validation?</p>

Conclusion – establishing credibility of the claims of competence frameworks

This article, which builds on a methodological approach for constructing a competence framework (Child & Shaw, 2019), argues that validity is an overall judgement as to the degree to which the claims regarding the use or uses inferred from a competence framework are credible. Any validation endeavour will be “a professional exercise, involving insight, judgement and understanding” (Newton, 2017, p. 6).

Competence frameworks are meaningful in relation to the context(s) in which they are intended to function. A judgement of validity will be contextually driven, and evidence relevant to one context may not be relevant to another. This raises the difficult issue of *who is making the judgement* (e.g., the framework developer(s), groups of other stakeholders who may be users of the framework, or those who may be directly affected by its outcomes) and by what process or means.

Credibility is best achieved by identifying the most appropriate audience at the outset of validation, by acknowledging and taking seriously their concerns and values, and providing them with acceptable evaluative evidence (Newton, 2017). Developers of competence frameworks tend to base their design decisions or validation methods on what is meaningful to them. This has led to significant variability in validation approaches even within the same claim category (Batt et al., 2019). But validation is a “social decision procedure” (House, 1980, p. 249) and there is, therefore, a requirement to engage different competence framework users in social discourse. Affected stakeholders need to be active participants in the practices of the wider validation community by directly engaging in and contributing to validation practice. While this may pose a potential threat to the framework developer in the sense that they might be professionally challenged (Konrad, 1999), the increased transparency will ultimately be beneficial for the competence framework.

Finally, a validation study should not be taken as a “one off” event; contexts change, purposes evolve, as do stakeholder concerns. On-going validation of competence frameworks acts as a defence against contextual changes over time. Continuing validation, however, implies a shared and sustained mutual relationship between all relevant, affected parties. The validity argument for a competence framework does not need to be watertight. It does, however, need to be based on assumptions that are credible to those with a stake in demonstrating its valid use.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Baczynska, A. K., Rowinski, T., & Cybis, N. (2016). Proposed core competencies and empirical validation procedure in competency modelling: Confirmation and classification. *Frontiers in Psychology, 7*, 1–13. doi.org/10.3389/fpsyg.2016.00273

Batt, A. M., Tavares, W., & Williams, B. (2019). The development of competency frameworks in healthcare professions: a scoping review. *Advances in Health Sciences Education, 25*, 913–987. doi.org/10.1007/s10459-019-09946-w

Bok, H. G. J., Jaarsma, A. D. C., Teunissen, P. W., van der Vleuten, C. P. M., & van Beukelen, P. (2011). Development and validation of a competency framework for veterinarians. *Journal of Veterinary Medical Education, 38*(3), 262–269. doi.org/10.3138/jvme.38.3.262

Chartered Institute of Professional Development. (2021, December 15). [Competence and competency frameworks](#).

Chartered Insurance Institute. (2015, August). [The insurance competency framework](#).

Child, S. F. J., & Shaw, S. D. (2019). A purpose-led approach towards the development of competency frameworks. *Journal of Further and Higher Education, 44*(8). doi.org/10.1080/0309877X.2019.1669773

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer & H. I. Braun (Eds.), *Test Validity* (pp. 3–17). Lawrence Erlbaum.

European Commission Education & Culture. (2007). [Key competencies for lifelong learning – A European framework](#). European Communities.

European Commission Education & Culture. (2018). [Commission staff working document. Proposal for a council recommendation on key competencies for lifelong learning](#).

European Parliament and Council of the European Union. (2006). Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning. *Official Journal of the European Union, L 394/10*, 30 December 2006.

Hoskins, B., & Fredriksson, U. (2008). [Learning to learn: What is it and can it be measured?](#)

House, E. R. (1980). *Evaluating with Validity*. Sage.

House, E. R. (2014). Origins of the ideas in evaluating with validity. *New Directions for Evaluation, 142*, 9–15. doi.org/10.1002/ev.20081

Hyland, T. (1994). *Competence, Education and NVQs: Dissenting Perspectives*. Cassell.

- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. doi.org/10.1111/jedm.12000
- Konrad, J. (1999, July 9–11). The assessment and verification of National Vocational Qualifications [NVQs]: a European quality perspective. [Paper presentation]. International Lifelong Learning Conference, University College Worcester.
- Ling, S., Watson, A., & Gehrs, M. (2017). Developing an addictions nursing competency framework within a Canadian context. *Journal of Addictions Nursing*, 28(3), 110–116. doi.org/10.1097/JAN.0000000000000173
- Miranda, F. B. G., Mazzo, A., & Pereira-Junior, G. A. (2018). Construction and validation of competency frameworks for the training of nurses in emergencies. *Revisita Latino-Americana de Enfermagem*, 26. doi.org/10.1590/1518-8345.2631-3061
- Mulder, M., Weigel, T., & Collins, K. (2007). The concept of competence in the development of vocational education and training in selected EU member states: A critical analysis. *Journal of Vocational Education and Training* 59(1), 67–88. doi.org/10.1080/13636820601145630
- Muniz-Rodriguez, L., Alonso, P., Rodriguez-Muniz, L. J., & Valcke, M. (2017). Developing and validating a competence framework for secondary mathematics student teachers through a Delphi method. *Journal of Education for Teaching*, 43(4), 383–399. doi.org/10.1080/02607476.2017.1296539
- Newton, P. E. (2017, October). *An approach to understanding validation arguments*. Ofqual/17/6293, pp. 1–77.
- Patterson, F., Tavabie, A., Denney, M., Kerrin, M., Ashworth, V., Koczwara, A., & MacLeod, S. (2013). A new competency model for general practice: Implications for selection, training and careers. *British Journal of Medical Practice*, 63(10), e331–e338. doi.org/10.3399/bjgp13X667196
- Priestley, M., & Sinnema, C. (2014). Downgraded curriculum? An analysis of knowledge in new curricula in Scotland and New Zealand. *Curriculum Journal*, 25(1), 50–75. doi.org/10.1080/09585176.2013.872047
- Rissi, J. J., & Gelmon, S. B. (2014). Development, implementation, and assessment of a competency model for a graduate public affairs program in health administration. *Journal of Public Affairs Education*, 20(3), 335–352. doi.org/10.1080/15236803.2014.12001792
- Sastre-Fullana, P., Morales-Asencio, J. M., Sese-Abad, A., Bennasar-Veny, M., Fernandez-Dominguez, J. C., & De Pedro-Gomez, J. (2017). Advanced practice nursing competency assessment instrument (APNCAI): Clinimetric validation. *British Medical Journal Open*, 7, e-13659. doi.org/10.1136/bmjopen-2016-013659
- Soh, T. M. T., Osman, K., & Arsad, N. W. (2012). M-21CSI: A validated 21st century skills instrument for secondary science students. *Asian Social Science*, 8(6), 38–44. doi.org/10.5539/ass.v8n16p38
- Vitello, S., Greateorex, J., & Shaw, S. D. (2021). *What is competence? A shared interpretation of competence to support teaching, learning and assessment*.

Teachers' and students' views of access arrangements in high stakes examinations

Carmen Vidal Rodeiro (Research Division) and Sylwia Macinska (English Language Learning and Assessment)

Introduction

High stakes assessments pose challenges to some students' ability to demonstrate their knowledge and skills. These challenges may stem from assessment features not related to the constructs being measured. In such cases, students' performance may be affected by the access to the assessment, which can obscure the knowledge and the understanding of the content being assessed and be a threat to the validity and fairness of the assessment. To address this, many countries have introduced access arrangements (also known as test accommodations) to support the needs of students struggling with standard assessment procedures.

Access arrangements are pre-exam arrangements that help students with specific needs (e.g., special educational needs, disabilities, temporary injuries) to access the assessment and demonstrate their knowledge and skills by removing unnecessary barriers without changing the assessment demand or reducing its validity. For example, students who cannot concentrate for extended periods or fatigue easily may be awarded extra time to complete their assessments.

Evidence confirming the need for an arrangement (e.g., scores from psychometric assessments for the candidate; samples of the candidate's handwritten work; report from a medical professional outlining how a student's disability or illness is a barrier to the assessment) needs to be acquired by the centres. The evidence of need will vary depending on the special educational needs or disability of the students and on the access arrangement(s) being applied for (see Cambridge Assessment International Education (2022) for more details). Centres must, therefore, undertake the necessary steps to gather the evidence of need and demonstrate that a requested arrangement represents as much as possible the normal way of working for a student. The principle of the arrangement to align with the students' normal way of working aims to ensure that students are not introduced to an unknown procedure or technology during the assessment.¹

Evaluating access to access arrangements and how effective the arrangements are is important to ensure that the diverse learning needs of students are addressed and that the performance outcomes are a true reflection of students'

¹ Although students with temporary injuries (e.g., a broken arm a week before the exam) can request access arrangements, it is not expected that the approved arrangements in these situations align with the students' normal way of working. Access arrangements due to temporary injuries are out of the scope in this research.

knowledge and skills (Sireci et al., 2003). The provision and effectiveness of access arrangements can be evaluated from multiple angles. In particular, such evaluation should look at: (1) the types of students who are granted access arrangements, (2) the results of assessments when the arrangements are used, and (3) the students' and teachers' perceptions of how well the arrangements are working.

A large body of research has already examined the uptake of access arrangements (overall and broken down by students' characteristics) and their impact on performance. However, evidence regarding students' and teachers' perceptions of how well access arrangements work is rarely gathered. The aim of the research described in this article was, therefore, to gather stakeholders' views on access arrangements. In particular, the focus was on teachers' and students' understanding of the current provision of access arrangements and their perspectives on the strengths and weaknesses of access arrangements (including usefulness, fairness, consistency, implementation, and perceived effectiveness of use).

The research focused on access arrangements provided by Cambridge Assessment International Education (henceforth "Cambridge International"). Cambridge International is a large provider of international education programmes and qualifications for 5 to 19 year olds. They work with more than 10 000 centres in 160 countries around the world and offer pre-exam arrangements for all their qualifications (see Cambridge International, 2022).

In the following, we provide a summary of existing research evaluating access arrangements from the three different angles mentioned above.

Types of students granted access arrangements

There is a large body of research looking at the types of students who are granted access arrangements and the equality of students' access to such arrangements. The research has pointed out that the identification of special educational needs is not a scientific process and that there might not be an unbiased route to the provision of access arrangements. In fact, many studies have found that there is a relationship between the uptake of access arrangements and students' characteristics such as gender, type of school attended, attainment, socio-economic background, ethnicity and being identified as having special educational needs or disabilities (Fuchs et al., 2000; Lerner, 2004; Lindsay et al., 2006; Ofqual, 2020; Twist et al., 2006; Vidal Rodeiro, 2021; Yull, 2015). For example, some of the studies have shown that a disproportionately low number of access arrangements are awarded to students of minority ethnic groups, students who receive free or reduced-price school meals, or have low attainment at school. More recent research (Hutchison, 2021) has also shown a negative effect of attending school in a local authority with high levels of disadvantage; this made students less likely to be identified with special educational needs or disabilities than children of similar backgrounds in more affluent areas.

Impact of access arrangements on performance

The majority of the research looking at the effectiveness of access arrangements is based on experimental studies, often suffering from methodological limitations acknowledged by the authors (e.g., Duncan & Purcell, 2019; Gregg & Nelson, 2012; Liu et al., 2019). Additionally, as most of these studies are conducted in the United States, their results cannot be easily extrapolated to the context of high stakes assessments in other countries.

Furthermore, the practice of providing access arrangements is not without controversy. It is not clear to what extent access arrangements work as intended (i.e., create equity and level the playing field for candidates with disabilities and learning difficulties). Some argue that access arrangements may potentially lead to an unfair advantage for some students, rather than simply levelling the playing field (Elliott & Marquart, 2004; Zuriff, 2000). If that were the case, the test scores of the students with access arrangements would be inflated, which would have a detrimental effect on the validity of the assessment. Others, however, claim that access arrangements provide gains for students with special educational needs and disabilities, but do not seem to unduly advantage them (Cohen et al., 2005; Sireci, 2008). An overview of studies evaluating the effectiveness of some of the most common access arrangements (e.g., extra time; reading assistance; writing assistance; word processor) is given in Vidal Rodeiro and Macinska (2022).

Students' and teachers' perceptions of access arrangements

Research looking at students' and teachers' perceptions of access arrangements is scarce.

In the context of examinations in England, Woods (2007) highlighted that the lack of data relating to the perspectives of students with special educational needs or disabilities, and the perspectives of their parents/carers and teachers, is a particular obstacle to the effective evaluation of access arrangements.

Hipkiss and Robertson (2016) and Woods et al. (2018) have argued that user feedback on provision of access arrangements is particularly important to develop effective arrangements and highlighted the lack of students' perspective within the process of identifying students' needs. In particular, Woods et al. (2018) recommended continuous collaboration between awarding bodies and centres to ensure “*enhanced shared understanding of the purpose, place and limitations of access arrangements*”.

Lovett and Leja (2013), who reviewed empirical literature on students' perceptions of access arrangements in the United States, also discussed the importance of students' feedback on the usefulness of access arrangements. They mentioned that, in particular, if students are provided access arrangements that they do not believe to be helpful, they might not want to use them and that students' feedback is needed to determine how well the arrangements are working.

Regarding teachers' perceptions, Meadows (2012) carried out some work to measure teachers' attitudes towards the use of access arrangements for students with special educational needs in public schools in the United States. The data

collected in the study showed varying attitudes among teachers, differing by their position at the school (regular or special education teacher), the level of education taught, and the teacher's education and experience. For example, the research showed that special education teachers had a more positive attitude towards the use of arrangements than regular teachers and there was also a more positive attitude by teachers with lower education levels towards access arrangements. Teachers with positive attitudes tended to use access arrangements correctly and effectively to help improve student learning in the classroom and to improve student performance on examinations.

Data and methods

Data for this research was gathered via an online survey questionnaire. The survey included a mixture of closed and open-ended questions covering the following themes:

- awareness of access arrangements
- resources to provide access arrangements
- students' views on access arrangements
- overall views on access and inclusion.

We sent the survey to centres in eight countries: Indonesia, Italy, Malawi, Malaysia, Myanmar, Oman, South Africa, and Switzerland. In these countries there was a particular interest to find out about the use, implementation of, and views on access arrangements. All centres offering Cambridge International qualifications in each of the eight countries were invited to take part in the survey.

We carried out descriptive analyses for each question. Responses to open-ended questions were coded and analysed in an attempt to bring together recurring themes.

Findings

There were 258 responses to the questionnaire out of 587 invitations sent, resulting in a participation rate of 44 per cent. A mixture of Cambridge co-ordinators,² exams officers, senior management and teachers took part in the research. Their roles in the centres suggested that they were likely to be relatively well informed and able to respond to the survey questions.

Table 1 below shows that around three-quarters of the centres that responded to the questionnaire were in Indonesia (24 per cent), Italy (28 per cent) and Malaysia (26 per cent). The participation rate was highest in Italy, where over half of the invited centres (51 per cent) took part in the research, and lowest in Oman and Malawi, where 27 per cent and 30 per cent of the centres, respectively, responded to the survey.

² The Cambridge co-ordinator is the member of staff that is in charge of communications between the centre and Cambridge International. They are usually familiar with Cambridge procedures and carry out the administrative work within the centre relating to Cambridge qualifications.

Table 1: Participating centres, by country.

Country	Number of centres responded	Number of centres invited	Per cent centres (of responses)	Per cent centres (of invitations)
Indonesia	62	133	24.0	46.6
Italy	73	142	28.3	51.4
Malawi	6	20	2.3	30.0
Malaysia	66	160	25.6	41.3
Myanmar	4	9	1.6	44.4
Oman	8	30	3.1	26.7
South Africa	32	77	12.4	41.6
Switzerland	7	16	2.7	43.8
<i>Total</i>	<i>258</i>	<i>587</i>		<i>44.0</i>

Awareness of access arrangements

The first section of the survey investigated the respondents' awareness of access arrangements and the most common access arrangements used by centres offering Cambridge International qualifications.

Table 2 shows the responses to the question "Are there provisions in the education system of your country for access arrangements in examinations?". Just below 60 per cent of the respondents said that they were aware of such provision, but almost 30 per cent did not know. Only 13 per cent of the respondents mentioned that there was no provision for access arrangements in the education system in their country. Table 2 also shows the answer to the question broken down by country. The countries with the highest percentages of respondents being aware of provisions for access arrangements were South Africa, Italy and Malaysia. The highest lack of awareness was in Indonesia and Myanmar.

Table 2: Are there provisions in the education system of your country for access arrangements in examinations?

	All (N=258)	Indonesia (N=62)	Italy (N=73)	Malawi (N=6)	Malaysia (N=66)	Myanmar (N=4)	Oman (N=8)	South Africa (N=32)	Switzerland (N=7)
Yes	58.1	30.6	69.9	33.3	65.2	0.0	62.5	81.3	57.1
No	13.2	17.7	12.3	33.3	13.6	25.0	12.5	3.1	0.0
Don't know	28.7	51.6	17.8	33.3	21.2	75.0	25.0	15.6	42.9

Figure 1 shows the access arrangements available to request (in the centres that were aware of access arrangements provisions).

