



OECD Education Working Papers No. 253

Upper-secondary education
student assessment
in Scotland: A comparative
perspective

Gordon Stobart

<https://dx.doi.org/10.1787/d8785ddf-en>

DIRECTORATE FOR EDUCATION AND SKILLS

Cancels & replaces the same document of 27 August 2021

**UPPER-SECONDARY EDUCATION STUDENT ASSESSMENT IN
SCOTLAND: A COMPARATIVE PERSPECTIVE**

OECD Education Working Paper No. 253

Gordon Stobart, Emeritus Professor of Education, University College London and Honorary Research Fellow at the Oxford University Centre for Educational Assessment (OUCEA).

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

Beatriz Pont, beatriz.pont@oecd.org
Romane Viennet, romane.viennet@oecd.org.

JT03480286

OECD EDUCATION WORKING PAPERS SERIES

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcome, and may be sent to the Directorate for Education and Skills, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

www.oecd.org/edu/workingpapers

Acknowledgements

This paper was prepared as a complementary analysis to the OECD assessment of the implementation of Scotland's *Curriculum for Excellence*. The author would like to thank Beatriz Pont and Romane Viennet of the OECD Directorate for Education and Skills, for their invaluable guidance and contributions towards the development of the paper. The author also acknowledges the valuable feedback received from Paulo Santiago and Hannah Kitchen, of the OECD Directorate for Education and Skills, and Professor Graham Donaldson, through their comments; and from the stakeholders of Scottish education, through the online consultation seminar. Martyn Ware and colleagues at the Scottish Qualifications Authority (SQA) provided helpful data on Scottish qualifications and Scottish government officials kindly provided a factual review of the paper. Jacqueline Frazer (OECD Directorate for Education and Skills) performed the final editing and formatting.

Abstract

Scotland's *Curriculum for Excellence* (CfE) is a pioneering example of 21st century curriculum reform. With Scotland positioned within the historic British examination tradition, qualifications for upper-secondary school students have seen far less reform. The cancellations of examinations in 2020-21, and the crises these generated, have provided an opportunity to reconsider the upper-secondary assessment system. This paper compares the Scottish system to five other legacy traditions, as well as four other British legacy systems, to offer insight for how Scotland could further improve the alignment between CfE and upper-secondary assessments. The analysis is guided further by theoretical considerations on what constitutes a dependable and trustworthy assessment system, to refine the reflection around possible options for the Scottish system. Three major themes emerge from this comparative review. One focuses on how the external assessments could be more innovative in order to capture a wider range of student capabilities. The second is to rethink the role of teacher assessment, with more emphasis placed on continuous school-based assessment. The third is to better integrate the academic and vocational strands with the assessment system which, given SQA's responsibilities for both, would offer a broader range of curriculum options.

Table of contents

Acknowledgements	3
Abstract	4
1. Introduction	7
1.1. Methodology	8
1.2. Structure of the paper	9
2. An overview of the Scottish assessment system	9
3. Student assessment systems adapting to educational change	13
3.1. Aligning with a broader curriculum	13
3.2. Digital and technical innovations may increase alignment	14
3.3. Developing school based assessment	16
3.4. Changes to the school leaving age	17
3.5. Student diversity	18
4. Student assessments in context: the impact of historical legacies	19
4.1. British examination systems	19
4.2. Alternative legacy traditions	21
4.3. Inhibiting factors in assessment reform	30
5. The COVID-19 ‘stress test’ revealed the fragility of assessment systems	33
5.1. Approaches to adapt upper-secondary examinations across countries	33
5.2. Opportunities for long term change	35
6. What is needed for a dependable assessment system that supports learning?	36
6.1. Dependable assessments: validity, reliability and manageability	36
6.2. What do students want?	38
7. Options from Scotland and beyond for the future of student assessment in upper-secondary education	39
7.1. International examples of alternatives to the British examination legacy	39
7.2. Options for moving beyond legacy models	40
8. Conclusions	48
References	49

Tables

Table 2.1. Typical secondary education study patterns in Scotland	11
Table 2.2. Breakdown of course assessment in National 5, Higher and Advanced Higher Qualifications	11
Table 4.1. Summary of student examination systems	26
Table 4.2. National/central assessments at the upper-secondary level (2015)	28

Table 4.3. Main purposes and features used for reporting results of national/central assessments at the upper-secondary level (2015)	29
--	----

Figures

Figure 6.1. Dependability: the one-handed clock	38
---	----

Boxes

Box 4.1. Reforming the Junior Certificate in Ireland	30
Box 7.1. The Ontario Secondary School Diploma (OSSD)	42

1. Introduction

The aim of this working paper is to examine the opportunities for *constructive alignment* (Biggs and Collis, 1982^[1]) between forward-looking curricula such as *Curriculum for Excellence* (CfE) in Scotland, United Kingdom, the accompanying pedagogy, and their assessment in the upper-secondary education (ISCED 3). The paper complements the broader OECD Review of Scotland's *Curriculum for Excellence* (2021). The review noted the powerful influence of the secondary school assessment system on the curriculum and on pedagogy, which warranted a further study. The paper considers whether, and how, comparative assessment systems are adapting to these broader educational aspirations.

Nine systems are drawn upon which are relevant for Scotland's upper-secondary assessment system (secondary years S4-S6). This relevance is either in terms of family resemblances (England (United Kingdom), henceforth referred to as "England"; Hong Kong, China; Ireland; Wales (United Kingdom), henceforth referred to as "Wales") or through instructive differences (France; New Zealand; Norway; Ontario (Canada) henceforth referred to as "Ontario"; Queensland (Australia), henceforth referred to as "Queensland").

The assumption underlying this paper is that student assessment systems are essentially a social, rather than scientific, process which reflects the history and culture within which they occur (Stobart, 2008^[2]). These traditions are socially embedded and any reforms may well be opposed by both the public and policy makers. Scotland is located within the British tradition of school examinations with its emphasis on single subject end-of-course external examinations, while other systems reveal different origins and traditions. These range from the devolved and teacher led Nordic traditions to the centralised French system. The American system is another devolved system, with college entrance based on a wide range of evidence.

As cultures and education systems change, so upper-secondary assessment may have to change. For example, education systems have been moving towards so-called 21st century curricula focused on student capabilities and competencies (OECD, 2019^[3]); staying-on rates after the end of compulsory schooling have increased; and the student population has been growing more diverse. These transformations may involve, among others, technical innovations and developments in school-based assessments.

Assessment reforms may be inhibited by a number of factors, including the resistance of teachers and parents. In Ireland, opposition from teacher unions to the reform of the Junior Certificate, taken at the end of lower-secondary education (at 15-16 years old), provides a case study. Successful change requires public confidence and trust, something not easily achieved in historically embedded systems. In jurisdictions such as England, the reliance of the high-stakes school accountability system on schools' performance in national examinations limits innovative assessment formats. In systems such as in Hong Kong, China, where results are central to highly competitive university selection, the pressures to ensure comparable results limit what changes can be made.

Additionally, unanticipated events can also deeply affect student assessment systems. In 2020-2021, this is best illustrated by how the COVID-19 pandemic disrupted education and forced decision makers and education stakeholders to find emergency solutions to adapt or replace examinations and student assessment processes overall. Scotland and England sought to generate comparable grade distributions to previous years. In Scotland students were initially awarded grades based on teacher judgements, which were then moderated by the Scottish Qualifications Authority (SQA) using a statistical algorithm. The

outcome was student and public outcry at the unfairness of individual results and a late switch back to teachers' grades for their students, emphasising the fragility of the assessment system. Other systems with a greater emphasis on school-based assessments proved more resilient. Either way, the COVID-19 stress test opened additional opportunities for jurisdictions internationally to reflect on the future of student assessment and to seek possible alternatives to their own system.

How student assessment systems balance the demands of validity, reliability and manageability is a concern for every education system (OECD, 2013^[4]), because these demands come with necessary trade-offs. This paper seeks to contribute to the reflection in Scotland and beyond, by considering several options to develop dependable assessment systems. A dependable assessment is one that can reliably give a trustworthy estimate of students' capabilities. It involves an optimal trade-off between construct validity, reliability, and manageability. Examining comparative approaches to student assessment should provide encouragement and examples for Scotland's decision makers as they consider possible directions for assessment developments, to improve alignment with *Curriculum for Excellence* and the capacities expected of students.