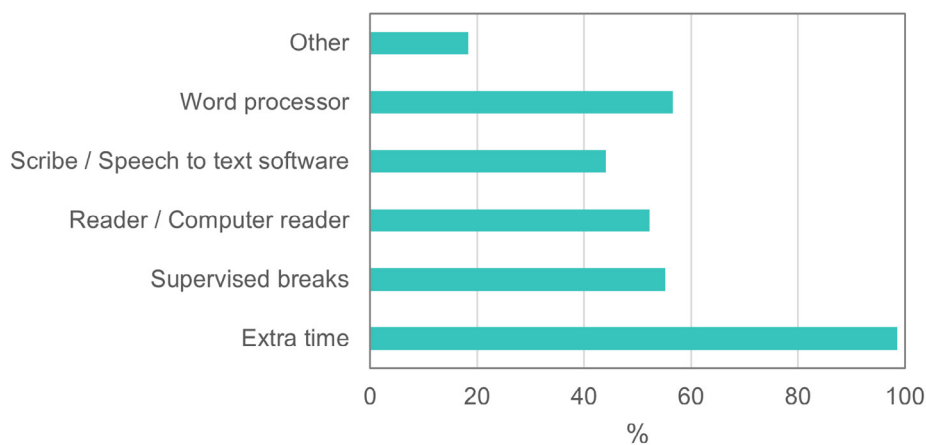


Figure 1: What access arrangements are available to request for your students for any examinations? (N = 136).

Extra time was available in 98.5 per cent of the centres. Using word processors, readers, or having supervised breaks was available in more than half of the centres. Just below 20 per cent of the respondents said that other access arrangements were available to request in their centres. These included coloured overlays, separate invigilation, colour naming and prompters.

The remaining questions in this section of the survey related to the provision of access arrangements for Cambridge International examinations. For details on all the different access arrangements (e.g., what they are and how to use them), see the *Cambridge Handbook for centres* (Cambridge International, 2022). Note that not all respondents answered all questions.

When asked the question “Do you know that Cambridge International offers access arrangements for their examinations?”, 93 per cent of the respondents to the question (135 out of 145) gave a positive response. Among those who were aware of Cambridge International’s offer, 75 per cent (N = 101) had applied for access arrangements for their students at some point. The majority of these centres had less than 5 per cent of students with access arrangements, although 11 per cent had between 6 per cent and 10 per cent of students with arrangements. Very few centres had higher percentages (only four centres reported having more than 10 per cent of their students with access arrangements).

Centres that never applied for access arrangements for Cambridge International examinations (N = 34) were asked the reasons for that. All respondents to this question (27 out of 34) selected as a reason for never having applied for access arrangements for Cambridge International examinations “None of the students required access arrangements”. It is encouraging that they did not select any of the other available options (e.g., the school does not have the resources to provide access arrangements; the school did not know how to apply for access arrangements; the school is not able to provide the required evidence of the students’ need for the access arrangements; the school lacks confidence to make judgements about students’ needs).

Resources to provide access arrangements

The questions in this section of the survey were only presented to those participants who knew that Cambridge International offered access arrangements for their examinations ($N = 135$). They were related to the availability of the centre resources to either request or provide access arrangements.

The first question asked participants if their centres had the appropriate resources to provide the access arrangements required by their students. Figure 2 shows that the vast majority of the centres did not have resourcing issues: 94 per cent of the respondents agreed or strongly agreed with the statement “My school has appropriate resources to provide the required access arrangements”. Only eight respondents disagreed with the statement. These respondents were in centres in Malaysia, Oman and South Africa. There were no reported resourcing issues in centres in Indonesia, Italy, Malawi and Switzerland. No participants from Myanmar provided an answer to this question.

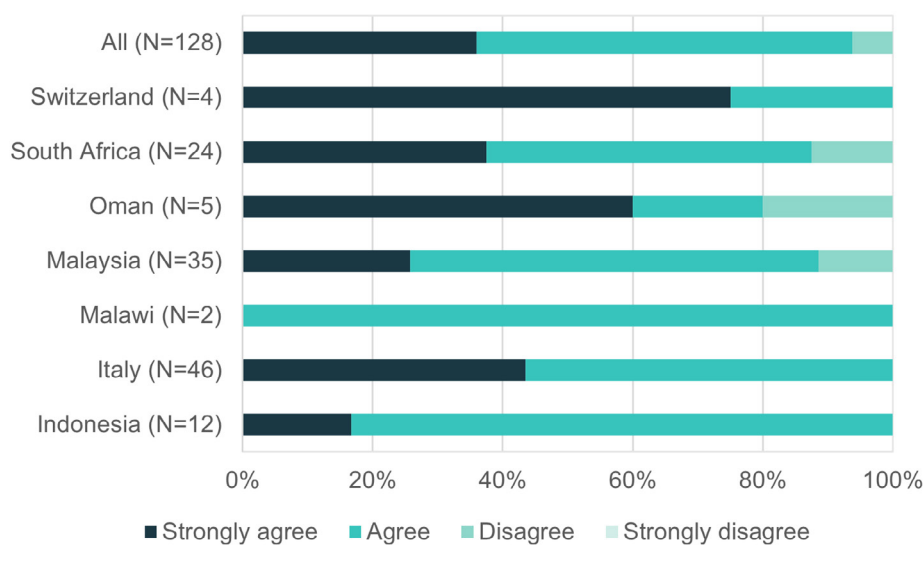


Figure 2: My school has appropriate resources to provide the required access arrangements.

Only those respondents who did not agree or strongly agree with the above statement were asked about the specific resourcing constraints with regard to providing access arrangements. Their answers included money, lack of staff, lack of physical space and technology issues as the main resourcing constraints.

Figure 3 shows that the vast majority of the centres did not experience resourcing problems when trying to provide Cambridge International with the required evidence to apply for access arrangements. 96 per cent of the respondents agreed or strongly agreed with the statement “My school has appropriate resources to provide the required evidence to apply for access arrangements”. Only five respondents disagreed with it. These respondents were in centres in Italy and Malaysia. There were no reported resourcing issues in centres in any of the other countries. As previously, no participants from Myanmar provided an answer to this question.

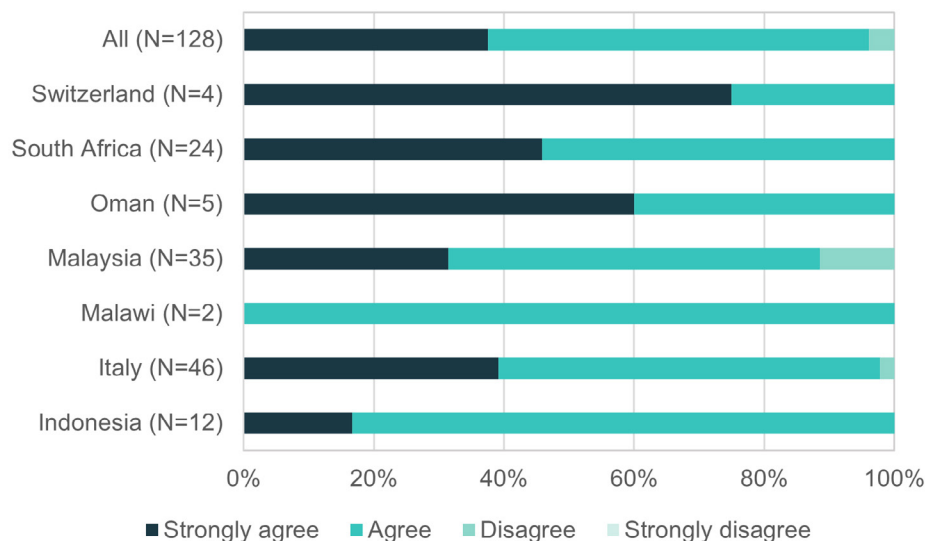


Figure 3: My school has appropriate resources to provide the required evidence to apply for access arrangements.

Students' views on access arrangements

Questions in this section were asked only to those participants who knew that Cambridge International offered access arrangements for their examinations ($N = 135$). They related to students' views on access arrangements, with participants asked to rate their agreement with several statements using a four-point scale from strongly agree to strongly disagree. To avoid data protection and consent issues in different countries, teachers (instead of students) were asked to answer the questions based on their observations and/or feedback from their students. No participants from Myanmar provided answers to questions in this section.

Overall, the majority of the responses confirmed that the alignment principle (that is, students only used the access arrangements in their examinations if that was their normal way of working in the classroom) was met. However, there were 20 centres where that was not the case (see Figure 4). Respondents also confirmed that, when access arrangements are requested, students use them in their examinations. Only a small number of centres in Indonesia and Italy disagreed with the statement "Students use the access arrangements that have been requested for them and approved for them".

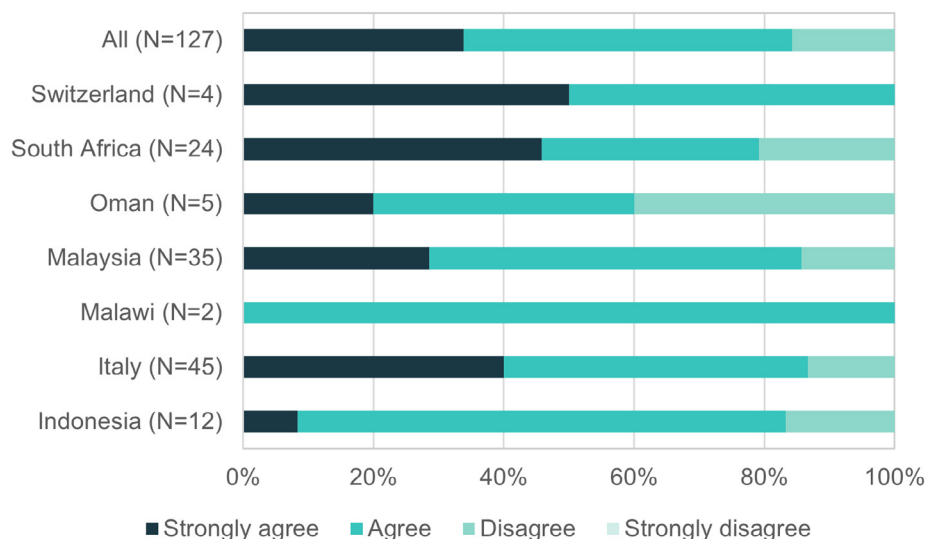


Figure 4: Students only use access arrangements in their examinations if that is their normal way of working in the classroom.

Teachers were next asked if students awarded access arrangements for their examinations find them useful. Their responses, summarised in Figure 5, were very reassuring. All but one respondent provided positive replies (61 per cent agreed and 38 per cent strongly agreed with the statement “Students awarded access arrangements find them useful”).

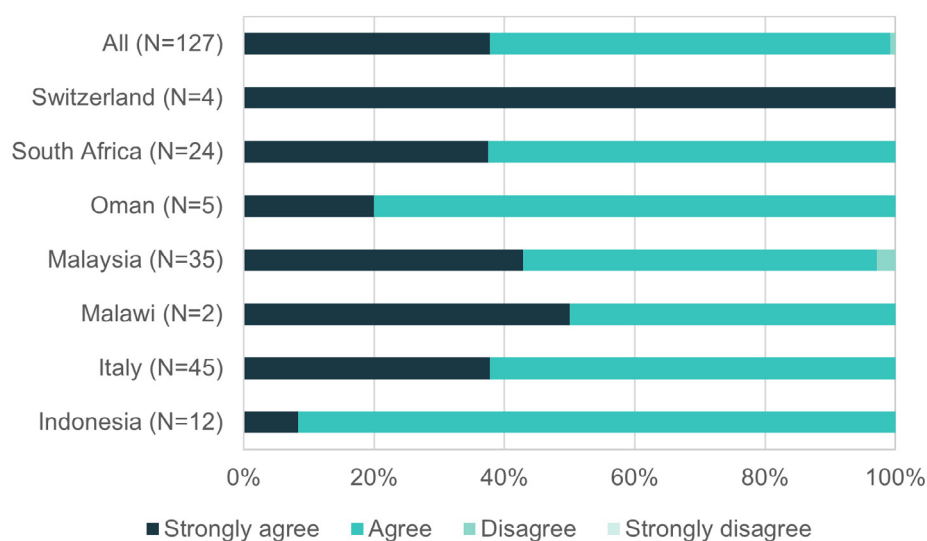


Figure 5: Students awarded access arrangements find them useful.

The majority of respondents, based on students’ feedback or on their own observations, did not think that students awarded access arrangements feel ashamed or embarrassed because they need assistance in their exams (Figure 6): 66 per cent disagreed with the statement “Students awarded access arrangements feel ashamed or embarrassed that they need assistance with exams”, and 17 per cent strongly disagreed. However, 22 respondents, that is 17 per cent of the respondents, agreed with the statement. The percentage was higher in Italy and Malaysia than in the other participating countries in the survey.

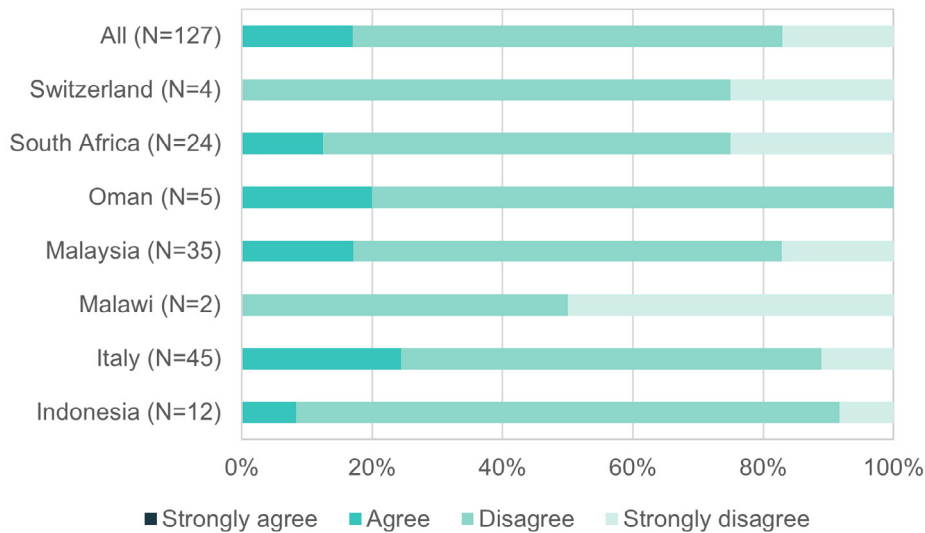


Figure 6: Students awarded access arrangements feel ashamed or embarrassed that they need assistance with exams.

Finally, Figure 7 shows that the majority of respondents to the questionnaire disagreed with the statement “Students without access arrangements regard such arrangements as unfair” (70 per cent disagreed and 22 per cent strongly disagreed). There were 10 centres that reported the opposite view. Such centres were in Indonesia, Italy, Malaysia, South Africa and Switzerland.

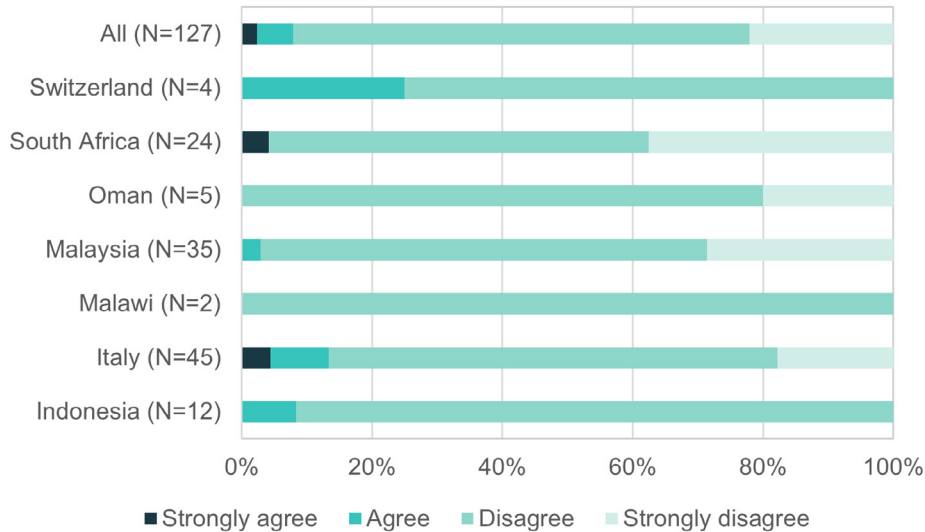


Figure 7: Students without access arrangements regard such arrangements as unfair.

To better understand the responses to the questions in this section of the survey questionnaire and to gather any further viewpoints, participants were asked if there was anything that they would like to tell us about their students’ views on access arrangements. There were 18 participants who left some comments.

Several respondents reported that students understood why some have access arrangements and others do not and that students are usually supportive of their peers. It was mentioned that the centre culture and ongoing practices of providing support through the year via different ways of working helped encourage this. This understanding from other students was also helped where the local system allowed access arrangements, as this normalised the concept. It was noted, however, that some cultures stigmatised learning difficulties making parents reluctant to have their child assessed, and that some students felt embarrassed and either refused access arrangements or had to be persuaded to accept them.

A couple of the further comments related to the use of extra time in examinations. Participants mentioned that students with dyslexia did not use (or need) the extra time awarded to them and that, in some cases, students awarded extra time do not use it unless encouraged by the teachers.

Overall views on access and inclusion

The final section of the survey asked all participants about their views on access and inclusion more generally, whether or not their students used access arrangements for their Cambridge International examinations. All participants were asked to rate, from strongly agree to strongly disagree, several statements and were given the opportunity to explain their answers or provide comments in a free-text question at the end.

Figure 8 shows the agreement of the respondents with the statement “Access arrangements are a fair means for helping students with disabilities and/or special needs”. Only one centre, in South Africa, disagreed with the statement. Overall, a higher percentage of centres strongly agreed (53 per cent) than agreed (47 per cent) with the statement.

Just below 50 per cent of the respondents did not agree with the statement “Students who need access arrangements in exams should be taught in special education schools” (see Figure 9). A further 27 per cent strongly disagreed. However, this overall pattern was not seen in all countries. For example, in Indonesia and Malaysia high percentages of participants (50 per cent and 32 per cent, respectively) agreed with the statement above. In contrast, in Italy, only 4 per cent of the participants did.

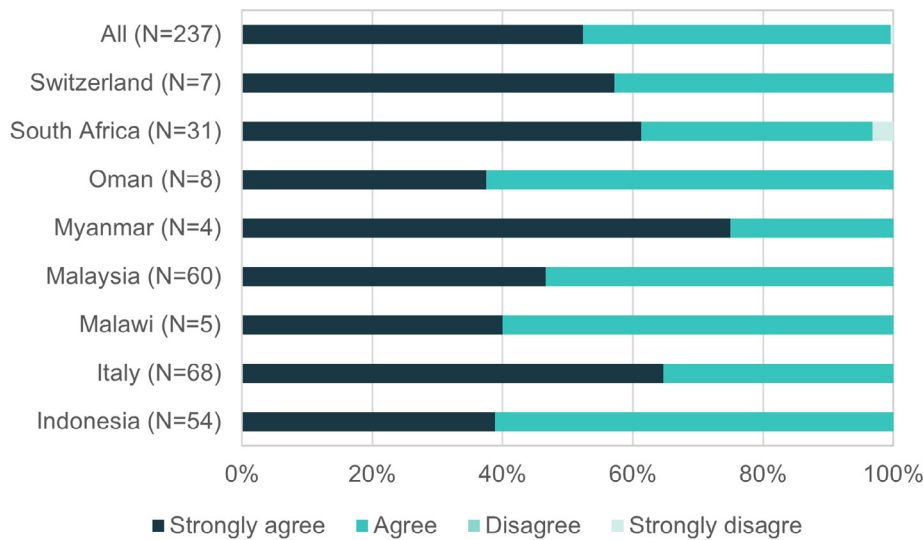


Figure 8: Access arrangements are a fair means for helping students with disabilities and/or special needs.

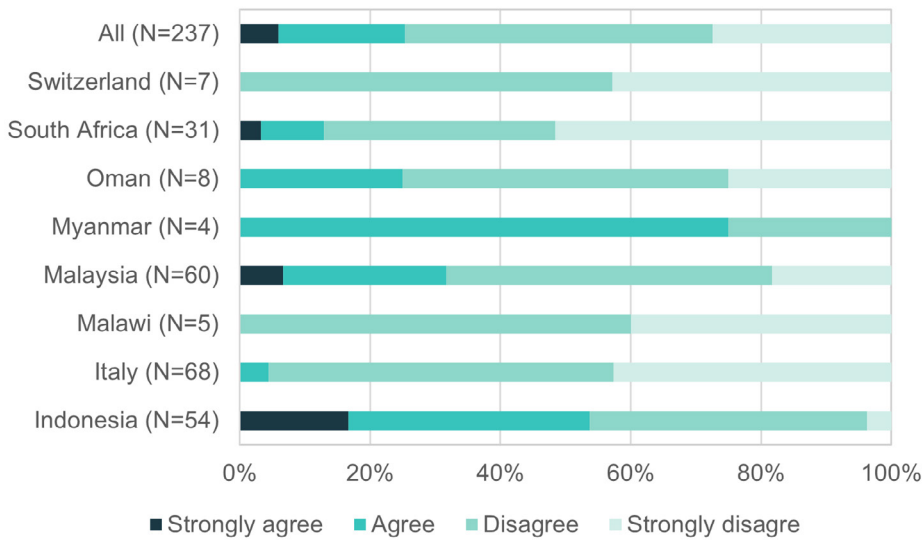


Figure 9: Students who need access arrangements in exams should be taught in special education schools.

Regarding access arrangements giving students with disabilities and/or special needs an unfair advantage, Figure 10 shows that, in general, participants did not think that was the case. Over 50 per cent of the participants disagreed with the statement “Access arrangements in exams give students with disabilities and/or special needs an unfair advantage”, with a further 34 per cent strongly disagreeing. There was a small minority of respondents who believed that access arrangements provide an unfair advantage. These respondents were mainly in Indonesia and Malaysia.

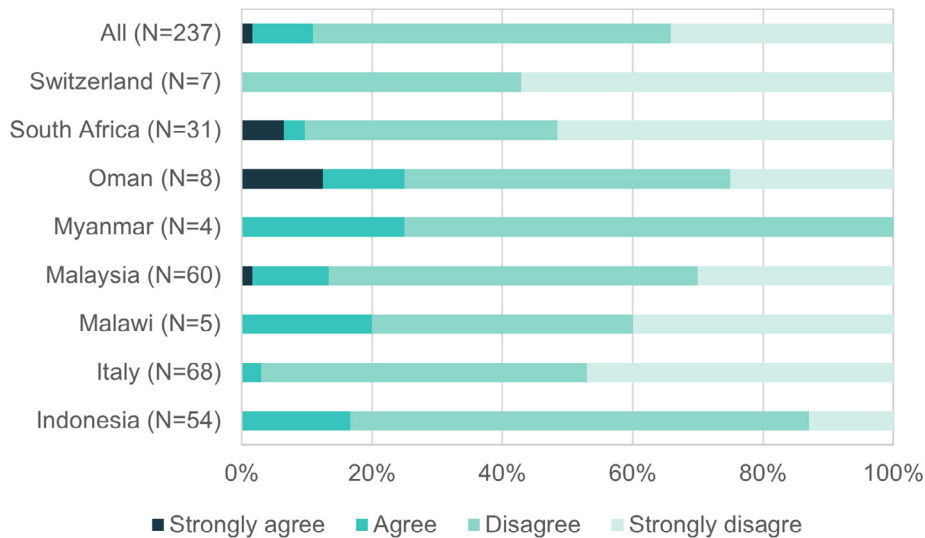


Figure 10: Access arrangements in exams give students with disabilities and/or special needs an unfair advantage.

The vast majority of the respondents thought that access arrangements make a difference in the education of students with special needs and/or disabilities and that they help provide an accurate picture of students’ abilities and knowledge. Italy and Malaysia were the countries with the highest numbers of respondents disagreeing with the above two statements.

Over 70 per cent of the respondents to the survey were not in agreement with the statement “Some students who have access arrangements in place do not really need them” (Figure 11). However, it is concerning that almost 30 per cent of the respondents (67 participants, mainly in Indonesia, Malaysia and Italy) thought that students granted access arrangements for their examinations do not need them.

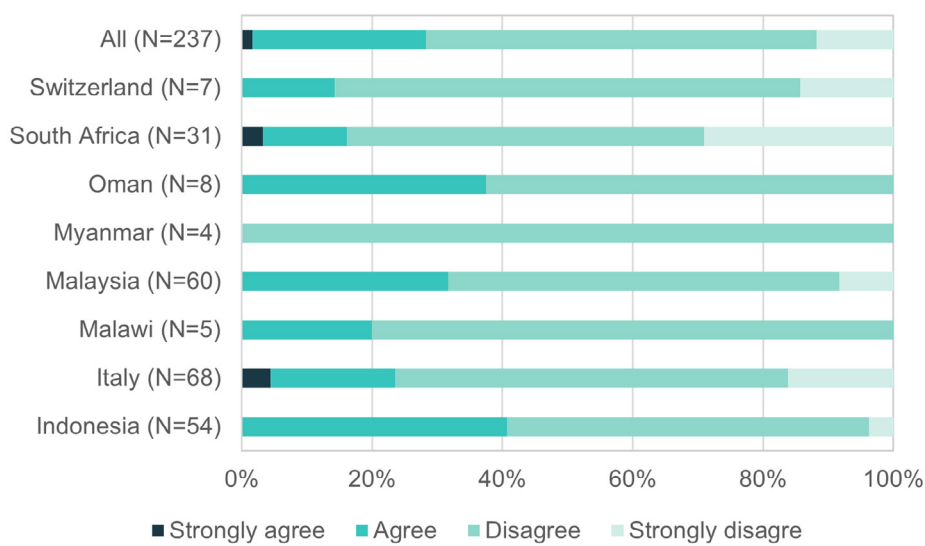


Figure 11: Some students who have access arrangements in place do not really need them.

Finally, we asked participants if they wanted to add any further comments about their views on access arrangements. There were 33 participants who responded to this question. There was a wide range of comments, which have been grouped into “positives”, “issues” and “suggestions”.

Positive points included that access arrangements allowed students with difficulties to show what they can do and that the centres are generally happy with access arrangements provision.

Issues included cost hindering professional assessments (e.g., to gather evidence of need), not understanding the guidance to request or deliver access arrangements, and special educational needs teachers reportedly being too generous. Some of the respondents also mentioned that providing access arrangements can be hard if the centre only needs them rarely and that special educational needs are a taboo topic with parents/society.

Respondents made suggestions relating to additional guidance (e.g., when arrangements are necessary; how to request arrangements) and exams officer training. Some further points from participants suggested a cautious view of access arrangement use (e.g., that it should only be given where really needed, that they were not “a must”). Three participants commented that students needing access arrangements should be in special needs schools.

Discussion and conclusions

Using a survey questionnaire, the present study reported on the views of 258 centres in eight countries around the world regarding awareness and views of access arrangements for students with special educational needs taking Cambridge International examinations.

Despite some variation within and between countries, the levels of awareness and provision of access arrangements found in this research are quite reassuring. As Griffiths and Woods (2010) mentioned “*awareness and availability of access arrangements may also support flexibility of teaching and learning opportunities for students experiencing special educational needs*”. This is important as inclusion of students with disabilities or special educational needs into general education settings means that they would receive access to the curriculum and assessments through the use of access arrangements.

Lack of resourcing to provide access arrangements in Cambridge International centres usually related to staffing or physical space needs to implement/deliver the access arrangements, to the cost of gathering the evidence of need, and to technology. Limited resourcing might cause centres to have thresholds for eligibility for access arrangements. For example, Woods et al. (2018) reported that, in the English context, there were numerous indications that resource constraints were inhibiting the process and the identification of students with special needs and therefore not having the appropriate provision of access arrangements. However, an effective access arrangements system should take as its starting point the individual student, not resourcing issues.

For arrangements to be beneficial, students need to be familiar with them. In fact, lack of practice in the use of the access arrangements could make students reluctant to use them during the examinations (Woods, 2007). This research confirmed that the access arrangements for Cambridge International examinations reflected the student's normal way of working in the classroom.

Previous research has shown (e.g., Bolt et al., 2011; Finn, 1998; Lewandowski et al., 2014; Sharoni & Vogel, 2007) that access arrangements are a positive feature of the services that awarding bodies provide to support students with special educational needs or disabilities and that access arrangements make students feel more comfortable and relaxed when taking the tests. Indeed, Lovett and Leja (2013) reported that, by removing access barriers to the tests (e.g., time limits or settings with distractions), access arrangements make students' testing experiences more positive, and Woods et al. (2010) and Elliott and Marquart (2004) mentioned that the use of access arrangements reduced students' experience of exam anxiety, thus leading to improved performance. Other research, has shown, however, that although arrangements such as extra time or rest breaks would benefit students, other access arrangements were not seen as having a particularly positive effect (Lewandowski et al., 2014) or could be distracting (Woods, 2007).

The majority of survey respondents, based on students' feedback or on their own observations, reported that students awarded access arrangements found them useful and did not feel ashamed or embarrassed by their need for assistance in their exams. Moreover, students without arrangements did not think that arrangements provided an unfair advantage. This contrasts with findings in previous research where students regarded access arrangements as "cheating" or felt embarrassed to be seen using them (e.g., Woods, 2007; Woods et al., 2010).

Teachers' attitudes and beliefs have a powerful influence on how successfully inclusive education practices (e.g., the use of access arrangements in exams) are implemented. Positive attitudes towards inclusion are, according to previous research, among the strongest predictors of the success of providing equal opportunities and access for all students (e.g., Alghazo & Gaad, 2004; Forlin et al., 2008; Forlin et al., 2011). The final section of the survey showed that views on access and inclusion were, overall, positive. The majority of the respondents did not think that students who need access arrangements in exams should be taught in special education schools. This resonates with recent policy changes, which have led to the integration of students with special educational needs or disabilities within mainstream schools so that they receive the same education and opportunities as their peers. Meadows (2012) reported that many schools have adapted to meet the distinctive needs of the individual students. As a result, students with disabilities have been moved from separate, special education classrooms, into general education classrooms where they receive access to the general education curriculum through the use of access arrangements.

Recently, in education systems around the world, there have been moves towards making assessments as inclusive as possible, rather than improving "access" through reasonable adjustments and there has been some research looking at alternatives to some of the current access arrangements. For

example, Lewandowski et al. (2014) suggested that increasing the time allowed in assessments will increase accessibility beyond the provision of specific access arrangements. “Universal test design” approaches that would be fair and valid for all students should also be considered. Examples of the requirements for universally designed assessments are: accessible, non-biased test items; simple, clear and intuitive instructions and procedures; maximum readability and comprehensibility. Future research could investigate further alternatives to the use of access arrangements that provide all students with the opportunity of equal access to the assessment and of displaying their full knowledge and skills.

In conclusion, teachers’ views on access arrangements should be gathered more regularly. This would provide evidence for timely and effective evaluation of the provision, the administration and the impact of access arrangements. Students’ (and parents’) views are also important and, although more difficult to gather due to access and consent issues should also be gathered as frequently as possible.

Acknowledgements

We would like to thank our colleague Vicki Crisp for her advice and support on the analysis of the survey qualitative data.

References

- Alghazo, E. M., & Gaad, E. E. N. (2004). General education teachers in the United Arab Emirates and their acceptance of the inclusion of students with disabilities. *British Journal of Special Education*, 31(2), 94–99. doi.org/10.1111/j.0952-3383.2004.00335.x
- Bolt, S. E., Decker, D. M., Lloyd, M., & Morlock, L. (2011). Students' perceptions of accommodations in high school and college. *Career Development for Exceptional Individuals*, 34, 165–175. doi.org/10.1177/0885728811415098
- Cambridge Assessment International Education. (2022). *Cambridge Handbook 2022. Regulations and guidance for administering Cambridge exams*.
- Cohen, A. S., Gregg, N., & Deng, M. (2005). The role of extended time and item content on a high-stakes mathematics test. *Learning Disabilities Research and Practice*, 20(4), 225–233. doi.org/10.1111/j.1540-5826.2005.00138.x
- Duncan, H., & Purcell, C. (2019). Consensus or contradiction? A review of the current research into the impact of granting extra time in exams to students with specific learning difficulties. *Journal of Further and Higher Education*, 44(4), 439–453. doi.org/10.1080/0309877X.2019.1578341
- Elliott, S. N., & Marquart, A. M. (2004). Extended time as a testing accommodation: Its effects and perceived consequences. *Exceptional Children*, 70(3), 349–367. doi.org/10.1177/001440290407000306
- Finn, L. L. (1998). Students' perceptions of beneficial LD accommodations and services at the postsecondary level. *Journal of Postsecondary Education and Disability*, 13, 46–67.
- Forlin, C., Earle, C., Loreman, T., & Sharma, U. (2011). [The sentiments, attitudes, and concerns about Inclusive Education Revised \(SACIE-R\) Scale for measuring pre-service teachers' perceptions about inclusion](#). *Exceptionally Education International*, 21(3), 50–65.
- Forlin, C., Keen, M., & Barrett, E. (2008). The concerns of mainstream teachers: Coping with inclusivity in an Australian context. *International Journal of Disability, Development and Education*, 55(3), 251–264. doi.org/10.1080/10349120802268396
- Fuchs, L. S., Fuchs, D., Eaton, S.B., Hamlett, C., Binkley, E., & Crouch, R. (2000). Using objective data sources to enhance teacher judgments about test accommodations. *Exceptional Children*, 67(1), 67–81. doi.org/10.1177/001440290006700105
- Gregg, N., & Nelson, J. M. (2012). Meta-analysis on the effectiveness of extra time as a test accommodation for transitioning adolescents with learning disabilities: More questions than answers. *Journal of Learning Disabilities*, 45(2), 128–138. doi.org/10.1177/0022219409355484
- Griffiths, D., & Woods, K. (2010). School processes in providing reading support in GCSE examinations. *British Journal of Special Education*, 37(4), 180–189. doi.org/10.1111/j.1467-8578.2010.00478.x

Hipkiss, A., & Robertson, C. (2016). *The impact of access arrangements organisation on Special Educational Needs Co-ordinators (SENCOs)*. University of Manchester.