1.1. Methodology

This working paper was commissioned to the author as part of the OECD's work on Scotland's education system. It complements the OECD report *Scotland's Curriculum for Excellence: Into the Future* (2021^[5]). This paper is a comparative study of upper-secondary school assessment systems with Scotland as the focal point. The comparative approach was adopted to position Scotland in relation to other systems, and to offer a range of approaches to upper-secondary school assessment that may be informative for stakeholders in Scotland. The paper is largely organised around broad themes emerging as major factors affecting student assessment, rather than by case-by-case studies of the different jurisdictions. These themes include current educational changes, historic legacies, responses to the COVID-19 pandemic, and the integration of vocational assessment.

The selection of a sample of nine assessment systems from various traditions of student assessment was purposive, as Scotland's legacy plays a significant role in its own system's evolutions. Four of the systems are within the historic British tradition, as is Scotland. These are England, Wales, Ireland and Hong Kong, China. The intention was to consider instructive similarities and differences between them. The remaining five jurisdictions were selected to illustrate contrasting approaches to upper-secondary assessment from other traditions. These particular systems were selected for their relevance to Scotland. New Zealand and Queensland, Australia have broken away from the British model of single-subject examinations and of extensive external examinations. Norway is of similar size to Scotland and represents the Nordic tradition, which has historically put teacher-based assessment at the centre of its system. This is also the case for the province of Ontario in Canada, a high-achieving jurisdiction which represents features of the American system. Although the French Baccalauréat involves a larger student population than Scotland's, it represents another highly centralised system of examinations and has had a wide influence. Both Scotland and Wales have introduced an additional Baccalaureate qualification, and the International Baccalaureate is also a highly regarded alternative upper-secondary qualification. Of relevance to Scotland's *Curriculum for Excellence* is how the French system assesses the breadth of Baccalauréat curriculum.

Information on these international systems was gathered through the analysis of policy documents and research publications. Key sources include the recent international research study *Examination standards: How measures and meanings differ around the world* (Baird et al., 2018^[6]) and articles in the international journal *Assessment in Education, Principles,*

Policy and Practice. Reports from the Organisation of Economic Co-operation and Development (OECD) were also utilised, as well as the information packs containing key statistics and chronological accounts of the evolution of student assessment in upper-secondary education in Scotland provided by the Scottish Government, the Scottish Qualifications Authority (SQA), and independent experts.

A series of online stakeholder meetings was organised with Scottish educationalists (including teachers, learners and other school stakeholders), policy makers and officers from the Scottish Qualifications Authority (SQA). Their purpose was both fact-checking and sounding out suggested options for the current system.

1.2. Structure of the paper

The paper begins with a brief overview of the Scottish upper-secondary assessment system, which involves a complex series of examinations, pathways within them, and teaching approaches (Section 2). Section 3 considers how upper-secondary assessment systems have responded to some of the 21st century curricular reforms currently underway in jurisdictions such as Scotland, Wales, Norway, Hong Kong, China and New Zealand.

Section 4 considers the main reasons why assessment systems have been slow to adapt to these educational changes. A key element among factors inhibiting assessment reforms is that examination systems are a product of historical traditions, often with their origins in the 19th century or earlier.

Section 4 considers how the different jurisdictions managed the 2020 COVID-19 ‘stress test’ when schools closed and, in some countries, national examinations were cancelled.

This leads to considerations in Section 5 of what is needed in a dependable system and the balance of validity, reliability and manageability that is involved. This includes consideration of how Scottish students may want to be assessed, given they are the ones most directly affected by the assessment system.

Section 6 builds upon the argument and examples to consider some of the options, based on comparative analysis, that are available for the Scottish upper-secondary assessment system. It presents six options for consideration that may assist policy makers and educators in further developing an assessment system that is better aligned with *Curriculum for Excellence* and better reflects the demands of 21st century teaching and learning.

2. An overview of the Scottish assessment system

Scotland was an international leader in adopting broader ‘21st century’ capacities for education and life outside schooling. Similar aspirations are found in most of the jurisdictions reviewed. These broader educational goals have led to curriculum reform, of which Scotland’s *Curriculum for Excellence* is a pioneering example. *Curriculum for Excellence* offers a framework for education between the ages of 3 and 18, covering all cycles until upper-secondary education (referred to as “the Senior Phase” in Scotland). Where examinations are, or aim to be, curriculum-related (as opposed, for instance, to skill-based examinations), such reforms directly impact assessment systems. This is especially the case where the curriculum encourages new types of skills (such as socio-emotional and transversal skills), as well as content; and where it seeks to respond to the increasing number and diversity of the student cohort post-16 years old in terms of both cultural and academic diversity (OECD, 2020^[7]; OECD, 2010^[8]).

Wales, France, and Norway are currently rolling out such forward-looking curricula. These widespread curriculum reforms seek to encourage broader skills and capacities. To align with these, assessment systems may need to be modified, especially in those jurisdictions with traditional examinations such as New Zealand and France.

Many jurisdictions around the world found challenges in how to translate these aspirations into their upper-secondary school assessment policies. The new curriculum intentions are often hard to align with historic assessment practices that are embedded in that society. In this paper Scotland is positioned, with England, Wales, Ireland and Hong Kong, China, within the British legacy system (see Section 4.1). Upper-secondary school examinations in this system have their origins in the expanding educational systems of the 19th century. Hunter (1963^[9]) identifies three main purposes of the Scottish Certificate examinations that were introduced in 1888:

- To supplement the inspection of certain secondary-type schools;
- To establish uniform standards of attainment;
- To provide a Certificate acceptable to the universities and other examining or professional bodies. (Hunter, 1963, p. 322^[9]).

The first purpose was a temporary expedient at a time to deal with the expansion of secondary education. However, the use of examination data in evaluating school performance has continued. The role of examinations in setting and raising standards and in selection and certification is still central to their current functions.

The current National Qualifications were immediately preceded by the Standard Grade and Intermediate Grade. When introduced in the 1980s, the intention was to give all young people recognition at the end of compulsory schooling and to drive curriculum and teaching change. These changes included teacher assessments of practical and enquiry elements along with inter-disciplinary courses in science, social science and social and vocational skills. These ambitions were not fulfilled, and the teacher-assessed elements were removed as they were seen as inflating the overall grades. The Intermediates, which were introduced for those staying on in S5 (secondary year 5) but for whom Highers were not suitable were also removed, leaving the prospect of increased failure.¹

Scotland, unlike many jurisdictions, does not provide a school leaving certificate but issues subject-based qualifications. Students who have had limited success in their examinations may therefore have little to show for their school achievements. The current upper-secondary examination system in Scotland is also distinctive in having four different levels of mainstream qualifications over the three years of the upper-secondary phase (S4-6). The current system consists of the teacher assessed National Qualifications 3 and 4, and the externally examined National 5, Higher and Advanced Higher. These replaced the legacy system Access 3, Standard Grade, Intermediate 1 and 2, Higher and Advanced Higher. Other qualifications are also available, with their level indicated within the Scottish Credit and Qualifications Framework (SCQF), with, for example, SCQF 5 incorporating National 5.

Students prepare for qualifications throughout secondary education (see Table 2.1), and a major concern has been whether the qualification study programmes replace *Curriculum for Excellence* as the *de facto* curriculum (OECD, 2021^[5]). The new National Qualifications were introduced in 2013 in order to align more closely with *Curriculum for Excellence*. They were modular, with each unit incorporating examinations and

¹ I am grateful to Graham Donaldson for directing me towards this, and other, Scottish examination history.

teacher-assessed coursework. The assessment of each unit, along with its quality assurance demands led to an unsustainable workload for teachers and in 2016, the Scottish Government requested that the Scottish Qualifications Authority (SQA) remove the requirement for mandatory unit assessment in the qualifications. The Scottish Credit and Qualifications Framework (SCQF) required courses of 160 hours, which meant there was a reduction in the courses that could be studied and a pushback into preparation in S3.

Table 2.1. Typical secondary education study patterns in Scotland

School Phase	Ages	Curriculum phase and <i>typical qualification prepared</i>	Typical number of subjects
Secondary 1-3 (S1-3)	11-14/15 (flexible multi-level teaching in some schools)	Broad General Education	12-15
S4	14-15 /15-16	Senior Phase <i>National 4</i> (Teacher assessed) <i>National 5</i> (Exam/coursework)	6-9
S5	15-16/16-17	Senior Phase <i>Higher</i>	5
S6	16-17/17-18	Senior Phase <i>Advanced Higher</i>	2-3

Source: Scottish Government (2021_[10]) OECD Independent review of Curriculum for Excellence 2020-2021 - Initial evidence pack, and information provided by the SQA (unpublished).