Hutchison, J. (2021). *Identifying pupils with special educational needs and disabilities*. Education Policy Institute.

Lerner, C. (2004). "Accommodations" for the learning disabled: A level playing field or affirmative action for elites? *Vanderbilt Law Review*, 57(3), 1043–1124.

Lewandowski, L., Lambert, T. L., Lovett, B. J., Panahon, C. J., & Sytsma, M. R. (2014). College students' preferences for test accommodations. *Canadian Journal of School Psychology*, 29(2), 116–126. doi.org/10.1177/0829573514522116

Lindsay, G., Pather, S., & Strand, S. (2006). *Special Education Needs and ethnicity: Issues of over- and under-representation*. Department for Education and Skills Publications.

Liu, K. K., Thurlow, M. L., Press, A. M., & Dosedel, M. J. (2019). *A review of the literature on computerized speech-to-text accommodations*. NCEO Report 414. University of Minnesota, National Center on Educational Outcomes.

Lovett, B. J., & Leja, A. M. (2013). Students' perceptions of testing accommodations: What we know, what we need to know, and why it matters. *Journal of Applied School Psychology*, 29(1), 72–89. doi.org/10.1080/15377903.2013.751477

Meadows, M. P. (2012). *Teacher attitudes toward the use of accommodations in the classroom and on standardized tests*. The University of Southern Mississippi.

Ofqual. (2020). *Access Arrangements for GCSE, AS and A level: 2019 to 2020 academic year*. Office of Qualifications and Examinations Regulation.

Sharoni, V., & Vogel, G. (2007). Entrance test admissions, admission and enrolment of students with learning disabilities in teacher training colleges in Israel. *Assessment and Evaluation in Higher Education*, 32, 255–270.

Sireci, S. G. (2008). *Validity issues in accommodating reading tests*. *Journal of Educators and Education*, 23, 81–110.

Sireci, S. G., Li, S., & Scarpati, S. (2003). *The Effects of Test Accommodation on Test Performance: A Review of the Literature*. Center for Educational Assessment Research Report no 485. University of Massachusetts.

Twist, L., Donahue, B., Lewis, K., & Keogh, N. (2006). *Report on the impact of additional time in National Curriculum tests*. National Foundation for Educational Research.

Vidal Rodeiro, C. L. (2021). *Equality of students' access to Access Arrangements and their impact on students' performance*. Cambridge Assessment Research Report. Cambridge Assessment.

Vidal Rodeiro, C. L., & Macinska, S. (2022). Equal opportunity or unfair advantage? The impact of test accommodations on performance in high-stakes assessments. *Assessment in Education: Principles, Policy and Practice*, 29(4), 462–481. doi.org/10.1080/0969594X.2022.2121680

Woods, K. (2007). Access to the General Certificate of Secondary Education (GCSE) examinations for students with Special Educational Needs: What is best practice? *British Journal of Special Education*, 34(2), 89–95. doi.org/10.1111/j.1467-8578.2007.00461.x

Woods, K., James, A., & Hipkiss, A. (2018). Best practice in access arrangements made for England's General Certificates of Secondary Education (GCSEs): Where are we 10 years on? *British Journal of Special Education*, 45(3), 236–255. doi.org/10.1111/1467-8578.12221

Woods, K., Parkinson, G., & Lewis, S. (2010). Investigating access to educational assessment for students with disabilities. *School Psychology International*, 31(1), 21–41. doi.org/10.1177/0143034310341622

Yull, A. (2015). [The impact of race and socioeconomic status on access to accommodations in post-secondary education](#). *Journal of Gender, Social Policy & the Law*, 23(2), 353–392.

Zuriff, G. E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education*, 13(1), 99–117. doi.org/10.1207/s15324818ame1301_5

Who controls what and how? A comparison of regulation and autonomy in the UK nations’ education systems

Pia Kreijkes and Martin Johnson (Research Division)

Introduction

Prior to devolution in 1999 the education systems of England, Scotland, Wales and Northern Ireland differed, and many aspects have gradually diverged even further since (Sibieta & Jerrim, 2021). Changes in education policy have also affected the distribution of regulation and autonomy across different levels of their education systems. This has led to variations in the ways that institutions and actors exert control over aspects of school governance and curriculum in the different nations, leading to complex and differing relationships between local autonomy and central control. A crucial part of this relationship is the “middle tier”, which operates in the space between the central government and individual schools. It includes a range of actors, which differ across nations, such as local authorities (LA), school clusters and education partnerships. There is sometimes ambiguity about what the middle tier includes. For example, Woods et al. (2021b) suggest that school inspection bodies may or may not reside in the middle tier in different nations. To resolve this, it is perhaps useful to define the middle tier according to its function rather than through the agencies that populate it. The middle tier is a space where agencies mediate the process of government policy through to its enactment in schools. As such, it provides a link between policy and practice. This article examines and reflects upon such complex relationships and the distribution of power and autonomy within the educational systems of the nations of the United Kingdom.

Autonomy can be defined as “the condition in which a person or an entity, such as a country or organisation, can exercise self-rule or self-governance” (Woods et al., 2020, p. 118). It implies a considerable degree of freedom in making decisions and determining one’s conduct. Regulation, in contrast, describes how rules or directives from others determine one’s conduct. Regulation from higher levels directly affects the autonomy of actors at lower levels. Consequently, the autonomy of individual schools and teachers depends on how much freedom is *granted* to them by the powers to which they are held accountable. Therefore, it has been argued, one always needs to question how genuine any apparent state of autonomy really is (Woods et al., 2020). Even if schools may have a relatively high degree of rhetorical autonomy, features of the wider school system, including

those of the middle tier, can impinge on that autonomy and effectively prevent schools from enacting it fully. Such features can reduce the effective autonomy by exerting power explicitly, such as through inspection and accountability measures, or through implicit influences, for instance the pressure to conform to the educational values endorsed by those in positions of power (Woods et al., 2020).

This article first provides a brief overview of how the education systems of England, Scotland, Wales and Northern Ireland have developed in relation to each other. It then compares the education policy-making styles of the nations. This is followed by a descriptive comparison of the distribution of power in four broad areas of the system: 1) school governance, 2) curriculum, 3) assessment, and 4) school improvement and accountability. The comparison sets the scene for discussing the interplay of autonomy and regulation within the education systems as a whole.

It is important to caveat this paper by recognising that our analysis is accurate at the time of writing (late 2022), and we have ignored policy changes that fall within the academic year 2022/23. We acknowledge that systems are continually developing, for example we are aware that there are plans for substantial changes to the regulatory bodies in Scotland as three new national organisations will be created (e.g., see <https://www.gov.scot/news/new-national-education-bodies/>). Rather than considering this a limitation of the paper, we argue that the timeliness of our analysis adds to the body of literature that describes the changing relations between the constituent countries of the UK. This point is important because there has been a long tradition of convergence and divergence between the education policies of the UK nations since the early 19th century, and this analysis helps to take stock of this process at the present time.

Brief overview of how UK nations' education systems developed in relation to each other

Figure 1 presents a simplified version of how the national governance structures of the four nations have evolved in relation to each other in the modern period. During the period from the 16th century to the end of the 20th century, the four nations have moved through processes of convergence, incorporation, devolution and divergence, and these shifts have inevitably influenced many areas of policy, including education.

The late 19th and early 20th centuries saw a gradual period of divergence in policy-making across the four nations. There was a devolution of education powers to the Scottish Office department of the UK Government in 1872, and in 1922 there was the creation of a devolved government in Northern Ireland which had responsibility for education in the province. Although there were some convergent pressures brought about by the imposition of Direct Rule¹ in the province in 1972, there was a very significant divergent shift in the later 1990s when a series of referenda were held in Scotland, Wales, and Northern Ireland. These referenda led to the establishment of a Scottish Government, a Welsh

¹ Direct Rule is the administration of Northern Ireland directly by the UK Government. Apart from a period in 1974, it was in place for 27 years between 1972 and 1998 (Torrance, 2022).

Government, and a Northern Ireland Executive. These devolved institutions held a range of powers that differed from each other, but they all had responsibilities for developing their own education policies. For example, since devolution each nation has undergone a curriculum review process, resulting in curricula that have varied to different degrees from the National Curriculum for England.

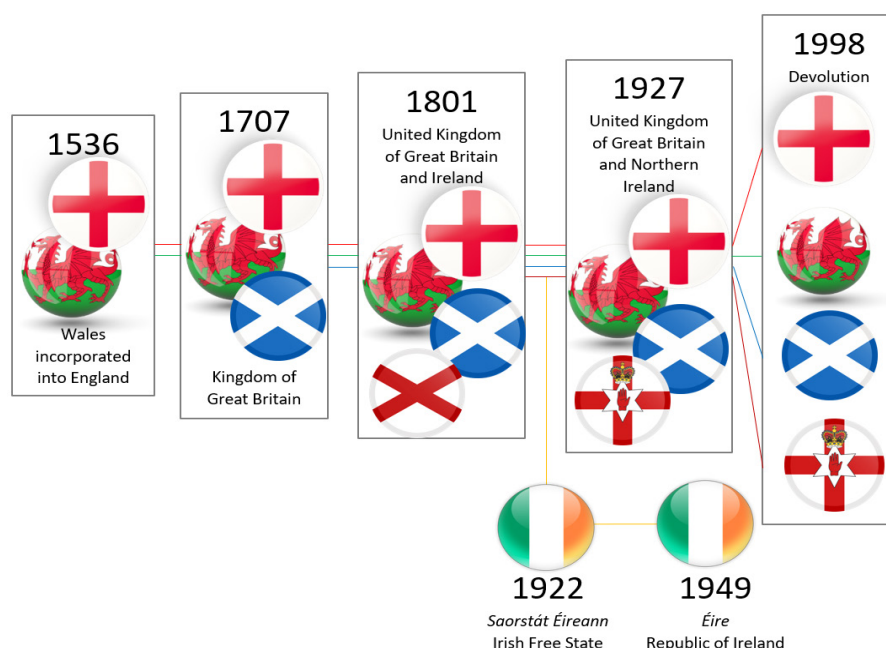


Figure 1: Four nations’ governance – historic links.

While policy development variances are evident across the four UK nations, it has been observed that there are a number of common structural and social features that mean that the nations can be perceived as being more similar than different – especially when compared with systems in other non-UK nations (Croxford & Raffe, 2014; Machin et al., 2013; Paterson & Ianelli, 2007). Raffe et al. (1999) argue that comparing close neighbours’ education policies can elicit nuanced analyses. This is supported by Krause (2018) who notes that close case analysis benefits from reducing the number of variables that need to be considered, and which Arnott & Menter (2007) highlight, can help to make apparent any interdependencies between nations.

Policy-making styles: (de)centralisation, hierarchy and policy mediation

Policy has hierarchic and centralised potential as policy decision-making (including budgeting decisions) tends to coalesce around the upper levels of government in the UK (e.g., see Clark, 2012; Richards et al., 2022). The execution of such policy may rely on different structures to communicate it to those enacting it. These structures may be organised hierarchically, prone to centralised direction, or through networks that distribute the responsibilities for policy enactment via mediating processes. One can thus conceive education systems to have three levels or tiers: 1) central government (tier 1 or the macro level), where national policy-making happens, 2) the middle tier (tier 2 or the meso level), where policy is mediated from central government to schools, and 3) individual schools (tier 3 or the micro level), where school leaders and teachers enact policy.

Greany (2022) broadly considers “any aspect of statutory or non-statutory support and influence which operates between individual schools and central government” as part of the middle tier (p. 249). Given that the space between central government and individual schools is vast, it is perhaps not surprising that there are various ways in which education systems can employ the mediating tier. The areas where it can play a substantial role include finance (such as allocating funding), accountability (including support for improvement), access (ensuring provision for all children) and people (professional development, staffing) (Bubb et al., 2019).

Governance of education that involves a middle tier in policy mediation is likely to include a network of institutions, such as consortia of local authorities in Wales and Regional Improvement Collaboratives (RICs) in Scotland. Middle tier functions help to mediate policy implementation through bringing together existing elements of government-funded activity, such as school staff, university researchers and members of local authorities. The different roles of the multi-faceted middle tier within the four UK nations will be expanded upon throughout this article, and an overview of the various bodies belonging to the tiers is provided in the Appendix.

In contrast to networked governance processes, hierarchic policy enactment might reduce the role of the middle tier by capitalising on direct communication between the central government and those enacting policy (which in education would be school managers and leaders). Education policy in England may be seen to have hierarchic characteristics, relying to some extent on a flat hierarchy with a high degree of social regulation (Malin et al., 2020). Flat hierarchies can benefit from communicative clarity as messages between the executive and those enacting policies can be less susceptible to degradation. A negative consequence of flat hierarchies in education is that they might lead to a sense of isolation for some schools as there is little local-level policy mediation beyond a school, and this could lead to variations in practice (Teelken, 2000).

Scottish policy enactment exhibits a less hierarchic policy approach than England. It has been observed that Scotland has shifted from “a dominant culture of high social regulation, with its associated bureaucratic, managed organisations, to ... [a] culture with high levels of social cohesion manifested through partnership” (Chapman, 2019, p. 561). This networked approach aims to support a consensual process that is heavily populated with national organisations, professional bodies and interest groups (Grek & Ozga, 2009), where central government sets overall direction but leaves implementation to regional and local actors (Chapman, 2019).

Despite the appearance that networks distribute responsibilities for policy enactment across various partners, which may suggest a more cohesive approach, it is important to explore the way that power works in practice. It has been observed that multi-stakeholder partnership networks, which seem to display less bureaucratic authority, may exacerbate established hierarchic power relations rather than undermine them (Faul, 2016), an analysis that chimes with recent comment on Scottish policy enactment (Humes, 2020).

A brief comparison of key elements of the education systems across the UK

In this section we describe and compare key elements of the education systems of the UK nations: 1) school governance, 2) curriculum, 3) assessment, and 4) school improvement and accountability. Note that these areas overlap. In addition to some other sources, our description strongly draws on the excellent comparison of school institutions and policies by Sibieta and Jerrim (2021), and we refer interested readers to their report for more detail. The description lays the foundation for a discussion of how autonomy and regulation are distributed through the systems as a whole.

1) School governance

Governance structures, and the role of the middle tier within these, differ between as well as within nations depending on school type. This section examines who is responsible for the running of schools, including who has control over staffing, teacher pay and school spending. First, an overview of the school types is provided (Table 1).

Table 1: School types within the UK nations.

England ^a	Scotland ^b	Wales ^c	Northern Ireland ^d
<ul style="list-style-type: none"> • Community (LA maintained) • Foundation and Voluntary-controlled • Voluntary-aided • Academies • Free schools • Grammar • Faith schools • Private/independent 	<ul style="list-style-type: none"> • Public/Local (state funded) • Grant aided • Independent/private • Denominational • Gaelic-medium 	<ul style="list-style-type: none"> • Community • Voluntary-controlled • Voluntary-aided (often religious or faith schools) • Foundation • Welsh-medium 	<ul style="list-style-type: none"> • Controlled • Catholic maintained • Grant-maintained integrated • Voluntary grammar • Integrated • Independent • Irish-medium

Note. School types are not all mutually exclusive. For example, denominational schools in Scotland can be state-funded or independent, and grammar schools in England can be LA-maintained, foundation schools or academies. Descriptions of the type of schools for each nation can be found here:

a <https://www.gov.uk/types-of-school>

b <https://www.citizensadvice.org.uk/scotland/family/education/school-and-pre-school-education-s/types-of-school-s/>

c <https://law.gov.wales/schools-maintained-local-authorities>

d <https://www.education-ni.gov.uk/articles/information-school-types-northern-ireland>

There are some notable differences in school type attendance between nations (see Atkins et al., 2021). A higher percentage of pupils in Northern Ireland attend selective secondary schools (43 per cent) compared to England (5 per cent), and no pupils attend such schools in Wales and Scotland. Attendance of independent schools is lowest in Northern Ireland (0.2 per cent) and highest in England (7 per cent). In addition, the vast majority of pupils in Northern Ireland attend religiously affiliated schools (91 per cent), which is much lower in England (24 per cent), Wales (18 per cent) and Scotland (14 per cent). Beyond England, it is possible for pupils to attend schools where English is not the primary medium of instruction. In Wales, a

relatively high percentage of pupils attend such schools (23 per cent), but this is lower in Northern Ireland (2 per cent) and Scotland (1 per cent).