The roll-out of the Revised National Qualifications began in 2017-18, with more emphasis placed on the final course examinations and a reduced contribution in teacher-assessed coursework in some subjects. The current proportions of exams and coursework in three major National Qualifications (*National 5*, *Higher* and *Advanced Higher*) can be found in Table 2.2.

Table 2.2. Breakdown of course assessment in National 5, Higher and Advanced Higher Qualifications

Qualification level	Proportion of final exam in the qualification
Number of qualifications by examination weighting	
National 5 (53)	
1	No exam
17 (inc. 1 with no exam)	Exam less than 50%
19	Exam 51-75%
16 (inc. 4 with no coursework)	Exam 76%+
Higher (41)	
0	No exam
14	Exam less than 50%
18	Exam 51-75%
9	Exam 76%+
Advanced Higher (29)	
4	No exam
9	Exam less than 50%
13	Exam 51-75%
3	Exam 76%+

Source: Information provided by the Scottish Government and SQA.

The Revised National 5 are typically taken in Secondary 4, the end of compulsory education at age 16. However, 88% of students continue into Secondary 5 where most will take Highers, the basis for most selection to Higher Education and Training. In 2018/19 26.8% of students left at this point, with 61.2% leaving after S6 (Scottish Government, 2021_[10]). Of these 19.1% of school leavers gained one pass or more at SCQF Level 7. University education in Scotland has traditionally begun at 17 years, followed by a four-year degree programme, though increasingly, students are staying on in S6 to take Advanced Highers in order to widen their access to Higher Education.

The system is made more complex by the flexibility offered to schools to prepare students for the different levels (National 4 & 5; Highers; Advanced Highers) at different rates. This may lead to multi-level teaching in the same class and different ages of entry, particularly in smaller entry schools. This is problematic in some subjects in which transition from one qualification to the next may involve additions to both the curriculum and skills.

Scotland offers a wide, and increasing, range of vocationally oriented awards, some of which are incorporated in the National Qualifications. Most awards are taken in the form of National Group Awards which, in more occupationally related areas, include National Certificates and National Progression Awards. Along with the recently introduced Foundation Apprenticeships, these can lead to more advanced study at Higher National Certificate and Diploma levels.

There is concern about how this traditional examination system aligns with *Curriculum for Excellence*. At present, there are limited comparative examples of how national examinations manage to assess broader global skills such as creativity, collaboration, and communication. For instance, Scotland and other education systems successfully participate in the Global Competence assessments of the OECD's Programme for International Student Assessment (OECD, 2020_[11]), and countries such as Norway, New Zealand and Finland are currently developing approaches which offer opportunities to use more extensive online resources as part of examinations. Part of the appeal of the International Baccalaureate Diploma (see Section 4.2.4) has been the inclusion of a 1 600 word Theory of Knowledge essay (plus oral), along with an extended essay of 4 000 words and a 'Creativity, activity and service' component as part of its core requirements. These sample a much fuller range of the aspirations of the 21st century curriculum. The evidence from this paper points to those systems with a wider range of assessment approaches being better placed to both meet the demands of curriculum reform and to offer more resilience in the face of massive disruption, such as that seen in the COVID-19 pandemic.

There appears to be limited information about how students in Scotland would prefer to be assessed. The anecdotal evidence from reviews such as the Priestley (2020_[12]) and *Insight #SQAfutures* (SQA and YoungScot Observatory, 2018_[13]) is that students would prefer a system which includes some forms of continuous assessment by their teachers. Student interviews conducted by the OECD team as part of its 2021 assessment of CfE implementation also found similar attitudes from students, who valued the idea of continuous assessment by their teachers. They were also concerned about the 'teaching to the test' they were receiving in preparation for the National Qualifications. There were more positive attitudes towards the Advanced Highers as they were perceived as encouraging greater depth and creativity.

3. Student assessment systems adapting to educational change

Assessment systems are having to adapt to widespread educational and curricular reforms. Scotland's *Curriculum for Excellence* is a pioneering example of broader 21st century educational goals which have led to curriculum reform. Where examinations are curriculum-related, such reforms directly impact assessment systems. This is especially the case where the curriculum encourages new skills as well as content, and responds to the increasing diversity of the student cohort. Wales, France, and New Zealand and Norway are all currently rolling out new curricula. Other transformations in education also affect assessment systems, including digital and technical innovations, developments in school-based assessments, changes to the school leaving age, and increasing student diversity.

3.1. Aligning with a broader curriculum

Education is witnessing widespread curriculum reform which seeks to encourage broader skills and capacities. These encompass not only knowledge acquisition, but the development of a range of skills such as captured in the four fundamental capacities of Scotland's *Curriculum for Excellence*:

- Successful learners;
- Confident individuals;
- Responsible citizens;
- Effective contributors.

Similarly, in Wales, which launched its *Curriculum for Wales* in 2020 for implementation in 2022, the educational goals are to develop learners who are:

- ambitious, capable learners, ready to learn throughout their lives;
- enterprising, creative contributors, ready to play a full part in life and work;
- ethical, informed citizens of Wales and the world;
- healthy, confident individuals, ready to lead fulfilling lives as valued members of society (Welsh Government, 2020^[14]).

The key capacities in this process are creativity and innovation; critical thinking and problem-solving; personal effectiveness; and, planning and organising.

Hong Kong, China summarises its educational goals as:

“The school curriculum should provide all students with essential life-long learning experiences for whole-person development in the domains of ethics, intellect, physical development, social skills and aesthetics, according to individual potential, so that all students can become active, responsible and contributing members of society, the nation and the world.” (Hong Kong Curriculum Development Council, 2001^[15])

The 2020 curriculum reforms in Norway were based on the 2017 *Education Act*. This, too, emphasised a broader vision of education and of the curriculum:

The pupils and apprentices must develop knowledge, skills and attitudes so that they can master their lives and can take part in working life and society. They must have the opportunity to be creative, committed and inquisitive. The pupils and apprentices must learn to think critically and act ethically and with environmental

awareness. They must have joint responsibility and the right to participate.
(Norwegian Ministry of Education and Research, 2017_[16])

In all these jurisdictions, these broader visions for education led to renewed approaches to curriculum and assessment. Norway, for instance, based its approach on competency (Norwegian Directorate of Education and Training, 2018_[17]). Competencies are understood as the ability to acquire and use knowledge and skills in order to overcome challenges and solve problems in *familiar and unfamiliar circumstances* and situations, and to involve *understanding and the ability to reflect and think critically*.

What jurisdictions around the world have found challenging is how to translate these aspirations into their upper-secondary school assessment policies (Gouédard et al., 2020_[18]; Geisinger, 2016_[19]; OECD, 2013_[4]). At present there are few examples of how national examinations can assess broader global skills such as creativity, collaboration, and communication (OECD, 2020_[20]; Hopfenbeck and Stobart, 2015_[21]).

3.2. Digital and technical innovations may increase alignment

The 21st century has also seen dramatic advances in Information Technology and progress in digitalisation in and around education. From integrating computer-based tasks, to developing adaptive assessments for a formative and personalised learner experience, this offers new possibilities for how subjects can be examined. The potential of these tools, and of Artificial Intelligence (AI) in general, to help reintegrate learning and assessment, is significant, and a matter of international interest (OECD, 2021_[22]).

Norway provides an example of this. A decade ago, students at the University of Oslo, soon backed by students in other Norwegian universities, pushed for computer, rather than pen-and-paper, based exams. Students claimed they no longer had the handwriting skills for the traditional type of assessment. Over 2014 and 2015, Norwegian universities pioneered the National Project for Digital Exams, which digitised all exam processes from admittance to final grade. By 2020, 90% of Norwegian examinations were computer-based and this has filtered down to upper-secondary school examinations (Emmertsen, 2020_[23]). The significant development since 2015, has been that students taking secondary school examinations have had access to online resources in some centrally administered examinations. From 2012 to 2015, there was a trial of unrestricted internet access during the examination. In 2018 this was expanded to all candidates in one subject.

There have been national pilot studies in New Zealand, Israel, Norway and Finland which have used computer-delivered examinations. These showed that online and on-screen assessments could be implemented in external, sessional examinations in schools and colleges – though there were considerable logistical challenges in each case (Ofqual, 2020_[24]). These challenges fell into three main groups: the IT provision in schools; implementation issues; and equity and fairness concerns for all students.

In each country there have been concerns with IT hardware resources: were there sufficient computers and space (in Israel public libraries were also utilised) and how compatible was the hardware with the examination specifications? Solutions have ranged from state financed procurement of ‘exam ready’ devices to bring your own device (BYOD) solutions. Online delivery of the examinations also depended on broadband capacity and reliability. In rural Finland, where power cuts are frequent, the assessments are encrypted and downloaded ahead of time and distributed through local networks.

Joint research from the United Kingdom (CEA, SQA and the Welsh Government, 2014_[25]) anticipated some of these challenges, noting the underuse and underfunding of ICT for assessment purposes in schools. The research noted that in individual qualifications,

particularly in vocational areas, there has been a steady increase in the use of digital assessment tools such as e-testing and e-portfolios. For instance, the Middle Years Baccalaureate offered as part of the International Baccalaureate (IB) requires candidates to complete three e-Portfolios of coursework. One of the goals of the 2016 Digital Technology national strategy for Scotland (Scottish Government, 2016^[26]) was to ‘Support, develop and embed approaches to assessment that make effective use of digital technology’. Scotland launched a new digital strategy in 2021 (Scottish Government, 2021^[27]).

Most of the technical innovations by UK examination boards have been in relation to the marking and processing of examination papers, for example, scanning candidates’ papers for marking online and standardising markers. This is done by sending scripts for which the marks were agreed by the chief and senior examiners, as unmarked scripts, allowing them to monitor the examiner’s accuracy. However, this does not impact directly on the candidates for whom the examinations operate in their traditional pen-and-paper format. At this developmental stage, the question is whether the demands of the examination, its content, and the responses required, are essentially ‘business as usual’ with the paper version transferred to a screen or whether new kinds of questions and stimulus materials are being introduced.

As a consequence of the 2020 COVID-19 closure of schools, the US College Board decided that it would not be possible for students to sit the Advanced Placement Tests [see Section 4.2.1] in the traditional way. New examinations for all 28 courses were written specifically in response to the crisis. Most of the tests took 45 minutes and had one or two free-response questions. Students wrote and submitted their responses within the allotted time for each question. Students were able to take tests on any computer, tablet, or smartphone they had access to. They could either type and upload their responses or write responses by hand and electronically submit a photo, creative procedures which inevitably raised reliability issues (College Board, 2021^[28]; Compass Education Group, 2020^[29]).

More radical examination formats are being developed in various commercial settings. Information technology has made new assessment approaches possible in terms of what can be presented on screen by way of:

- Manipulable screen-based stimulus materials. The World Class Tests in Mathematics and Problem-Solving involve manipulating objects on screen in order to solve problems². In Norway, online resources are provided for use during secondary school examinations;
- Access to the internet during the examination. Pilot secondary school examinations in Norway gave candidates unlimited access to the internet;
- ‘Tailored’ adaptive testing based on students’ responses: for instance, Alberta Computer Adaptive Assessment System (CAA); Measures of Academic Progress (MAP);
- Collaborative responses: for instance, the OECD Programme for International Student Assessment collaborative problem-solving skills in which students work collaboratively online.

Such approaches are resource-intensive and expensive to develop – often beyond the reach of national examination boards. Their resource implications, such as laptops and high-capacity broadband, have meant that, at present, they have not been utilised in national

² <https://www.worldclassarena.org/world-class-tests>

examinations in the United Kingdom. These resource demands also raise issues of fairness in relation to socio-economic factors.

Where there is a policy intention to assess a broader range of skills, an option is to add additional assessments to the existing forms. In Wales there is a proposal to further roll-out the *Key Stage 4 Skills Challenge Certificate* taken at 16 in order to address ‘wider skills’ alongside revised single subject-based General Certificates of Secondary Education (GCSE) (Qualifications Wales, 2021^[30]). However, this may do little to address the issue of broader skills within the GCSEs themselves.

3.3. Developing school based assessment

If it is difficult to capture elements of the broader curriculum in conventional ‘pen-and-paper’ examinations (Frederiksen and Collins, 1989^[31]), one solution is to entrust more of the summative assessment to classroom teachers, under certain conditions. This may require teachers’ further professional development in order to increase their validity and to minimise problems of reliability and bias (Moss, 2013^[32]; Black et al., 2010^[33]; Martínez, Stecher and Borko, 2009^[34]). Performance in the classroom and other settings than conventional examinations allow for adaptable tools to assess skills such as the ‘4Cs’ (critical thinking, communication, collaboration, and creativity). This is routinely done in vocational qualifications and in ‘high trust’ cultures in which teachers’ assessments are the basis of certification (for example, Norway, Queensland and Ontario). In classroom-based assessment, the outcomes are determined by the teacher. These can be tailored to local circumstances and can be designed to demonstrate a wider range of skills.

Beyond professional development of teachers themselves, additional tools may be needed to increase validity and minimise the risk of bias or lesser reliability with school-based assessment. An effective moderation process that ensures consistency across the system is one such tool. Moderation can take many forms, including teachers cross marking each other’s assessments, or a competent external organisation systematically checking school-based marking, for instance. Effective moderation processes must strike a delicate balance between rigour and manageability, in order not to overload the assessment system (OECD, 2013^[4]). Examination agencies already have experience of handling and moderating school-based examination components such as:

- Research Projects, including for example, projects about local history, geography fieldtrips;
- Extended assignments (e.g. *International Baccalaureate*, *Scottish Baccalaureate* (see Section 4.2.4));
- Teacher Assessment and teacher set examinations (e.g. Queensland (see Section 4.2.5), Norway (see Section 4.2.2), Hong Kong, China (see Section 4.3.2));
- Portfolio assessment (e.g. preparatory sketches in Art and Design subjects);
- Performance assessment and direct assessment of skills themselves (e.g. drama and music performance, science practical tasks);
- Practical skills, for example those found in vocational qualifications.

Teacher-based assessments can address a broader range of skills, and curricular reforms have generally encouraged teachers to assess these broader skills. This is the case for non-examination years in British systems. In Scotland, teachers are responsible for assessments from Primary to National 4 and teachers’ coursework marks contribute to most

subjects in National 5. England also relies on teachers' assessment at Key Stage 3 (secondary years 7-9, 11-14 year-olds).

In England's accountability system, secondary schools are judged on the basis of their examination results (see Section 4.3.3). As a consequence, teacher-assessed coursework has been largely removed from the examination system, the policy rationale being that teachers may inflate grades to improve outcomes for their schools.

In other systems, teachers continue to be largely responsible for assessing students throughout their secondary schooling. This may be in the form of *continuous assessment* in which teachers monitor and assess everyday performance. This is the *Grade Point Average* (GPA) found in North America, including Ontario from our sample. In New Zealand, Norway and Queensland teachers are trusted to provide the major mark contribution to secondary school qualifications.

The historical ebb and flow of the role of the teacher-assessed coursework which contributes to qualifications is reviewed later [Section 4.3.2]. While it offers better alignment with curriculum developments such as *Curriculum for Excellence*, the demands of school accountability found in England, of 'fair' selection in Hong Kong, China, and the workload demands on teachers in Scotland and Ireland have all had an impact on the degree to which school-based assessment contributes to qualification systems.

3.4. Changes to the school leaving age

The incremental raising of the school leaving age is a further example of an educational reform which affects examination systems. In England students can leave school at 16 but must stay in full time education, apprenticeship, or be in part-time education accompanied by 20 hours a week of work or volunteering. In Scotland, while students can still leave at 16, 88% of students continued in education or training in 2018/19 (Scottish Government, 2021_[10]). The United Kingdom saw the largest increase (5%) in staying-on rates in OECD countries between 2010 and 2018 (OECD, 2020_[7]). In Ireland 90% of students stayed in education in 2015, having taken the Junior Certificate (formerly the Junior Leaving Certificate) at 16 (MacPhail, Halbert and O'Neill, 2018_[35]).