The middle tier plays a comparatively small role in school decisions in England compared to the rest of the UK. The regulation of schools by local authorities has reduced over time, and considerable autonomy over school staffing, pay and spending decisions has been extended to individual local authority-maintained schools (Sibieta & Jerrim, 2021). This trend has been exacerbated since 2010 when the Academies Act 2010 was introduced. Academies and free schools possess very high levels of autonomy as they can set their own admissions arrangements (subject to legislation) and deviate from national pay and conditions for staff. In addition, they are not obligated to follow the National Curriculum. In Wales, local authorities have a greater influence on school decisions compared to England, as they retain control over staffing and teacher pay. Governing boards also play a considerable role in staffing as well as admissions policies. All publicly funded schools in Wales must follow national pay and conditions. Importantly, schools in England and Wales have considerably more autonomy over their spending compared to Scotland and Northern Ireland, where middle tier organisations have more control.

In Scotland, local authorities and national government are largely in control of school decisions, such as staff recruitment and retention. Local authorities play a key role in financial decisions. Over time, the influence of school governors has diminished as school boards have been replaced by consultative parent councils and parent forums. In Northern Ireland, the Education Authority has much control over spending. Northern Ireland stands out because all schools are managed by boards of governors who, alongside the national government, have considerable control over school governance. In controlled schools, which are mostly attended by Protestants, the board of governors acts on behalf of the Education Authority. Catholic maintained schools are also funded by the Education Authority, but schools are managed by a board of governors while the Council for Catholic Maintained Schools is the employer.

2) Curriculum

A brief overview of curriculum organisation in the four nations is provided in Table 2. Of the four curricula, the National Curriculum for England stands out as the only one that can be characterised as ‘subject-based’, while the other three curricula focus on cross-cutting areas of learning and competencies (Atkins et al., 2021; Sibieta & Jerrim, 2021). England’s curriculum clearly sets out the minimum material that teachers need to cover at each stage of education (except for those at academies and free schools – although it is acknowledged that many academies still follow the National Curriculum (Roberts, 2021)). Teachers have some flexibility over when they introduce the content, and they can go beyond the minimum requirements, but the National Curriculum remains highly prescriptive (Sinnema et al., 2020). By contrast, the curricula of Scotland, Wales and Northern Ireland give schools and teachers considerable autonomy over content choice (Sibieta & Jerrim, 2021). For example, Scotland’s Curriculum for Excellence sets out expectations for learners’ experiences and outcomes at different educational

stages, but it does not prescribe specific content. Similarly, schools and teachers in Northern Ireland can set content that is appropriate to learners' interests and abilities.

Interestingly though, these differences in the level of prescription do not seem to be reflected in headteachers' perceptions of who is shaping the curriculum in practice. For a detailed comparison, we refer readers to Sibieta and Jerrim (2021), who analysed data from the Programme for International Student Assessment (PISA) 2006, 2009 and 2015. Teachers in Scotland were less frequently perceived to have considerable responsibility over course offerings and content compared to the other UK nations. In addition, the national government and local education authorities were perceived to play a larger role in shaping course offerings and content in Scotland compared to England, with Northern Ireland and Wales falling in between. This is at odds with the emphasis on teacher autonomy in the Scottish Curriculum for Excellence, which demonstrates the challenge of achieving a teacher-led curriculum when national and local agencies play a considerable role in the wider school governance (Sibieta & Jerrim, 2021). The reduced role of the national government and the middle tier in shaping course offerings and content in England might partly be explained by the prevalence of academies.

Table 2: Curriculum organisation.

England	Scotland	Wales	Northern Ireland
National Curriculum for England (2014)	Curriculum for Excellence (2010)	Curriculum for Wales (2022)	The Northern Ireland Curriculum/Statutory Curriculum at Key Stage 3 (2007)
Subject-based model: Organised around disciplines, such as English, Maths, History, Geography, Science and Physical Education	Cross-cutting areas of learning: Expressive arts; Health and wellbeing; Languages; Mathematics; Religious and Moral Education; Sciences; Social Studies; Technologies Aims to help young people become successful learners; confident individuals; responsible citizens; and effective contributors	Cross-cutting areas of learning: Expressive arts; Health and Well-being; Humanities; Languages, Literacy and Communication; Mathematics and Numeracy; Science and Technology Aims to help young people become ambitious, capable learners; enterprising, creative contributors; ethical, informed citizens; healthy, confident individuals	Cross-cutting areas of learning: Language and Literacy; Mathematics and Numeracy; Modern Languages; The Arts; Environment and Society; Science and Technology; Learning for Life and Work; Physical Education; Religious Education

3) Assessment

While there was a general reduction in external testing across the UK in the period from devolution to 2010, national assessments resurged from 2010 onwards (Sibieta & Jerrim, 2021). An overview of the major internal and external assessments that pupils take throughout primary and secondary school is provided in Table 3.

In contrast with the other UK nations, England has maintained and expanded the use of school performance league tables (Sibieta & Jerrim, 2021). The other nations abolished such league tables in the early 2000s (although Wales adopted a publicly available school categorisation system in 2011 which triggers different levels of external support). League tables expose schools in England to strong market pressures (Machin et al., 2013) which place schools in direct competition with each other and concentrate decision-making outside of the middle tier.

Table 3: Primary and secondary assessment.

England	Scotland	Wales	Northern Ireland
<ul style="list-style-type: none"> • Baseline check (age 5) • Phonics check (age 6) • SATs (Maths & English, internally marked, age 7) • SATs (Maths & English, age 11) 	<ul style="list-style-type: none"> • Scottish National Standardised Assessments: Literacy & Numeracy (P1, P4, P7) 	<ul style="list-style-type: none"> • National Literacy and Numeracy Tests (Y2–Y9) 	<ul style="list-style-type: none"> • Levels of Progression (Literacy, Numeracy and ICT in Y4, Y7) • Unofficial transfer tests to secondary school
<ul style="list-style-type: none"> • GCSE uses 9–1 system • GCSEs and A levels focus on linear end of course assessments 	<ul style="list-style-type: none"> • SNSAs: Literacy & Numeracy (S3) • National 1–5s, Highers, and Advanced Highers 	<ul style="list-style-type: none"> • National Literacy and Numeracy Tests (Y9) • GCSE uses A*–G system • Retained AS levels and a modular system of assessment for both GCSEs and A levels 	<ul style="list-style-type: none"> • Levels of Progression (Literacy, Numeracy and ICT in Y10) • GCSE uses A*–G system (including a C* grade) • Retained AS levels and a modular system of assessment for both GCSEs and A levels

Note. This table is based on Sibieta and Jerrim (2021). SATs = Standardised Assessment Tests.

Importantly, the fact that Scotland reduced external assessment as well as participation in large-scale international studies such as the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS) after 2007, makes it very difficult to evaluate the success of its curriculum (OECD, 2015). The Scottish government has also been criticised for the replacement of the Scottish Surveys of Literacy and Numeracy (SSLN), which ran from 2011 to 2016, as the lack of data hinders scrutiny of the education system (Education and Skills Committee, 2019). The role of assessments in school improvement and accountability measures is addressed in the next section.

Again drawing on PISA data, Sibieta and Jerrim (2021) found that Scotland stands out when it comes to headteachers' perceptions of what agencies and actors are responsible for establishing internal and external pupil assessment policies. Across nations, headteachers reported they themselves play the largest role, followed by teachers. Yet, in Scotland, the national and regional government were perceived to have substantially more responsibility compared to the other three nations, whereas governing boards were perceived to have substantially less responsibility. There thus seems to be less autonomy at the school level and more regulation at the national level and middle tier in Scotland.

4) School improvement and accountability

Assessments and league tables are an important factor in school improvement and accountability measures. According to Arnott and Menter (2007), the culture of performativity²

“... effectively ensures compliance within the system and enables government to ‘be accountable’ for its policies. So the combination of testing, league tables for schools, targets and target setting, key performance indicators, standards and inspection creates a discourse where comparison becomes simple and where ‘failure’ and ‘success’ can be identified very easily” (p. 255).

Schools in all four nations seem to use assessments to make comparisons with regional and national performance, to monitor the school’s progress from year to year, and to identify aspects of instruction or curriculum that could be improved (Sibieta & Jerrim, 2021). In England, schools are more likely to use assessment data to judge teachers’ effectiveness compared to the rest of the UK, in particular Scotland.

All four nations employ regional or national agencies as part of their school improvement and accountability systems. An overview of such bodies and the areas they inspect is provided in Table 4. While school inspection plays a key role across the UK, Sibieta and Jerrim (2021) note some interesting differences. In Scotland, national agencies rather than middle tier organisations appear to play a larger role in the accountability system. Education Scotland has joint responsibilities for school improvement, the curriculum and inspection, whereas inspection is in the hands of agencies that are separate from those responsible for school improvement and the curriculum in the other nations. The areas that are inspected are quite similar across nations, although Scotland stands out because they only examine between two and four of their 15 quality indicators at each inspection. Notably, inspections in England and Wales occur more frequently compared to those in Scotland and Northern Ireland.

Poor inspection results tend to result in formal notices as well as additional interventions and support (often provided by the middle tier) and follow-up inspections across the UK nations, which is described in detail by Sibieta and Jerrim (2021). For example, if schools in Wales require significant improvements or special measures, Estyn must inform the Minister for Education, and the school and local authority must submit action plans to address the problems. Schools that require urgent improvements enter a Formal Intervention Process involving external support, and the Regional Consortia play an important role in providing that support, which again highlights the importance of the middle tier.

² A culture of performativity in education is characterised by an emphasis on performance evaluation, quantifiable targets and comparisons (see Ball, 2003).

Table 4: School inspection.

England	Scotland	Wales	Northern Ireland
Ofsted	Education Scotland	Estyn	Education and Training Inspectorate
Areas include: <ul style="list-style-type: none"> • Quality of education • Behaviour and attitudes • Personal development • Leadership and management 	Areas include 2–4 of 15 quality indicators from 3 key themes: <ul style="list-style-type: none"> • Leadership & Management • Learning Provision • Successes & Achievement 	Areas include: <ul style="list-style-type: none"> • Standards • Wellbeing and attitudes to learning • Teaching and learning experiences • Care, support and guidance • Leadership and management 	Areas include: <ul style="list-style-type: none"> • Achievement and standards • Provision for learning • Leadership and management
<ul style="list-style-type: none"> • Schools are inspected about once every four years 	<ul style="list-style-type: none"> • About 10 per cent of schools were inspected in 2018–19 	<ul style="list-style-type: none"> • Legislation stipulates that all schools must be inspected at least once every seven years 	<ul style="list-style-type: none"> • About two-thirds of primary and post-primary schools were inspected between July 2016 and June 2018

Note. This table is largely based on Sibieta and Jerrim (2021).

In England, there seem to be more severe and immediate consequences for school governance compared to the rest of the UK. The Academies Act 2010 gave a statutory duty to the Department for Education to request all maintained schools who received an Ofsted “Inadequate” rating to convert to an academy (Atkins et al., 2021). This means that sponsors take over the school and appoint an independent board of governors. If academies or free schools receive such a rating, the Regional Schools Commissioner can implement various improvement measures such as transferring an academy to a new academy trust or sponsor (Sibieta & Jerrim, 2021). Multi-academy trusts (MATs) are responsible for the performance of schools belonging to their trust and are thus another important part of the accountability system that is located in the middle tier.

A recent report highlighted key similarities and differences between the four nations in the intended mechanism of school improvement at the school level (Munoz-Chereau & Ehren, 2021). In England, Wales and Northern Ireland, the provision of feedback through the inspectorate is thought to lead to improvement. In Scotland, Wales and Northern Ireland, the promotion of school self-evaluation is regarded as crucial. In England, self-evaluation seems to play a much smaller role in the inspection process. Lastly, in Scotland and Northern Ireland, inspection is thought to enhance professional dialogue and a culture of school self-reflection, which, in turn, is considered to lead to school improvement.

Discussion

After describing various individual aspects or elements of educational systems, it is important to consider the systems’ “ecology” as a whole. Similar functions are covered in each system, but the profile of where these are carried out differs. For example, the regulator of qualifications and examinations in Northern Ireland (Council for the Curriculum, Examinations and Assessment)

can be considered a tier 1 organisation as it has a closer relationship to policy-making than the Office of Qualifications and Examinations Regulation (Ofqual) in England, which performs the same regulatory function but resides notionally in tier 2. It is likely that a series of compensations exist, where a function (such as monitoring performance) is covered by an inspection agency in one nation but is covered by schools in another system. This means that trying to make sense of a system through describing its elements (e.g., school types) without considering relational links is of limited use.

This observation gives our discussion a cultural perspective. Some of the variances across national systems are less about the functions that are performed in the systems (as these share a high degree of commonality) but in the ways that the agencies that deliver these functions relate to each other. Woods et al. (2021a) highlight how education system reform in England and Northern Ireland reflects a managerialist style where state funding is largely passed directly to school leaders who are then tasked with educational improvement responsibilities. In Scotland and Wales there appears to be a greater culture of consensus and consultation to encourage policy implementation. Similarly, there are variations of policy emphasis when looking at relationships around the middle tier in the different nations. It has been observed that Scotland, Wales and Northern Ireland prioritise policies that support relationship building across schools and communities, while the focus in England is on enhancing relationships across MATs (Woods et al., 2021a).

None of the four UK nations just lets its schools “run wild” to act completely independent of governments’ intentions or aims for education, and each government exerts control in some way through a variety of agencies. Accordingly, although the extent of autonomy that is granted to individual schools may differ across countries, there has to be some kind of regulation to ensure that schools across the country provide good education to all pupils. Scotland and England present particularly interesting cases for exploring where power resides in the system. In Scotland, input regulation (e.g., prescribed content) is low and output regulation (e.g., inspection, evaluative use of achievement) is also relatively low whereas in England, both input and output regulation are high (Leat et al., 2013). Despite the low curriculum input regulation in Scotland, a greater percentage of Scottish headteachers report that national and local government have considerable control over course offerings and content compared to England (Sibieta & Jerrim, 2021). This apparent mismatch between policy intentions and headteachers’ perceptions indicates the presence of other factors that determine teachers’ perceived autonomy over the curriculum.

Sinnema et al. (2020) noted that there can be a conflict between the autonomy that is granted to teachers and other regulatory mechanisms that undermine their ability to enact such autonomy in practice. While there may be a tendency that input regulation is replaced by output regulation (Nieveen & Kuiper, 2012), this does not seem to be the case in Scotland, at least not in an intentional and explicit way. Instead, there are more implicit and perhaps unintended regulatory

pressures at play. First, little prescription, and as some argue a lack of clearly articulated learning progressions, in the curriculum (e.g., Drew, 2013; Priestley & Minty, 2010) encourage teachers to seek guidance elsewhere. Examining the Scottish curriculum, Smith (2019) found that content selection can increasingly depend on the demands of external assessments rather than educational priorities, a sentiment supported by recent empirical work (Ritchie et al., 2022). There is a greater perceived role of local and national government on assessment policies in Scotland compared to the other nations (Sibieta & Jerrim, 2021). As such, teachers' perceptions of diminished agency might be a recognition that the locus of assessment control has been ceded to the Scottish Qualifications Authority (SQA) and formal examination demands. In the light of reduced curriculum input regulation, the power of the SQA is relatively strong in shaping the enacted curriculum (Priestley & Drew, 2017).

A second, more fundamental issue regarding the enactment of autonomy is that merely granting autonomy to teachers is insufficient if the system does not support them in enacting it. The idea of teachers as curriculum developers and agents of change has been fundamental in the development of the Curriculum for Excellence in Scotland (see Priestley & Minty, 2013 for various examples). And yet, Priestley et al. (2012) described it as an irony that agency is considered something that can be "demanded" from teachers. Teachers are put in a situation in which they must take on more responsibility, but they are not provided the support to develop the skills required to fulfil it. This may be another reason why teachers turn to assessments to seek guidance and gain confidence in the way that they plan learning.

It is clear, but perhaps deserves further highlighting, that a high level of teacher autonomy over the curriculum is not inherently positive or negative. This has been noted by Sinnema et al. (2020), who stated that flexibility can be considered a burden as well as a gift. For example, increased autonomy can increase teachers' sense of control, commitment and satisfaction, as well as allow them to adapt the curriculum to local needs and interests. Yet, as discussed, it can also leave teachers with a lack of guidance and encourage them to orientate their teaching towards assessments. It also means that pupils' learning experiences are highly dependent on what individual schools and teachers regard as important, which can lead to a patchwork of content that lacks coherence and leads to high degrees of variability across the country.