High completion of upper-secondary education is a shared policy aim across all OECD countries. This raises questions about the nature and purpose of national examinations at age 16 and the message they send. If they are intended to certify the successful completion of the curriculum in the first five years of secondary school education, a curriculum which now involves a wider range of skills, are there more valid ways of assessing educational progress? Are traditional single-subject examinations outdated at this stage? The original function of examinations at 16 was to certify the completion of compulsory schooling for school leavers, and to select the minority who would progress further. These examinations now have more of a 'staging post' or 'practice run' function with most students then preparing for further qualifications which are used in selection for Higher Education or training.

National testing at 16 is a feature of British systems, which is found in relatively few other assessment systems (scarce examples include mathematics and Japanese in Japan; the Junior High School Diploma in Chinese Taipei; and the National Examination Certificate in Poland (Suto and Oates, 2021_[36])). In the United States 26 states and territories have a minimum leaving age of 18, with a further 11 states set at 17, with graduation certification usually at 18. In Canada, Ontario has a school leaving age of 18. In France, where the end of compulsory education is 16, 90% of 17-year-olds remain in education with the Baccalaureate leaving qualification taken at 17-18 years (Jeantheau and Johnson, 2019_[37]).

There are national examinations at age 15 in Norway, when students graduate and move from lower-secondary schools to different upper-secondary ones. However, a student takes only one national external examination and one local oral or practical examination, marked by the teacher. Marks in other subjects are determined by teacher assessment. The examination results will influence the route taken in upper-secondary education and training and are carried forward into the final School Leaving Certificate (Tveit, 2013^[38]).

The French *brevet (Diplôme national du brevet)* is awarded at the end of lower-secondary (*collège*) at 15 and is based on equal contributions from teacher assessment and examinations. The examination component incorporates an oral test based on an inter-disciplinary project by the student which accounts for 100 of the brevet's 800 marks. It is not necessary to pass it in order to continue to upper-secondary education (Jeantheau and Johnson, 2019^[37]).

3.5. Student diversity

Education now meets the needs of a far more varied range of students within the same school or college. Scotland's inclusive practices were demonstrated by its early move away from selective to comprehensive schools. A typical classroom will have students from a variety of cultural backgrounds as well as a range of interests and attainments. Increasing cultural diversity has led to more sensitivities around the curriculum. What should be taught in history or literature? The *Black Lives Matter* movement in the United Kingdom has focused attention on Britain's colonial heritage and how it should be understood. Russia in the Gorbachev era, and post-Apartheid South Africa both had to suspend their history syllabuses while history was "re-written".

This diversity also raises issues of fairness and equity in assessment. The historic, and contemporary, appeal of examinations is that they are the fairest way of assessing students. Candidates get the same questions under the same conditions and are marked using the same mark scheme. Two challenges here are whether there has been equal access to assessment, do some have a more privileged preparation, and is the format of the assessment, for example timed written responses, the fairest way of capturing someone's understanding and skills? (Stobart, 2005^[39]). The use of more varied formats, for example school-based assessments and practical work in vocational qualifications, represent ways of making qualifications more fit-for-purpose for a more diverse candidature (see Section 6.1).

Those students, the majority in England, who 70 years ago would have left school at 15 without taking any examinations are now staying in education and training until 18. One response is to offer a broader range of courses and qualifications, with increased vocational offers in occupationally related areas, for example Hospitality and Catering, Sport and Leisure, or Business. Countries such as Norway offer a strong vocational track at 16 with over half of students taking this route. This involves two years of schooling in one of nine vocational programmes followed by two years of apprenticeship or a third year of schooling (Tveit, 2013^[38]). Routes in the Netherlands and Germany are reflected in the division, at age 11, of the school system into general and vocational pathways. In New Zealand vocational courses can contribute, along with academic ones, to the upper-secondary National Certificate of Educational Achievement. In Ireland, vocational preparation, general education, and vocational education are integrated into the national Leaving Certificate (see Section 4.1).

The French *Baccalauréat*, a highly centralised qualification, has progressively broadened its scope to include many more students. This has included incorporating vocational and professional baccalauréats alongside the traditional academic tracks. In 2018, 39% of the

student cohort enrolled in upper-secondary education were enrolled in a vocational or professional track (OECD, 2020^[7]), and over 80% of the candidates to all *baccalauréats* successfully complete the qualification every year. The *Baccalauréat* is currently undergoing further reform, including a role for teacher assessment (see Section 4.2.3).

Within British educational systems, these routes are not as clearly identified, particularly up to age 16, and are more likely to be taken as individual ‘applied’ qualifications. There is a legacy in these systems of vocational awards being seen, and treated, as less prestigious, despite policy announcements about parity. In England, the General National Vocational Qualification (GNVQ) was given parity with the GCSEs and Advanced Levels (A Levels) only to be absorbed as ‘Applied’ GCSEs and A Levels and made more ‘academic’. In Scotland, there are vocationally oriented courses within the mainstream National 4 and 5, Highers and Advanced Highers, though they have a relatively low take-up. For example, Business Management National 5 had an entry of 7 576 in 2019 (total N5 entries 288 552). Other separate vocational courses, National Certificates and National Progression Awards made 7 061 awards. In 2019 a range of vocational awards (SCQF 1-6) were awarded to over 44 000 students (Scottish Government, 2021^[10]).

4. Student assessments in context: the impact of historical legacies

Contemporary examination systems are the product of their histories, with a basic architecture that may go back centuries. Husen (1967^[40]) argued that any educational system can only be fully understood in the context of the culture, traditions, history and general social structure of the nation it is designed to serve. These legacies may constrain reform, even where curriculum and pedagogy may have seen substantial change. In Scotland, the current system of Nationals and Highers have evolved from their 19th and 20th century precursors, with their origins in the 1888 Scottish Leaving Certificate. This brief review of different assessment systems is intended as a reminder that there is a variety of ways in which secondary school students can be assessed. These are summarised in Table 4.1.

4.1. British examination systems

The legacy for Scotland is that of the ‘British’ examinations system. Education has been distinctively Scottish since the 1707 Act of Union and made a devolved power of the Scottish Parliament in 1999. However, Scotland’s upper-secondary school assessment policies sit within a broad ‘British’ tradition of student assessment that features public examinations at key points in secondary schooling. This is also the case for other countries in the British Isles as well as, for example, Ireland and for legacy countries such as West and East Africa, Hong Kong, China, and Singapore.

Many of the distinctive features of the Scottish examination system have their origins in the structure of secondary schooling, with its transfer to higher education at age 17, and in continuous reforms to offer a more coherent progression through the different levels of examination. Recent reforms, which built on the 1977 Munn and Dunning reports, the Howie Report (1992) and the Government’s own ‘Great Debate’ in 2002, sought to reduce the examination load in the last three years of secondary schooling (McVittie, 2008^[41]). The move to new National Qualifications in 2014 sought to better align the examinations with the goals of CfE.

This does not imply uniformity in how examinations are perceived and used within this historical tradition. Commentators have emphasised the difference in approach between Scotland, with its emphasis on inclusion, and England (United Kingdom) with its emphasis

on schools as a competitive market. Arnott and Ozga (2016^[42]) in their *Education and nationalism in Scotland: governing a 'learning nation'* observe that:

“People in Scotland are offered the opportunity to identify as inheritors of a tradition that values fairness and inclusivity, while also achieving academic excellence, combined, since the introduction of the Curriculum for Excellence, with a judicious dose of personal and practical development.” (p.256).

By contrast, England’s Education Reform Act of 1988 set a new direction for policy, ‘embracing a market-driven approach, greater competitiveness, and “choice and diversity among schools”’. These policies have also led to the marginalisation of local authorities and of democratically elected local government with an increase in direct government control (Furlong and Lunt, 2016^[43]).

Historically, the key purposes for these assessments have been for selection and certification, and for setting curriculum-based standards. To be successful students needed to have demonstrated the level of knowledge required by the syllabus. A more recent emphasis in this tradition has been the increased use of results for school and system accountability, though there were antecedents for this in the 19th century ‘Payment by Results’ scheme that operated in England for over 30 years.

While the origins of the written examination are found in the Chinese civil service selection examinations which were in place for over a thousand years (Stobart, 2008^[2]), the format of ‘British’ school assessments can be traced back to the 19th century. The Victorians’ enthusiasm for examinations came from the rapidly expanding professions, which saw them as a fairer way of selecting entrants for training, and for certification. Because examinations were used in relation to high-status occupations such as medicine, the model of the written, theoretical test was invested with similar status in education (Broadfoot, 1979^[44]). The formal written exam has progressively been given priority over alternative approaches such as the more applied and practical.