England is also a very interesting case for examining the relationship between teacher autonomy, input regulation and output regulation. Both input regulation through curriculum prescription and outcome regulation through accountability measures are considered high (Leat et al., 2013). However, in 2022, about 39 per cent of primary schools and 79 per cent of secondary schools are academies (Plaister, 2022), which are exempt from following the National Curriculum, and thus do not experience this form of input regulation. Nevertheless, academies are still accountable to the Department for Education, being monitored by Regional Schools Commissioners and Ofsted. The conversion of local authority maintained schools into academies, which are outside of local authority control and National

Curriculum requirements, exemplifies that autonomy over school decisions has been a policy priority for at least two decades (Sibieta & Jerrim, 2021; Woods et al., 2020) despite high input and output regulation. The conversion into academies was strongly influenced by the idea of a “self-improving school system”, which is characterised by school-led improvement (Woods et al., 2020).

However, the reduction of local authority power has led to the creation of a new middle tier as more and more authority is transferred from individual schools to multi-academy trusts (MATs) (see Woods et al., 2020). MAT powers include “direction setting” (which focuses on school performance), holding the headteacher to account and ensuring financial probity (including setting staff pay). As such, they act as the governing body for groups of academies and have considerable control over pay, conditions, the curriculum and budgetary decisions (Sibieta & Jerrim, 2021). The majority of academies are part of a MAT (~75 per cent, Sibieta & Jerrim, 2021). Hence, this de-emphasises the idea of a self-improving school system and seems counter to the reasons that academies were created in the first place (Greany & Higham, 2018). In addition, there is now “a tighter level of prescription” from central government about how MATs operate as well as “a requirement for tight vertical accountability, both within MATs and between MATs and the government” (Greany & Higham, 2018, p. 86). After all then, it seems that despite political emphasis on school-led improvement and reducing local authority influence, much control still resides in the middle tier rather than in individual schools albeit in a different form (i.e., now in MATs rather than local authorities).

In conclusion, our analysis shows that the character of the middle tier differs across the four UK nations (e.g., England has less local authority involvement than the other nations). This raises a question about whether this variation reflects broader systematic characteristics of the nations, telling us something about the nature of the relationships between central and local government in the different nations, and the nations’ political cultures more generally. Our discussion also shows that the middle tier in each of the four nations varies in profile, but that they share some common functions that are key to mediating the way that policy links with schools. The middle tier, through its various agencies, has two key functions. These agencies increase system cohesion through trying to reduce variability across schools (e.g., through the MAT structure in England or the local or education authorities in Scotland, Wales and Northern Ireland). These agencies also seek to enhance efficient resource allocation, using regulatory mechanisms and data generation to highlight effective teaching and learning practices (e.g., through inspection and links to assessment outcomes). In this way, the middle tier has a direct and crucial role in empowering schools to reach their full potential, which supports the governments’ abilities to achieve their political goals at a distance (Rose & Miller, 2008, cited in Ozga & Lawn, 2014).

The middle tier can be perceived as both a support and a threat to teacher agency, and the line between these perceptions can be fine and blurred. The middle tier may seek to harness the agency of teachers in implementing policies, so that the responsibility for system development and improving standards

is shared across multiple stakeholders. From this perspective, the middle tier provides necessary support for teachers to enact their autonomy. From a different perspective, the middle tier can be considered a threat to teacher agency by assuming control over functions and decisions that could otherwise be in the hands of individual schools. Since the middle tier bridges the functions of government and schools/teachers, perceptions of whether the tier is a support or a threat to teacher agency may broadly depend on the nature of relationships within the education sector.

References

- Arnott, M., & Menter, I. (2007). The same but different? Post-devolution regulation and control in education in Scotland and England. *European Educational Research Journal*, 6(3), 250–265. doi.org/10.2304/eeerj.2007.6.3.250
- Atkins, G., Dalton, G., Phillips, A., & Stojanovic, A. (2021). *Devolved public services: The NHS, schools and social care in the four nations*. Institute for Government.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228. doi.org/10.1080/0268093022000043065
- Bubb, S., Crossley-Holland, J., Cordiner, J., Cousin, S., & Earley, P. (2019). *Understanding the Middle Tier Comparative Costs of Academy and LA-maintained School Systems*. Sara Bubb Associates. doi.org/10.13140/RG.2.2.23574.29761
- Chapman, C. (2019). From hierarchies to networks: Possibilities and pitfalls for educational reform of the middle tier. *Journal of Educational Administration*, 57(5), 554–570. doi.org/10.1108/JEA-12-2018-0222
- Clark, G. (2012). *Decentralisation: An assessment of progress*. HM Government.
- Croxford, L., & Raffe, D. (2014). Social class, ethnicity and access to higher education in the four countries of the UK: 1996–2010. *International Journal of Lifelong Education*, 33(1), 77–95. doi.org/10.1080/02601370.2013.873214
- Drew, V. (2013). *Curriculum for Excellence in the Secondary School*. In T. Bryce, W. Humes, D. Gillies, & A. Kennedy (Eds.), *Scottish Education: Referendum* (pp. 501–511). Edinburgh University Press.
- Education and Skills Committee. (2019). *Report on Scottish National Standardised Assessments* (p. 93). Scottish Parliament.
- Faul, M. V. (2016). Networks and power: Why networks are hierarchical not flat and what can be done about it. *Global Policy*, 7(2), 185–197. doi.org/10.1111/1758-5899.12270
- Greany, T. (2022). Place-based governance and leadership in decentralised school systems: Evidence from England. *Journal of Education Policy*, 37(2), 247–268. doi.org/10.1080/02680939.2020.1792554
- Greany, T., & Higham, R. (2018). *Hierarchy, Markets and Networks: Analysing the 'self-improving school-led system' agenda in England and the implications for schools*. UCL Institute of Education Press.
- Grek, S., & Ozga, J. (2009). Governing education through data: Scotland, England and the European education policy space. *British Educational Research Journal*, 36(6), 937–952. doi.org/10.1080/01411920903275865
- Humes, W. (2020). Re-shaping the policy landscape in Scottish education, 2016–20: The limitations of structural reform. *Scottish Educational Review*, 52(2), 89–111.
- Krause, M. (2018). How fields vary. *The British Journal of Sociology*, 69(1), 3–22. doi.org/10.1111/1468-4446.12258

Leat, D., Livingston, K., & Priestley, M. (2013). Curriculum deregulation in England and Scotland—Different directions of travel? In W. Kuiper & J. Berkvens (Eds.), *Balancing Curriculum Regulation and Freedom across Europe*. SLO.

Machin, S., McNally, S., & Wyness, G. (2013). Educational attainment across the UK nations: Performance, inequality and evidence. *Educational Research*, 55(2), 139–164. doi.org/10.1080/00131881.2013.801242

Malin, J. R., Brown, C., Ion, G., van Ackeren, I., Bremm, N., Luzmore, R., Flood, J., & Rind, G. M. (2020). World-wide barriers and enablers to achieving evidence-informed practice in education: What can be learnt from Spain, England, the United States, and Germany? *Humanities and Social Sciences Communications*, 7(1), 1. doi.org/10.1057/s41599-020-00587-8

Munoz-Chereau, B., & Ehren, M. (2021). *Inspection across the UK: how the four nations intend to contribute to school improvement*. Edge Foundation.

Nieveen, N., & Kuiper, W. (2012). Balancing curriculum freedom and regulation in the Netherlands. *European Educational Research Journal*, 11(3), 357–368. doi.org/10.2304/eeerj.2012.11.3.357

OECD. (2015). *Improving Schools in Scotland: An OECD Perspective*. Policy Advice and Implementation Division Directorate for Education and Skills – OECD.

Ozga, J., & Lawn, M. (2014). Inspectorates and politics: The trajectories of school inspection in England and Scotland. *Revue Française de Pédagogie*, 186, 11–21. doi.org/10.4000/rfp.4388

Paterson, L., & Ianelli, C. (2007). Social class and educational attainment: A comparative study of England, Wales, and Scotland. *Sociology of Education*, 80, 330–358. doi.org/10.1177/003804070708000403

Plaister, N. (2022, May 18). *The size of multi-academy trusts*. FFT Education Datalab.

Priestley, M., & Drew, V. (2017). Curriculum for Excellence: Making the transition from policy intention to classroom practice. *The Scottish Educational Journal*, 101(3), 20–21.

Priestley, M., & Minty, S. (2010). *Developing Curriculum for Excellence: Summary of findings from research undertaken in a Scottish local authority*. University of Stirling.

Priestley, M., & Minty, S. (2013). Curriculum for Excellence: ‘A brilliant idea, but...’. *Scottish Educational Review*, 45(1), 39–52.

Priestley, M., Robinson, S., & Biesta, G. J. J. (2012). Teacher agency, performativity and curriculum change: Reinventing the teacher in the Scottish Curriculum for Excellence? In B. Jeffrey & G. Troman (Eds.), *Performativity in UK education: Ethnographic cases of its effects, agency and reconstructions*. Ethnography & Education Publishing.

Raffe, D., Branned, K., Croxford, L., & Martin, C. (1999). Comparing England, Scotland, Wales and Northern Ireland: The case for “home internationals” in comparative research. *Comparative Education*, 35(1), 9–25.
doi.org/10.1080/03050069928044

Richards, D., Warner, S., Smith, M. J., & Coyle, D. (2022). Crisis and state transformation: Covid-19, levelling up and the UK’s incoherent state. *Cambridge Journal of Regions, Economy and Society*, rsac038. doi.org/10.1093/cjres/rsac038

Ritchie, M., Peace-Hughes, T., Barnett, C., Priestley, M., & Shapira, M. (2022, September). *Student experiences of the Scottish Curriculum for Excellence*. British Educational Research Association Annual Conference, University of Liverpool.

Roberts, N. (2021). *The school curriculum in England* (Briefing Paper No. 06798). House of Commons Library.

Rose, N., & Miller, P. (2008). *Governing the present: Administering economic, social and personal life*. Polity Press.

Sibieta, L., & Jerrim, J. (2021). *A comparison of school institutions and policies across the UK*. Education Policy Institute.

Sinnema, C., Nieveen, N., & Priestley, M. (2020). Successful futures, successful curriculum: What can Wales learn from international curriculum reforms? *The Curriculum Journal*, 31(2), 181–201. doi.org/10.1002/curj.17

Smith, J. (2019). Curriculum coherence and teachers’ decision-making in Scottish high school history syllabi. *The Curriculum Journal*, 30(4), 441–463.
doi.org/10.1080/09585176.2019.1647861

Teelken, C. (2000). Market forces in education, a comparative perspective in England and Scotland. *Scottish Educational Review*, 32(1), 21–32.
doi.org/10.1163/27730840-03201003

Torrance, D. (2022). *Northern Ireland: Direct Rule*. UK Parliament.

Welsh Government. (2017). *Education in Wales: Our national mission. Action plan 2017-21*. The Education Directorate.

Woods, P. A., Roberts, A., Jarvis, J., & Culshaw, S. (2020). England: Autonomy and regulation in the school system in England. In H. Ärlestig & O. Johansson (Eds.), *Educational Authorities and the Schools: Organisation and Impact in 20 States* (pp. 111–130). Springer International Publishing. doi.org/10.1007/978-3-030-38759-4_7

Woods, P. A., Torrance, D., Donnelly, C., Hamilton, T., Jones, K., & Potter, I. (2021a). Constructions and purposes of school leadership in the UK. *School Leadership & Management*, 41(1–2), 152–170. doi.org/10.1080/13632434.2020.1859999

Woods, P. A., Torrance, D., Donnelly, C., Hamilton, T., Jones, K., & Potter, I. (2021b). Diverging systems of educational leadership in the four nations of the United Kingdom? *School Leadership & Management*, 41(1–2), 1–5. doi.org/10.1080/13632434.2021.1887644

Appendix

Education tiers and agencies across the UK nations.

	England	Scotland	Wales	N Ireland
Macro level/ Tier 1: Policy-making	Department for Education (DfE) <ul style="list-style-type: none"> Regulates the school system through 18 agencies and public bodies (some examples below) Sets the curriculum 	Education Scotland (ES) <ul style="list-style-type: none"> A Scottish Government executive agency and directly accountable to Scottish Government ministers Supports quality and improvement in education, including professional development Oversees the implementation of the curriculum (set by the Scottish Government) Inspection of schools and other education services 	Department for Education and Skills (DfES) <ul style="list-style-type: none"> Responsible for education, training and children's services (curriculum is set by the Welsh Government) 	Department of Education (DE) <ul style="list-style-type: none"> Duty to promote education and ensure the effective implementation of education policy Delivers its functions through 11 Arm's Length Bodies (some examples below)
				Council for the Curriculum, Examinations and Assessment (CCEA) <ul style="list-style-type: none"> Non-departmental public body of the DE Sets and develops the curriculum
Meso level/ Tier 2 (Middle tier): Policy mediation	National Schools Commissioner and eight Regional Schools Commissioners <ul style="list-style-type: none"> Supported by board of headteachers Provide oversight and support to under-performing schools Heavily involved in approving conversion to Academy status and new Free Schools 		Four Regional Consortia of local authorities <ul style="list-style-type: none"> Co-ordinate school improvement support across local authorities Responsible for professional development Distribute various grants to schools 	Education Authority <ul style="list-style-type: none"> Non-departmental body sponsored by the DE Provision of education and youth services Funding authority for all schools Oversees provision of education services

	England	Scotland	Wales	N Ireland
	<p>150 local authorities</p> <ul style="list-style-type: none"> Each LA appoints a Director of Children's Services Provide support services and brokering support between schools 	<p>32 local education authorities</p> <ul style="list-style-type: none"> Statutory duty to ensure adequate and efficient provision of school education in their area Spending and accountability for educational funding Can propose changes to education provision following a formal consultation process 	<p>22 local authorities</p> <ul style="list-style-type: none"> Duty to promote high standards of education and fair access to education 	
	<p>Multi-Academy Trusts</p> <ul style="list-style-type: none"> Governance of academies belonging to the trust 			
	<p>The Office for Standards in Education, Children's Services and Skills (Ofsted)</p> <ul style="list-style-type: none"> Non-ministerial government department Inspects and regulates schools and other education services Reports to Parliament but powers and duties reflect central government policies 		<p>Estyn</p> <ul style="list-style-type: none"> A Crown body independent of both the National Assembly for Wales and the Welsh Government but funded by the Welsh Government Inspect quality and standards in education and training providers 	<p>Education and Training Inspectorate</p> <ul style="list-style-type: none"> A "unitary" inspectorate and part of the DE Provides independent inspection services and policy advice for the DE
	<p>The Office of Qualifications and Examinations Regulation (Ofqual)</p> <ul style="list-style-type: none"> Non-ministerial government department Regulates qualifications, examinations and assessments 	<p>Scottish Qualifications Authority (SQA)</p> <ul style="list-style-type: none"> An executive non-departmental public body that reports to Scottish Ministers and the Scottish Parliament National accreditation and awarding body Regulates qualifications, examinations and assessments 	<p>Qualifications Wales</p> <ul style="list-style-type: none"> Independent statutory body funded by the Welsh Government Responsible for regulating general and vocational qualifications Regulates awarding bodies 	<p>CCEA Regulation</p> <ul style="list-style-type: none"> An independent function within CCEA Responsible for the accreditation and regulation of regulated qualifications

	England	Scotland	Wales	N Ireland
	Standards and Testing Agency (STA) <ul style="list-style-type: none"> • Executive agency • Provides a testing, assessment and moderation system to measure and monitor pupils' progress from reception to the end of Key Stage 2 • Develops and delivers the professional skills test for trainee teachers 			
	Education and Skills Funding Agency (ESFA) <ul style="list-style-type: none"> • Executive agency • Accountable for funding education and skills 			
	Teaching Regulation Agency <ul style="list-style-type: none"> • Executive agency • Regulates the teaching profession, including misconduct hearings and the maintenance of a record of teachers 	General Teaching Council for Scotland <ul style="list-style-type: none"> • Independent professional body • Maintains and enhances teaching standards • Promotes and regulates the teaching profession 	Education Workforce Council <ul style="list-style-type: none"> • Independent regulator for the education workforce • Contributes to improving the standards of teaching and the quality of learning 	General Teaching Council for Northern Ireland <ul style="list-style-type: none"> • The statutory, independent, regulatory body for the teaching profession dedicated to enhancing the status of teaching.
Micro level/ Tier 3: Policy enactment	Schools School leaders Teachers			

Note. Information was collected from various government websites, including those for the listed bodies. Except for Wales, the categorisation into levels/tiers was based on our own understandings. The Welsh Government (2017) published a document specifying the three tiers.

Assessment in England at a crossroads: which way should we go?

Tony Leech (Research Division)

Introduction

Assessment policy in England is closely scrutinised at public and political levels. School-level assessments, especially GCSEs and A levels and their vocational equivalents, have significant stakes for candidates and for wider society. These include school accountability purposes and selection to higher and further education. Such assessments, particularly given their high stakes, are frequently critiqued for, among other things, perceived unfairness in outcomes, deleterious effects on candidate wellbeing and other problems. In addition, assessment policy closely connects to politics at the level of ideology. GCSEs and A levels have a major cultural position and are often discussed in the media, especially around the time of the publication of their results every August.