In British education, regular testing was introduced in universities as early as the 18th century to monitor and improve standards. This approach then filtered down to schools, and by the mid-19th century, university boards were preparing examinations for schools, particularly the private schools for the middle classes (Stobart, 2008^[2]). Increasingly the state took control of the examining process, either directly by government agencies, as in Scotland, or through direct regulation of examination providers, as in England. In England, the university boards were progressively merged as a result of political pressures, so that only three GCSE and A Level examination boards remain.

Historically these examinations then shaped, or became, the school curriculum and how it was taught. A legacy of this tradition is the dominance of the examination system as the *de facto* curriculum for those in upper-secondary school. In England, the National Curriculum originally continued up to the end of Key Stage 4 in Year 11 (age 16) but soon defaulted to ending at Year 9 (age 14) with GCSE preparation becoming Key Stage 4. While systems such as England, Scotland and Wales have each developed a national curriculum, it is the examination syllabus or specifications that dominate teaching and learning during the upper-secondary years.

The evolution of current assessment systems shows a series of incremental changes in response to social and educational changes. The governance of examinations has shifted from universities to national bodies linked to government. This trend is partly the result of selection and accountability pressures which require comparable standards between those who offer examinations. The Scotland Qualifications Authority (SQA), a non-departmental government body, performs this role in Scotland, with the Office of Qualifications (Ofqual) ensuring comparable standards across the different awarding bodies in England. In Wales,

examinations are provided by its Welsh Joint Education Committee (WJEC), while in Ireland examinations are the direct responsibility of the State Examinations Commission.

A legacy of the British system is that of examinations being taken at age 16. The origins of examinations for 16 year-olds were in providing qualifications for school leavers – then the majority of students. Students in most developed countries are now expected to stay in education or training until 18. This is the point of formal assessment in countries such as France, North America and Australia. Outside the British tradition there are few examination systems that now test nationally at 16.

Contemporary features within the British tradition typically include:

1. Examinations are produced or regulated by a state agency, they are ‘nationalised’;
2. These are subject-based, with students taking a range of subjects, typically six or more at 16 and fewer at 18;
3. They are curriculum-related and expressed in term of a syllabus/specification;
4. Moderated teacher assessments of their students may also contribute to the final subject grade, though the weightings given to this can vary considerably;
5. Examinations are taken under standardised conditions at fixed times, often only once or twice a year;
6. Candidates’ responses are generally open-ended and marked by external examiners;
7. There is little or no pre-testing of the examinations, therefore setting grade boundaries has to take account of the relative difficulty of the papers; grade boundaries may change from year to year;
8. Grading is based on overall mark totals and, unlike some vocational qualifications, there are rarely hurdles within a subject that must be met in order to pass;
9. Grade distributions are described as standards-referenced (if there are improvements in standards more will get higher grades), though there are underlying normative assumptions (grade distributions should not change much from year to year).

Ireland has partially moved away from this framework by developing a modular (44 modules) and credit based (200 credits) system, with qualifications available in three curriculum areas (vocational preparation, general education, and vocational education). The Leaving Certificate Vocational Programme (LCVP) is integrated with the national Leaving Certificate and involves joint modules. The assessment system remains centralised, though there are credits for the satisfactory completion of modules and, unusually, 90% attendance. Moderated school-based coursework and practicals contribute to the credit system.

4.2. Alternative legacy traditions

These ‘British’ approaches to secondary school assessment can be contrasted with other assessment traditions. The intention in reviewing other traditions is to illustrate how the forms of secondary school student assessment reflect cultural differences in school assessment. Those embedded within a particular tradition may think that this is how assessment must operate. The position taken in this paper is that examining is essentially a social process. What is assessed, and how, is a cultural product and changes to assessment systems largely reflect social changes and demands.

4.2.1. American traditions

A widely adopted alternative tradition stems from the more psychometrically and reportedly objective (multiple-choice) focused approaches pioneered in the United States. These are based on statistical models which involve pre-testing, scaling of items for difficulty and machine marking (Baird et al., 2017^[45]). The administrative simplicity of this form of testing and the speed of marking have led to a world-wide uptake in education and many occupational spheres.

In American education, tests such as SATs (Princeton Review, 2021^[46]) are commercial products and operate in a market-based system with commercial alternatives such as the American College Test (ACT) (ACT, 2021^[47]). These are driven by college entrance requirements, they are not federally mandated, and there is no government oversight or accountability (Opposs et al., 2020^[48]). There are state-level tests, which may be in-house or commercially produced, and which may contribute to school graduation requirements. SATs are generally taken privately outside school; schools are not expected to prepare students for them (but they do) and they can be regularly re-taken. SATs are not directly curriculum related, they rather seek to assess how well one analyses and solves problems (CollegeBoard, 2021^[49]).

However, there is a move toward more curriculum related tests such as the Advanced Placement Tests which lead to college course credits (Morgan, 2018^[50]). The Advanced Placement Program allows students to take college-level courses while still in upper-secondary education, each course giving access to a standardised college-level assessment called “AP Exam”.

An important feature of the US system is that test scores are only one strand in the selection processes used by colleges and universities, one that is increasingly optional as a result of concerns about the fairness of the SAT in relation to disadvantaged students. Teachers’ continuous assessment (the Grade Point Average), teacher recommendations, class ranking and students’ activities all contribute. Schools issue graduation diplomas for students when students leave at 18 or if they complete their courses earlier.

In this paper, secondary school assessment in the Canadian province of Ontario offers an instructive assessment model that sits broadly within the American tradition. While education in Canada is devolved to the provincial level, the assessment system in upper-secondary schools in Ontario share many common features with the other ten provinces and three territories (DeLuca et al., 2017^[51]). It is related to the American system in using the award of credits for courses, which are combined into a graduation diploma for 18-year-olds at Grade 12. Teachers’ continuous assessments contribute the bulk of the marks in this process, 70% in Ontario’s case. Where it departs from much of the American tradition is in its limited reliance on psychometric tests (teachers generally construct their own tests) and its low-stakes school accountability systems.

Upper-secondary school students generally follow a programme leading to the *Ontario Secondary School Diploma* (OSSD). The goal in the government’s *Achieving Excellence* (2014) was for 85% of students to graduate. This involves achieving 30 credits, while students gaining over 14 credits can leave school with the *Ontario Secondary School Certificate* (OSSC). A credit is awarded for a mark of over 50% on a course. Of the 30 credits, 18 are in compulsory subjects (English, Mathematics, Science, Health and Physical Education, Career studies or Civics), while the other 12 come from a range of options. Students also need to complete 40 hours of community involvement activities and are required to pass the Ontario Secondary School Literacy Test (OSSLT). It is a school’s choice whether the OSSLT marks contribute towards the Diploma (Ontario Ministry of Education, 2010^[52]).

4.2.2. Nordic systems

School examinations were introduced in the early 19th century in the Scandinavian countries. Like the British system they were initially overseen by universities. Over the rest of the century there was a struggle for control in Norway, with teachers taking control of the examination system in 1884. This was an early sign of teachers being at the centre of the examination system, a feature which is still present, though with more central regulation of the system. The inclusive social values and traditions of Scandinavian countries have minimised the use of examinations in compulsory education. Instead, teachers are given ‘high-trust’ roles in assessing their students.

The debate around testing has often been an ideological one, with tests seen as competitive and divisive in a culture where co-operation and support are central values. This has led in Norway to the banning of testing and grading of pupils before age 11. There has been some softening of this position with the introduction of diagnostic ‘mapping tests’ to identify pupils with learning difficulties and of revised national tests (after initial opposition). The national tests are taken at the beginning of Years 5 and 8 in English, Reading and Mathematics and Reading and Mathematics in Year 9. They are intended to provide formative information for teachers. There are no national school ‘league tables’ as in England and Australia.

In upper-secondary education, assessment is still predominantly the responsibility of the teacher, who contributes over 80% of the overall marks. The only cohort-based central examination is in Norwegian. In other subjects, both centrally and locally (often practical or oral) devised examinations are administered to a sample of students chosen through the ‘exam lottery’ (*trekkordningen*). The exam lottery is meant so that students shall not be examined in every subject but shall instead prepare for exams in the subjects for which the exam is a possible final assessment in addition to the marks awarded for classwork. This involves around 20% of students in a subject, with all students in both Upper-Secondary 1 and 2 taking one such exam and in Upper-Secondary 3 taking two as well as one oral, practical, or combined practical and oral exam. Each of these contributes to the final certificate. Students find out which subject they may be examined in only a day before the examination. Schools run multiple examinations based on their lottery allocations. The fairness of this ‘lottery’ approach has been widely criticised by the School Student Union of Norway, as has the stresses associated with the 24 hour advanced notice of which subject is to be taken.