A significant reform of many aspects of England's assessment policy, including curriculum content changes and the removal of modularity in most qualifications, occurred in the early 2010s while Michael Gove was Secretary of State for Education. This reform, which resulted in new A levels and GCSEs being studied in the years from 2015, was controversial and substantial. Little major change in assessment policy has taken place since, particularly for general qualifications, except as required to deal with the emergency circumstances brought about by the pandemic. Partly as a result of this limited change, and as a consequence of the fact that the Covid-19 pandemic has upended previous certainties about assessment and education, a number of stakeholders have since 2020 published reports into how things might be done differently in the future. This article will briefly review these reports, and explore similarities and differences between them and the policy changes they recommend.

Reports under discussion

A total of seven different reports (two of which are a series from the same organisation) will be analysed. They have different remits and areas of interest. Some explore the education system (or systems – a number of the reports cover all four nations of the United Kingdom, despite their different education systems) more widely, rather than focusing on assessment.

The Independent Assessment Commission, funded by the National Education Union trade union, was chaired by Professor Louise Hayward and conducted a review of assessment and qualifications in England for learners aged 14–19. Its

final report, entitled *Qualifications for a New Era: Equitable, Reliable Assessment*, was published in 2022. It delineated five principles that the assessment system should meet, and 10 recommendations for how it should be designed in order to achieve those aims.

At the suggestion of Sir Anthony Seldon, *The Times* newspaper set up a 22-member Education Commission in 2021 led by its columnist Rachel Sylvester. According to the introduction to the Commission's report, written by the paper's then editor John Witherow, its aim was to "examine Britain's whole education system and consider its future in the light of the Covid-19 crisis, declining social mobility, new technology and the changing nature of work". Its final report presented a 12 point plan for education, with recommendations for the qualifications system as well as for school inspection, school technology, curriculum, universities, teacher training and other areas of the education system more generally.

Pearson, the publisher and provider of qualifications such as GCSEs, A levels and BTECs, undertook a Future of Qualifications & Assessment project, resulting in a report, *Qualified to succeed: Building a 14-19 education system of choice, diversity and opportunity*. The expert panel which advised the project included high-profile educationalists and former education secretaries. It explored the opinions of over 6000 stakeholders across the education system via surveys and built on these findings, through literature review and focus groups, to develop recommendations. It agreed four "guiding principles for reform" – empowerment, coherence, adaptability and innovation – and following these principles, made seven specific recommendations for reform in qualifications and assessment. This work sat beside more specific research on issues including standard maintaining, curriculum flexibility and mathematics and English post-16 resit policy, as part of a broader programme.

The education think tank EDSK, led by former Department for Education adviser Tom Richmond, has also contributed to this discussion. In two reports under the overarching title of *Re-assessing the future* (Richmond, 2021; Richmond & Regan, 2021), the organisation set out a vision for a wide-ranging reform of secondary assessment. Focusing on four major objectives – rigour, coherence, value and aspiration – it outlined a series of far-reaching proposed changes to the education system, including in relation to digital assessment, the institutions that deliver secondary education, accountability metrics and much else. Its approach is considerably more prescriptive and directive.

Another radical new vision for the qualifications system is given by the Tony Blair Institute for Global Change, *Ending the Big Squeeze on Skills: How to Futureproof Education in England* (Coulter et al., 2022). Given what it describes as a situation in which "the new technologies of the Fourth Industrial Revolution are profoundly altering society, the economy and the labour market", it argues for a major rebalancing of English education away from "passive forms of learning focused on direct instruction and memorisation" towards the greater development of transferable skills. It suggests that "at the core of a reformed system should

be a revised curriculum, more sophisticated modes of assessment and a new, rigorous accountability framework” as well as a “comprehensive edtech strategy” (pp. 3–4). The policy recommendations it makes include new performance and accountability measures, the development of a baccalaureate-style qualification featuring multimodal assessments, changes to inspection methods and a national curriculum reform.

The National Baccalaureate Trust has also made, unsurprisingly, proposals for a national baccalaureate for England to replace (or better organise) existing qualifications into one framework for upper secondary education. The approach it recommends is evocative of the International Baccalaureate. It contains core learning modules and compulsory personal development and extended project elements. This structure must be, it argues, universal for all learners, deliverable in existing institutions and reflective of current models of learning and subjects. Issues of credit and assessment structures are also explored in detail.

Finally, there are a number of other organisations devoted to change, such as Rethinking Assessment, a coalition of teachers and others, which has explored a variety of the themes that will be considered here in various short reports and blogs. Other stakeholders have also contributed to online and media discussions on similar assessment reform themes. For example, the Association of School and College Leaders (ASCL) also in 2021 published a *Blueprint for a Fairer Education System*. Drawing on research evidence and expert opinion, it proposes that the essential aim of the education system should be that: “All children and young people receive a high-quality, broad and challenging education. No child or young person receives a lower standard of education as a result of their background or where they live. Schools and colleges are supported to do everything they can to counteract the socio-economic disadvantages faced by some children and young people.” To institute this, it highlights five building blocks, one being assessment and qualifications. It does not propose radical changes to assessment systems, but focuses on streamlining GCSEs, a review of assessment methods, and the close integration of assessment and curriculum.

Some individuals have contributed to or consulted on more than one of the above-mentioned reports, and indeed some of the reports refer to others of those considered. Some ideas expressed are therefore common within the community of English educationalists.

Main themes of the reports

Some significant ideas are common to many of the reports. Most of the reports start from the premise that the present system is flawed in some major ways, including in relation to the volume of assessment and its alleged waning connection to the world of work, though the extent to which they argue that it requires *radical* reform varies. This article will focus on four specific areas as these are covered, to some degree at least, in all the reports. These are:

- whether high stakes assessment at age 16 is necessary, and if so to what degree

- the issue of how many subjects should be studied at what age
- whether digital assessments should be used more substantially
- the relationship between academic and vocational study.

Presenting what the reports say on these issues devoid of the specific contexts to which they respond provides only a partial view of the logic underpinning the recommendations offered. Each report, when read in full, can be seen to respond to particular issues to different degrees. For example, the Tony Blair Institute report is specifically animated by the question of reforming the education system to be able to provide learners with the skills they will need, it argues, to deal with “the new technologies of the Fourth Industrial Revolution” and to “thrive in a world increasingly shaped by automation and artificial intelligence”. Meanwhile, Pearson’s report exists in the context of it already being a major provider of qualifications, while the National Baccalaureate Trust seeks directly to propose a specific new system.

For readers to gain a stronger understanding of the logic of the specific reports, they are recommended to read them individually, as restating such logic is not the purpose of this article. Instead, it draws out and discusses general themes, and concludes by highlighting, with Freedman (2022), that the best way forward for policy in this area would be one based on “incremental” improvement and evidence-led reform.

Table 1: References in each report to specific issues.

Issue	<i>High stakes assessment at age 16?</i>	<i>Use of online and/or digital assessment</i>	<i>Breadth of study across subjects (especially post-16)</i>	<i>Academic/vocational relationship</i>
Report				
Times Education Commission	Exams in five core subjects (including maths and English) at 16. Continuous teacher assessment, online tests, presumably in other subjects	Use in core exams at 16 and in Digital Learner Profile, and in other contexts where possible, integrated into learning	British Baccalaureate at 18: academic and vocational, humanities/sciences, communication, critical thinking, extended project, community service, literacy, numeracy	Include academic and vocational options within Baccalaureate
Pearson	“Make GCSEs work better”; e.g. reintroduce varied types of assessment. Unnecessary to abolish GCSEs	Accelerate digital transformation. Technology can add value to assessment, but there are issues	Highlights pathway inflexibility and narrow curriculum in present system, but direct suggestions for change not offered	Notes that stakeholders say purely academic or vocational is not appropriate
EDSK	New online low-stakes exams one year earlier at age 15 in most curriculum subjects to replace GCSEs	Online assessment at 15. For upper secondary also considers it possible, but challenges	Baccalaureate from 15 to 18 with subject options at 3 difficulty levels (Foundation, Standard, Higher). English and maths to be compulsory	Three pathways (Academic, Applied and Technical)
Independent Assessment Commission	Assessment when ready (14–19). GCSEs in “present form ... need to change fundamentally”	“Deploy existing and emergent technologies to support high quality ... experiences in assessment”	Assessment when ready (14–19). Integrate academic and vocational subjects, extended project, community work into coherent 14–19 package	Academic and vocational subject options should be integrated into one package
Tony Blair Institute	Replace GCSE; some low stakes assessment at 16 to inform pupil choice and for school accountability	Build digital infrastructure (learner ID and digital profile) but no reference to digital exams	(Eventually) build a new qualification based on principles of IB. Continuous assessment from 16 to 18	General support for idea of greater esteem for vocational
National Baccalaureate Trust	Maintain assessment at 16 in a lower stakes context. Broadly retain existing qualifications	<i>No particular reference</i>	More subjects (less content). Extended project, PE, arts, community service, leadership, work experience. Could require study across subject groups	Proposed programmes can include existing vocational qualifications in the Baccalaureate

High stakes assessment at age 16

The question of the continued necessity of high stakes assessment at 16 is one which particularly animates discussion and is a significant feature of most of the reports mentioned here. In the wider media, this is often phrased as a binary question – should we scrap GCSEs or not? The majority of the reports discussed here engage with this argument more subtly, by focusing on the question of what high stakes assessment at 16 should achieve. This is in the context of young people in England being required since 2014 to attend full-time education in some form (which could include technical study or an apprenticeship) until they are 18; consequently GCSEs are no longer a point at which one’s school career can finish. GCSEs remain important as a mechanism of assessing what has been learned in the 14–16 phase, and certificating the wider range of subjects studied in secondary education before learners specialise in fewer areas. However, the necessity of them, given the other uses to which they (and their results) are put, including school accountability and selection for post-16 study, is perhaps more contestable. Many of the reports discussed are more sceptical of the need for them.

There is a general consensus across the reports (and other recent work from organisations including Rethinking Assessment) that GCSEs should change, though opinions as to what the extent of this change should be vary significantly. The National Baccalaureate Trust proposals appear to be largely compatible with existing GCSEs, while Pearson suggests that GCSEs could be made to work better. The Independent Assessment Commission, however, argues that GCSEs need to “change fundamentally” or be abolished. Different approaches are proposed. A number of the reports argue for a streamlined selection of exams in “core subjects” (with the smaller number of exams perceived to be beneficial for candidate wellbeing and efficiency, but still able to fulfil the functions of a wider suite). This is valuable, though it should be noted that streamlining in this respect could be viewed as describing having exams in fewer subjects, having fewer exams in the same number of subjects, or indeed doing both. Different arguments must be made for each of these options.

The desire to assess fewer subjects is an understandable one, especially given that, at younger ages, national assessment in England focuses only on English language, mathematics and science. At Key Stage 4, these are the core subjects in the national curriculum, though there are other compulsory subjects. However, deciding what should count as a “core” subject, and what should not, is fraught with controversy, especially when this contributes to school accountability. The EBacc performance measure (wherein schools are measured on both how many pupils take a specific set of GCSEs, and how well they do in them), contains English language and literature, maths, the sciences, geography or history and an ancient or modern language, on the basis that these are “considered essential to many degrees and open up lots of doors” (Department for Education, 2019). However, as Ashton & Ashton (2022) have found, the performance-measure focus on these subjects led to a narrowing of the curriculum: more schools spending more of their teaching time on these specific subjects, and consequently less on others such as

creative arts or design subjects. It is difficult to see how such curriculum narrowing would not also occur if the subjects on which accountability mechanisms were based were further limited (unless other ways in which it could be demanded that teaching time be set aside for other subjects were implemented, perhaps through inspection, or other accountability reforms).

Most of the writers of the reports we have considered are aware of this issue, and that of candidate wellbeing during exams, and therefore approach the question of the necessity of high stakes assessment at 16 as part of a wider reform package. For some, part of the solution is ensuring that assessment at 16 has lower stakes. For example, for Coulter et al. (2022), while GCSEs should be abolished, there is a role for “low-stakes assessments at 16 to inform pupil choice and hold schools to account”. EDSK proposes online low-stakes assessment at the end of the lower secondary phase of education.

One might reasonably question, however, the extent to which the stakes of an assessment can simply be *declared*. An assessment’s stakes for users are a function of the decisions which will be made using its outcomes, and the impact of these decisions on candidates’ lives. If an assessment at 16, for instance, restricts access to particular post-16 courses of study or the results are used in the allocation of funding to schools, then it will be over time taken more seriously (by candidates or teachers or both) and hence will take on greater stakes. For proposals to make assessment at 16 low in stakes to be meaningful, it would need to be more or less impossible to use those assessments for selection or accountability purposes (say, for example, if all post-16 students were able to study whatever they wanted regardless of results). There are certainly systems in the world which successfully use assessment at 16 for very different purposes and stakes can therefore be lower (Suto & Oates, 2021). What is critical is that all these different elements of the system (assessment, accountability and teaching, in particular) are aligned – not only in design, but in reality.

Subject breadth

Many of the reports argue that, particularly at post-16, too few subjects are studied by learners in England. While “too few” is a subjective statement, it is true that subject breadth in England post-16 is lower than in many similar countries. The average number of qualifications taken post-16 has fallen significantly since 2016, largely as a result of the decision to make AS levels standalone qualifications (thus meaning that, where previously, students were likely to start around four qualifications in Year 12, do AS levels in all of them and then drop one, but retaining the AS level as the exit qualification in that subject and proceeding to A level in the others, now students tend to start fewer qualifications). The average number of A levels taken has remained static at just over 2.6 for the last five years (Ofqual, 2022).

In many European countries, baccalaureate structures, by contrast, mean it is common for more subjects to be studied to 18. The same general approach is taken in the International Baccalaureate, whose Diploma Programme involves

study in six subjects (generally across six different subject groups), plus additional requirements around essay writing, community service and action, and study of the theory of knowledge. The National Baccalaureate Trust proposals for England (unsurprisingly) approximate this model, though with some flexibility to allow their baccalaureate to wrap around existing qualifications. In addition, however, both the *Times* commission and EDSK offer recommendations in a baccalaureate-like form, with quite specific detail offered about which subjects should be studied, while both the Independent Assessment Commission and the Tony Blair Institute argue for a new structure built more as a baccalaureate (with less explicit focus on what should be included, however).

Others have noted benefits to increasing post-16 subject breadth. Education Policy Institute and Royal Society research (Robinson & Bunting, 2021) found that students who took post-16 qualifications from more than one subject group had higher average earnings than those who didn't, by the time they were 26. The flexibility offered by study across subject groups is valuable for the workplace, in that transferable skills are increasingly cited as essential by employers (e.g., Hofman et al., 2022). There are cultural and social benefits to avoiding the bifurcation of skills, experiences and interests associated with a binary funnelling of individuals at 16 into primarily either STEM-only or humanities-only routes.

However, any substantial changes of this form would have major costs. A greater number of teachers would be required (where there are already challenges in teacher recruitment and retention), and those that remained would have to adjust to considerable changes in the structure and content of post-16 courses. There would likely also be knock-on effects to university study, as the slimming down of content in each subject at A level that would be necessary to allow students to study more courses in the same period of time would mean that they would be less well prepared for university in specific subjects. Also, students may not appreciate more constraints on what they can study.

There are options available to policymakers that would suit the goals of those seeking greater subject breadth without a radical transformation of the system. For example, the Core Maths qualification, which provides a basis for learners who want to use mathematical and statistical skills in everyday contexts, is designed to support mathematical skills required in other A levels and is equal in size and UCAS tariff points to an AS level, could be promoted more strongly as a fourth option for post-16 students. The Extended Project Qualification similarly provides a strong foundation in the kinds of writing, research and problem-solving skills necessary for success in further study. Embedding an expectation that both STEM and humanities subjects be continued post-16 for most students could be achieved using better careers and university application guidance. Approaches to ensuring that candidates are assessed more holistically, including on their community service and action, for example, could be built into "simple baccalaureate" schemes that wrap around existing post-16 qualifications and activities but present an overall score. Overall, the extent to which a new proposal would achieve the aim of increasing subject breadth in a meaningful and relevant way would likely depend on how much it was developed as part of a

coherent package co-produced by awarding organisations, schools and colleges, universities, workplaces and other stakeholders, not to mention potential students themselves.

Academic and vocational study

Another significant question concerning subject breadth is that of the relationship between the “academic” and “vocational” in learners’ programmes of study, and the assessment systems for them. It is a longstanding feature of the English education system that “vocational” education – that is, qualifications which are intended to prepare learners for the world of work, rather than further study, have suffered relatively to academic qualifications in terms of funding, esteem and support from government (Relly, 2021). For example, while A levels have been the main academic post-16 qualification of choice since the 1950s, recent decades have seen many short-term, not long-lasting, attempts to build new vocational qualifications including NVQs and Diplomas, a proliferation of short-term funding solutions for further education colleges and a diffusion of responsibility for vocational education.