4.2.3. Baccalaureate systems

The French *Baccalauréat* is a widely recognised historical system on which other Baccalaureate qualifications have been, often loosely, based. From its small-scale beginning in Napoleonic times as an elitist university entrance qualification, the *Baccalauréat général* has continuously developed in terms of scale and cost. The relatively recent introduction of new streams (*voies*), the *Baccalauréat technologique* in 1968 and *Baccalauréat professionnel* in 1985, represent politically driven reforms aimed at reducing social inequality by widening access to this nationally respected qualification. The Orientation Law of 1989 stipulated that 80% of 18-year-olds should reach *Baccalauréat* before 2000. By 2016, with two thirds of the 18-year-old cohort taking it, the pass rate was 88.6%. The significance of this is that success in the qualification gives the right to enter any university regardless of specialisation (Gautier, 2018_[53]). It was found, however, that the right of access is problematic, given 30% of university students drop out during the first year (Jeantheau and Johnson, 2019_[37]).

Despite this high-stakes function, there appears to be far less social and political interest in issues such as comparability, either over time or across streams, than in other jurisdictions. Marking standards similarly attract little research or comment, even though they are highly regionalised. The pass mark in every subject, every year, is 10 out of 20, and the qualification is awarded when the overall average is 10 or above. An averaged mark of 8 or 9 allows a candidate to take two resit tests or orals in order to reach the required 10. Grades (*mentions*) are also predetermined, 12 for an *assez bien*, 14 for *bien*, 16 for *très bien*, and 18 for *félicitations du jury*. Roger-Francois Gautier points out that:

“the standards are not known, the examination is not independently evaluated, and nobody seems to care” (Gautier, 2018, p. 126_[53]).

After limited change over many years, a major revision was planned to take place in 2021. The new academic diploma involves 10 specialist subjects. Students choose three of these in their second year of upper-secondary and just two of these in their third year. In the second year there will also be one written and one oral examination in French literature and culture. In their final year students will take examinations in Philosophy and their two specialist subjects. There will also be a 20 minute oral based on a project begun in the previous year.

The use of continuous assessments based on teacher and test-based judgements continues. These contribute 40% to the final grade, the pass mark remaining at 10 out of 20. The test content that teachers use in their tests is to be drawn from a computerised bank of nationally validated questions and will be marked by other teachers.

4.2.4. *The International Baccalaureate (IB)*

Better known outside the sphere of French education, this internationally recognised and high-status qualification resembles the *Baccalauréat* in incorporating a mix of subjects integrated through philosophical elements. The Diploma taken by 16-18 year-olds has three core requirements: the *Extended Essay*, an independent research essay of 4 000 words; *Theory of Knowledge (TOK)* assessed through an externally assessed essay of 1 600 words, plus an oral presentation; and *Creativity, activity and service (CAS)*. Unlike the *Baccalauréat*, it has, from the outset, included a significant role for teacher assessment. The examination and moderation procedures have much in common with the British model.

The International Baccalaureate resulted from requests for a qualification for private schools that were educating international and expatriate students. These required a broader and more portable qualification system than national ones. Beginning in Geneva in the 1960s, a seven-year pilot phase operated until 1970. The two-year Diploma Programme for 16-18 year-olds now operates in 140 countries and economies around the world and can be taken in English, French and, since 1983, Spanish.

There are Baccalaureate qualifications offered separately in Scotland, England, and Wales. In Scotland they are offered within four frameworks: Expressive Arts; Languages; Science; and Social Sciences. Their distinguishing feature is an Inter-disciplinary Project that seeks to help candidates develop and show evidence of initiative, responsibility, and independent working. 157 candidates were awarded the qualification in 2019.

The *Welsh Bacc*, introduced in 2015, is based on specified combinations of national qualifications plus a *Skills Challenge Certificate* which requires an individual project plus evidence of meeting enterprise, citizenship and community challenges.

In England, the *EBacc* is essentially an accountability device which directs schools into which subjects they should be offering at the General Certificate of Secondary Education (GCSE) taken by 16-year-olds (thus narrowing a school’s curriculum offer). There are no

other ‘broadening’ requirements outside the single subject GCSEs. Only those listed will be considered in the *EBacc* performance tables. This has largely been replaced by *Progress 8* measures, which serves a similar control function (UK Department for Education, 2019_[54]).

4.2.5. *The Queensland system*

Queensland, Australia, provides a well-established model of a teacher-driven examination system at the completion of upper-secondary education. In 1971 Queensland transitioned from traditional university-set and externally marked examinations to school-based assessments. These were based on the continuous assessment of student work, which was then moderated by panels of teachers (‘consensus moderation’) rather than by statistical moderation (Maxwell, 2010_[55]). The central role of teachers in grading was seen as a powerful form of professional development of assessment literacy. This, for many years, has provided a model of teacher professionalism in assessment and of how assessment can be directly linked to the curriculum.

Pressure from politicians and advocacy groups composed of university science and mathematics specialists led to a review of the system. Reporting in 2014 *Matters and Masters’* review (2014_[56]) proposed a shift towards introducing some external assessment, a change which was more the ‘result of advocacy rather than public dissatisfaction’ (Cumming, 2019, p. 171_[57]). This has led to changes in which, from 2020, upper-secondary students will complete four pieces of assessment in a subject, three internal and one external. It appears there was little teacher enthusiasm for these changes, with *Matters and Masters* commenting that their consultation responses showed “acceptance and resignation ... [that] the time had come for change” (*Matters and Masters*, 2014_[56]).

4.2.6. *New Zealand’s modular approach*

When New Zealand was facing an economic downturn and labour market crisis in the late 1960s, the government looked to reform the vocational education and training sector. This had been under-resourced and there was an imperative to improve young people’s skill levels. One consequence of limited work opportunities was that more students were staying on in school, the upper-secondary stage of which was dominated by the academic School Certificate Examination. Led by the secondary teachers’ union, and in collaboration with the Department of Education, there was a call for a radical overhaul of curriculum and assessment in upper-secondary school. Along with many other countries at the time, a more centralised and outcomes-based approach was adopted.

This was the background which led New Zealand to look to Scotland’s vocational provision. This too had developed out of concerns over school drop-out rates and high youth unemployment in the late 1970s. As a result of influential reports such as the Munn and Dunning reports published in 1977, and the *16-18s in Scotland: An Action Plan* published in 1983 (McVittie, 2008_[41]), a modular approach, which led to a National Certificate, was adopted. This was administered by the Scottish Vocational Education Council (SCOTVEC). The National Certificate provided a national and portable vocational qualification.

As Lee, Lee, and Openshaw (2013_[58]) observed:

“For New Zealand’s politicians and educationalists, the apparent success of a small society like Scotland’s in developing a viable workplace qualification that eased the difficult transitions between school and work was seen as highly relevant and thus worthy of emulation, even improvement.” (Lee, Lee and Openshaw, 2013, p. 34_[58]).

The significance of this policy adoption was that New Zealand developed a modular approach and later extended it to general qualifications as well. This incorporated the abolition of the University Entrance Examination in 1986, followed by the removal of the School Certificate Examination 15 years later.

As with the 2014 new National Qualifications in Scotland, there has been mounting concern about the manageability of the National Certificates of Educational Assessment (NCEA) modular curriculum and its assessment. This was partly because of the increasing complexity of the award, as new modules were added and complex rules of transfer from one level to another introduced. A package of reforms was approved in 2020, which involved seven changes to the NCEA (New Zealand Government, 2019^[59]). These included meeting separate literacy and numeracy requirements; having fewer, larger modules; keeping the Level 1 award optional (it is the highest level achieved for 10% of students); and, restricting the carrying over of credits to the next level.

These changes will mean that students will typically take up to six subjects, from which they will need 60 credits for a Level 2 award. Half the credits will be from external assessments, which have been broadened to include portfolios, reports, performances or common assessment tasks.

4.2.7. A summary of student examination systems

Table 4.1 summarises the characteristics of student examination systems in the nine jurisdictions used for comparison in this paper. A broader overview of the variety in OECD countries' choices regarding national assessments at upper-secondary level can be found in Table 4.2 and Table 4.3.