Many of the future of assessment reports argue for giving vocational qualifications parity of esteem with academic ones. For example, the *Times* commission suggests that both academic and vocational qualifications should be integrated within its proposed Baccalaureate “under the same umbrella”, with further prestige also given to vocational education by the creation of “high-quality technical and vocational sixth forms” and the ability for post-18 funding to support students in colleges as well as universities. EDSK proposes a system of different, equally prestigious, routes through the upper secondary education system (academic, applied and technical). In a different way, respondents to Pearson research rejected the idea of a false binary (or trinary) between the different routes, arguing instead for an approach that recognises choice at the subject level, with students to take a mixture of academic, vocational and applied subjects.

However, as Ewart Keep has long argued, “without active commitment and participation by a critical mass of employers” (Keep, 2020, p. 500), vocational education and training will struggle to reach its potential. Keep has highlighted how a key characteristic of the vocational training system in England which separates it from higher-performing systems such as that of Germany is the general unwillingness of employers to contribute as much as is needed to the training of their own employees (Keep, 2020). As a result, this task has mostly fallen to state education. In Germany, vocational education and training policy has been consistent for decades, as a result of being built on an established system of social partnership between governments, firms and workers, as represented through their trade unions. In England, no such partnership can be said to exist. Approaches to the development of vocational qualifications premised chiefly on a state-based top-down reform are likely to fail without much greater focus being placed on the employer’s role in training than has hitherto been the case, even within apprenticeships. Other described weaknesses in the UK apprenticeship

system include the short terms of many apprenticeships, the fact they tend to be largely classroom-based and a weak alignment between apprenticeships and labour market need. If implemented as part of a wider skills strategy, the gradual reform of vocational qualifications to meet these challenges would likely be generally commended.

In relation to Applied General qualifications such as OCR's Cambridge Technicals and Pearson's BTECs, a recent (2020–21) government attempt to largely eliminate them and reinforce a binary of academic A levels and vocational T levels and apprenticeships for post-16 students was largely defeated by a wide coalition of stakeholders. This highlights the extent to which student choice is regarded as a strength of the English system. Proposals to *forcibly* redesign the relationship between academic and vocational qualifications in candidates' programmes of study that do not take account of the value of student choice would be similarly vulnerable to attack. However, ensuring that Applied General and similar qualifications are popular with future candidates, are appropriate preparation for work in their subjects and recognised as such, are comparable in terms of difficulty to A levels and utilise a strong breadth of assessment types, such that they can be justifiably esteemed alongside A levels, are all valuable areas for further investigation.

Role of digital assessment

Finally, most of the reports highlight the many affordances of digital and/or online technology for improving the English assessment system. Particularly given the disruptions to education and assessment caused by the Covid-19 pandemic, digital assessment is seen by many as an important next step. Perceived benefits are varied, but are seen to include the following:

- personalisation of assessment (including for example the use of adaptive technology to ensure that questions are more appropriately targeted at candidates' ability)
- resilience of assessment (as assessments could be taken at different times throughout a course of study, rather than all at the end thereof)
- assessment of different skills (using technology to do things in assessment that are not possible with pen and paper, and ensuring that assessment is more relevant to the ways of working learners will experience in the workplace or further study)
- feedback through assessment (using digital assessment to demonstrate more directly to learners than in exams their areas of strength and weakness)
- wellbeing during assessment (as the use of adaptivity or other online assessment affordances could result in assessments of similar reliability being undertaken in less candidate time, and therefore potentially support candidates' mental health).

There would also likely be financial savings on the printing and the administration of pen-and-paper exams, if properly rolled out as part of a national system. Countries including New Zealand have in the last few years converted their assessment systems to digital, with significant benefits. In England, the most immediately relevant benefits for higher stakes assessments from the list above

may be those relating to system resilience, wellbeing and the ability to assess different skills, which may have important strengths in terms of validity.

However, rolling out a mass digital assessment system has many barriers. The issue of reliable access to the internet in order to conduct assessments, whether at schools or candidates' homes, is at the heart of equity concerns in this area. It would be necessary not only for candidates to have access to the digital assessment technologies during the assessments themselves, but also throughout teaching and learning periods related to them, so they can become familiar with the processes and requirements, and how to use the technology. The use of digital assessment in higher stakes contexts than hitherto would also require extensive testing and development. It is also likely that a single national procurement for the technology would be necessary for reasons of consistency and simplicity at centre level. At present, in the absence of this, each awarding organisation (including Cambridge University Press & Assessment's OCR exam board) is developing its own approach to digital assessment, which is challenging in terms of the ability to develop national standards.

Moreover, it should not be forgotten that digital assessment changes the constructs being assessed, even in a situation where pre-existing pen-and-paper assessments are merely "lifted and shifted" to a digital delivery system (Puhan & Kim, 2022). Thus, the greater use of digital assessment implies and requires considerable work on comparability of assessments before considering the technology ready to use, not only in relation to comparability between digital and paper systems *in general*, but in relation to candidates from particular identity groups, socioeconomic groups or ability groups (Hughes & Elliott, 2022).

An approach to the digital transformation that appropriately took account of these issues, and therefore had a clear focus, was devoted to equitable access and had a national development and testing model, would be a positive step for England. Areas of further thinking in this regard with particular relevance to formal assessments would include the possibility of streamlining assessments, especially at GCSE, to ensure more efficient and reliable grading while reducing the burden of assessment on candidates. There are also a number of significant affordances of greater digital *formative* assessment.

Conclusion

It is heartening that educational assessment in England is the subject of profound, broad, impassioned and often well-evidenced discussion and debate. It is right that the areas covered above are brought to the attention of policymakers in education, and debates within them supported by the best evidence and expertise. The most satisfying elements of these reports are those which start from clear premises – statements of what should be achieved by the education system, and particularly assessment within it – and consequently argue for a coherent but parsimonious set of reforms that can best achieve those aims.

As Freedman (2022) has argued, many of the more radical approaches set out do not necessarily have the strongest base of evidence behind them. In many cases

the perceived benefits of radical changes would have considerable costs and would themselves provoke further costly changes (e.g. to university admissions processes). The English education system would therefore benefit more from a model of “incremental improvement around assessment” (Freedman, 2022). As part of a wider expert-led and evidence-based strategy, there are valuable changes that could be made in terms of streamlining, updating and digitising assessment, as well as considering the breadth and depth of the curriculum and the relationship between different subjects. A model of evolution, not revolution, would allow policymakers and stakeholders the benefit of being able to carefully reflect on what works and what does not from the present system, and ensure that changes proposed have real value in making education and assessment better for all learners.

References

- ASCL. (2021). *A great education for every child: The ASCL blueprint for a fairer education system*.
- Ashton, H., & Ashton, D. (2022). Creativity and the curriculum: Educational apartheid in 21st Century England, a European outlier? *International Journal of Cultural Policy*. doi.org/10.1080/10286632.2022.2058497
- Coulter, S., Iosad, A., & Scales, J. (2022). *Ending the big squeeze on skills: How to futureproof education in England*. Tony Blair Institute for Global Change.
- Department for Education. (2019). *Guidance: English Baccalaureate (EBacc)*.
- Freedman, S. (2022). *The exam question. Changing the model of assessment reform*. Institute for Government.
- Hayward, L. et al. (2022). *Qualifications for a new era: Equitable, reliable assessment*. Independent Assessment Commission report.
- Hofman, J., Doyle, J., Haider, A., & Bruckmayer, M. (2022). *Skills needs in selected occupations over the next 5-10 years: Research report*. RAND Europe, for the Department of Education.
- Hughes, S., & Elliott, G. (2022, October 4). *How can we balance innovation and comparability in our digital high stakes assessments?* Cambridge University Press & Assessment blog.
- Keep, E. (2020). Employers, the ghost at the feast. *Journal of Education and Work*, 33(7-8), 500–506. doi.org/10.1080/13639080.2020.1852501
- National Baccalaureate Trust. (2022). *Proposals for a National Baccalaureate for England*.
- Ofqual. (2022). *Infographics for A level results 2022*.
- Parker, K. (2022). *Dylan Wiliam's vision for fair and accurate assessment*. TES Magazine.
- Pearson. (2022). *Qualified to succeed: Building a 14-19 education system of choice, diversity and opportunity*.
- Puhan, G., & Kim, S. (2022). Score comparability issues with at-home testing and how to address them. *Journal of Educational Measurement*, 59(2), 161–179. doi.org/10.1111/jedm.12324
- Relly, S. (2021). The political rhetoric of parity of esteem. *Oxford Review of Education*, 47(4), 513–528. doi.org/10.1080/03054985.2020.1866522
- Richmond, T. (2021). *Re-assessing the future. Part 1 – how to move beyond GCSEs*. EDSK.
- Richmond, T., & Regan, E. (2021). *Re-assessing the future. Part 2 – the final years of secondary education*. EDSK.

Robinson, D., & Bunting, F. (2021). *A narrowing path to success? 16-19 curriculum breadth and employment outcomes*. Education Policy Institute/The Royal Society.

Suto, I., & Oates, T. (2021). *High stakes testing after basic secondary education: How and why is it done in high-performing education systems?* Cambridge Assessment research report.

Sylvester, R. et al (2022). *Times Education Commission report*.

Research News

Lisa Bowett (Research Division)

The following reports and articles have been published since *Research Matters*, Issue 34:

Journal articles and other publications

Crisp, V., & Greatorex, J. (2023). The appliance of science: exploring the use of context in reformed GCSE science examinations. *Assessment in Education: Principles, Policy & Practice*. doi.org/10.1080/0969594X.2022.2156980

Greatorex, J. - Cambridge Partnership for Education (2022). *An analysis of cultural representations of India and the UK in English subject curricula*. British Council India.

Oates, T., Crato, N., & Patrinos, (2022). We cannot ignore the reality of global Covid learning loss. *TES Magazine*.

Research and statistics reports on our website

Carroll, M., & Gill, T. (2023). *Progression from GCSE to A Level, 2018 – 2020. Statistics Report Series No. 129*.

Crisp, V., & Ireland, J. (2022). *A structure for analysing features of digital assessments that may affect the constructs assessed*.

Gill, T. (2022). *Are students who take the Extended Project Qualification better prepared for higher education study?*

Gill, T. (2022). *Uptake and results in the Extended Project Qualification*.

Johnson, M., & Majewska, D. (2022). *Formal, non-formal, and informal learning: What are they, and how can we research them?*

Majewska, D., Rushton, N., & Shaw, S. (2022). *How did we get here? Timelines showing changes to maths education in England and the United States*.

Majewska, D. (2023). *Scientific literacy – what can we learn from high performing jurisdictions?*

Mouthaan, M., & Vitello, S. (2022). *What impacts success in proofreading? A literature review of text feature effects*.

Vidal Rodeiro, C. L., & Macinska, S. (2022). *Equal opportunity or unfair advantage? The impact of test accommodations on performance in high-stakes assessments*.

Vitello, S. (2022). *What impacts success in proofreading? A literature review of proofreading on screen vs on paper*.

Williamson, J. (2023). *The feasibility of on-screen mocks in maths and science*.

Conference presentations

Oates, T. (2023, February 10). *The textbooks of the future*. Stockholm University.

Oates, T. (2023, February 09). *Changing texts and educational media*. Designs for Learning.

Oates, T. (2023, January 15). *Rethinking Education from the Ground Up*. EA Sustain.

Oates, T. (2022, September 29). *School Improvement: Are you Inspection Ready?* Westminster Insight – School Inspections Conference.

Oates, T. (2022, October 21–22). *What personalised learning means in the modern world*. Nazarbayev Intellectual Schools. AEO XIII International Research-to-Practice Conference. Teaching, Educating, Loving: The Year of Children in Kazakhstan.

The AEA-Europe Conference 2022 took place online from 9 to 12 November 2022, with the theme ‘New Visions for Assessment in Uncertain Times’. Our researchers presented a total of 13 papers:

Constantinou, F. *Creativity in examination question writing: How novel can examination questions really be?*

Constantinou, F., & Carroll, M. *Online teaching during the Covid-19 pandemic: an exploration of the nature and quality of teacher–student communication*.

De Groot, C. E. *Estimation of component marks during a pandemic*.

Elliott, G. *The post-pandemic comparability narrative. What changes might we expect?*

Hughes, S. *Providing an evidence-base to inform digital assessment design*.

Jellis, C. *Cultural Challenges in developing an assessment for Indian children during a pandemic*.

Johnson, M., Tzagari, D., Richardson, M., Correia, C., & Child, S. *Symposium: Exploring the role of Assessment Literacy in times of uncertainty*.

Leech, T., & Chambers, L. *How do judges in Comparative Judgement exercises make their judgements?*

Mistry, S. *A learner centred approach to digital assessment item type design and development*.

Morley, F. *Annotation consistency, measured: A methodological poster*.

Oates, T. *A long weekend in Summer 2020 – exams in crisis*.

Vidal Rodeiro, C. L., & Chambers, L. *Online moderation of non-exam assessments: is Comparative Judgement a practical alternative?*

Vitello, S., & Leech, T. *Reflections on teacher assessment after the 2021 Teacher Assessed Grades process in England*.

The BERA Conference 2022 took place online from 5 to 8 September. Our researchers presented a total of three papers:

Greatorex, J., Kreijkes, P., & Majewska, D. *Exploring representations of culture in the UK nations' national curricula for English Language and English Literature.*

Johnson, M., & Coleman, V. *Teacher workload and wellbeing during lockdown in England: insights from a teacher diary study.*

Vidal Rodeiro, C. L., & Macinska, S. *Teachers' and students' views of access arrangements in high-stakes assessments.*

Kreijkes, P. (2022). *A bird's-eye view of curriculum publications concerning seven countries: A bibliometric analysis.* Proceedings of The European Conference on Education, United Kingdom, 2188–2162.

Majewska, D., Rushton, N., & Shaw, S. (2022, September 26–29). *A timeline is worth a thousand words: The history of maths education in England and the United States* [Poster presentations]. Annual conference of the International Society for Design and Development in Education, University of Nottingham.

Rushton, N., Majewska, D., & Shaw, S. (2022). *Different approaches to the curriculum mapping of mathematics through the lenses of two contrasting educational jurisdictions* [Paper presentation]. BAICE Conference 2022, University of Edinburgh, Scotland.

Blogs and podcasts

The following blogs and podcasts have been published since *Research Matters*, Issue 34:

Oates, T. (2023, February 15). [Sum find it tough: Why we struggle with maths.](#) The Bunker Podcast.

Greatorex, J., Walland, E., Vidal Rodeiro, C. L., Rushton, N., & Elliot, G. (2023, January 10). [How do our office buildings and environments influence working practice and culture?](#)

Rushton, N., Majewska, D., & Shaw, S. (2022, November 08). [Telling the story of maths education in England and the United States.](#)

Vitello, S. (2022, October 27). [We need research! Bringing research insights to our agile digital innovation team.](#)

Hughes, S., & Elliot, G. (2022, October 04). [How can we balance innovation and comparability in our digital high stakes assessments?](#)

Williamson, J. (2022, September 22). [Shifting maths and science assessments onto screen: what's different?](#)

Sharing our research

We aim to make our research as widely available as possible. Listed below are links to the places where you can find our research online:

Journal papers and book chapters: <https://www.cambridgeassessment.org.uk/our-research/all-published-resources/journal-papers-and-book-chapters/>

Research Matters (in full and as PDFs of individual articles): <https://www.cambridgeassessment.org.uk/our-research/all-published-resources/research-matters/>

Conference papers: <https://www.cambridgeassessment.org.uk/our-research/all-published-resources/conference-papers/>

Research reports: <https://www.cambridgeassessment.org.uk/our-research/all-published-resources/research-reports/>

Data Bytes: <https://www.cambridgeassessment.org.uk/our-research/data-bytes/>

Statistics reports: <https://www.cambridgeassessment.org.uk/our-research/all-published-resources/statistical-reports/>

Blogs: <https://www.cambridgeassessment.org.uk/blogs/>

Insights (a platform for sharing our views and research on the big education topics that impact assessment around the globe): <https://www.cambridgeassessment.org.uk/insights/>

Our YouTube channel: https://www.youtube.com/channel/UCNnkOpi7n4Amd_2afMUoKgw contains Research Bytes (short presentations and commentary based on recent conference presentations), our online live debates #CamEdLive, and podcasts.

You can also learn more about our recent activities from [Facebook](#), [Instagram](#), [LinkedIn](#) and [Twitter](#).

Contents / Issue 35 / Spring 2023

- 4 **Foreword:** Tim Oates
- 5 **Editorial:** Tom Bramley
- 6 **Creating Cambridge Learner Profiles: A holistic framework for teacher insights from assessments and evaluations:** Irenka Suto
- 27 **A conceptual approach to validating competence frameworks:** Simon Child and Stuart Shaw
- 41 **Teachers' and students' views of access arrangements in high stakes examinations:** Carmen Vidal Rodeiro and Sylwia Macinska
- 60 **Who controls what and how? A comparison of regulation and autonomy in the UK nations' education systems:** Pia Kreijkes and Martin Johnson
- 80 **Assessment in England at a crossroads: which way should we go?** Tony Leech
- 94 **Research News:** Lisa Bowett

Cambridge University Press & Assessment
Shaftesbury Road
Cambridge
CB2 8EA
United Kingdom

ResearchDivision@cambridge.org
www.cambridge.org