Table 4.1. Summary of student examination systems

Jurisdiction	Legacy system	Suite of national exams by age	External exam system	External Assessment formats	Contribution of school-based assessments
England	British	16, 18	National, terminal written exams – fixed dates Graded on total marks	Mainly written	None in main subjects
France	Baccalauréat	18	National, terminal, written exams – fixed dates Fixed pass marks (10/20)	Written, oral	40% (inc. tests based on item bank)
Ireland	British	16, 18	National, terminal, written exams – fixed dates Credit based	7 written exams 7 student tasks	Credits for module completion and attendance
New Zealand	British – now independent	17-18	External module assessments – fixed dates	Exams Portfolios Reports Performances	50% teacher assessment
Norway	Nordic	15, 18 (only 1 or 2 exams)	National – only taken by a sample of students	National and local exams (including practicals and orals)	Mainly teacher assessed – including continuous assessment
Ontario (Canada)	N. American	Not applicable	No province-wide subject exams. Credit based Provincial literacy and numeracy tests	No external subject exams	Teacher assessed, inc. school tests (30%)

Jurisdiction	Legacy system	Suite of national exams by age	External exam system	External Assessment formats	Contribution of school-based assessments
Queensland (Australia)	British – now Independent	18 (teacher assessed, locally moderated)	From 2020 – statewide exam – worth 25% of total marks	Written exams introduced (25%) in 2020	75% based on coursework – teachers mark, teacher moderated
Scotland	British	14-16; 15-17; 16-18 (students' progression rates vary)	National, timed, written exams – flexible age of entry, fixed exam dates	Mainly written	Coursework tasks (varying by subject teacher assessed, externally moderated)
Wales	British	16, 18	National, timed, written exams – fixed dates	Mainly written	Coursework tasks (around 20%)
IB (International Baccalaureate)	Baccalaureate	17-18	Timed, written exams – fixed dates	Written, oral, practicals, projects	20-50% in subjects; internal assessment of core elements

Table 4.2. National/central assessments at the upper-secondary level (2015)

In general programmes

	Notes	Existence (1)	Standardised at the central (C), state (S), regional (R), school (SC) or other (O) level of government (2)	Locus of authority at which they are designed/developed (3)	Locus of authority at which they are graded/marked (4)	Mechanisms to ensure reliability of marking across students if marking undertaken at the school level (5)	Based on norm-reference (N) or criterion-reference (C) test (6)	Year first established (7)	Compulsory for schools to administer		Percentage that administer them		Percentage of students exempted from taking them (%) (12)	Subjects covered: Mathematics (MAT), Natural sciences (NS), Reading, writing and literature (RWL), Social studies (SS), Other languages (OL), Physical education and health (PE), Information & communication technology (ICT), Technology (TEQ), Arts (ART), Religion/Ethics/Moral education (REL), Practical and vocational skills (PVS) or Other subject (OTH) (13)
									Public (8)	Government-dependent private (9)	Public (10)	Government-dependent private (11)		
OECD														
Australia		No	a	a	a	a	a	a	a	a	a	a	a	a
Austria		No	a	a	a	a	a	a	a	a	a	a	a	a
Belgium (Fl.)		Yes	S	3,11	3,11	a	C	2010	Sample*	Sample*	5	x(10)	m	MAT,RWL,SS,OL
Belgium (Fr.)		Yes	S	3	11	G	C	2009	All	All	1	1	0	MAT,NS,RWL
Canada		No	a	a	a	a	a	a	a	a	a	a	a	a
Chile		Yes	C	2	2	a	C	2011	All	All	1	1	0	MAT,RWL,OL,ICT
Czech Republic		Yes	C	2	2	a	C	2011/12	Sample	Sample	5	5	2	MAT,RWL,OL,NS,OL
Denmark		No	a	a	a	a	a	a	a	a	a	a	a	a
England		No	a	a	a	a	a	a	a	a	a	a	a	a
Estonia		No	a	a	a	a	a	a	a	a	a	a	a	a
Finland		No	a	a	a	a	a	a	a	a	a	a	a	a
France		No	a	a	a	a	a	a	a	a	a	a	a	a
Germany	1 2	No	a	a	a	a	a	a	a	a	a	a	a	a
Greece		No	a	a	a	a	a	a	a	a	a	a	a	a
Hungary		Yes	C	1	1	a	C	2001	All	All	1	1	0	MAT,RWL
Iceland		No	a	a	a	a	a	a	a	a	a	a	a	a
Ireland		No	a	a	a	a	a	a	a	a	a	a	a	a
Israel		No	a	a	a	a	a	a	a	a	a	a	a	a
Italy		Yes	C	2	2	a	N	2010	All	a	1	a	0	MAT,RWL
Japan		No	a	a	a	a	a	a	a	a	a	a	a	a
Korea		Yes	C	1,2	1,2	a	C	1986	All	All	1	1	m	MAT,RWL,OL
Luxembourg		No	a	a	a	a	a	a	a	a	a	a	a	a
Mexico		Yes	C	1	1	a	C	2008	All	a	1	a	0	MAT,RWL
Netherlands		No	a	a	a	a	a	a	a	a	a	a	a	a
New Zealand		Yes	C	1,2	2,7,8,9	G,M	C	2002	All	a	1	a	0	MAT,NS,RWL,SS,OL,PE,ICT,TECART,REL,PVS,OTH
Norway		Yes	C	1	1,7	G	N	2009,2010,2011	All	All	1	1	2	MAT,RWL,OL
Poland		No	a	a	a	a	a	a	a	a	a	a	a	a
Portugal		No	a	a	a	a	a	a	a	a	a	a	a	a
Scotland		No	a	a	a	a	a	a	a	a	a	a	a	a
Slovak Republic		No	a	a	a	a	a	a	a	a	a	a	a	a
Slovenia		No	a	a	a	a	a	a	a	a	a	a	a	a
Spain		No	a	a	a	a	a	a	a	a	a	a	a	a
Sweden		Yes	C	2	7,8,9	G	C	1996	All	All	1	1	m	MAT,RWL,OL
Switzerland		No	a	a	a	a	a	a	a	a	a	a	a	a
Turkey		No	a	a	a	a	a	a	a	a	a	a	a	a
United States		Yes	C	1,2	1,2	a	C	1969	Sample	a	5	a	a	MAT,NS,RWL,SS,ART

Note: Federal states or countries with highly decentralised school systems may have different regulations in states, provinces or regions. Please refer to Annex 3 for additional information. Please refer to the Reader's Guide for information concerning symbols for missing data and abbreviations.

1. National Assessment Study (Ländervergleich).

2. State-wide comparison tests (VERA: Vergleichsarbeit).

Source: OECD (2015^[60]) *Education at a glance 2015*, <https://doi.org/10.1787/eag-2015-en>. See Annex 3 for notes (www.oecd.org/education/education-at-a-glance-19991487.htm).

Table 4.3. Main purposes and features used for reporting results of national/central assessments at the upper-secondary level (2015)

In general programmes

	Main purposes or uses							Features used when reporting results						
	Notes	To evaluate school performance	To evaluate teacher performance	To provide teachers with student diagnostic information	To provide formative feedback to parents	To select or group students according to their achievement levels	To actively engage students in their own learning process	Other	Show level of performance for most recent year	Show "value added" or growth in student achievement over two or more years	Context sensitive	Compared with other groups or populations of students	Reported together with other indicators of school quality	Used by authorities external to the school for sanctions or rewards
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
OECD														
Australia		a	a	a	a	a	a	a	a	a	a	a	a	a
Austria		a	a	a	a	a	a	a	a	a	a	a	a	a
Belgium (Fl.)		No	No	No	Yes	No	No	Yes	Yes	No	Yes	Yes	No	No
Belgium (Fr.)		No	No	Yes	No	No	No	No	No	No	Yes	No	No	No
Canada		a	a	a	a	a	a	a	a	a	a	a	a	a
Chile		Yes	No	Yes	Yes	No	No	No	Yes	No	Yes	Yes	No	Yes
Czech Republic		Yes	No	Yes	Yes	No	No	No	Yes	No	No	No	No	No
Denmark		a	a	a	a	a	a	a	a	a	a	a	a	a
England		a	a	a	a	a	a	a	a	a	a	a	a	a
Estonia		a	a	a	a	a	a	a	a	a	a	a	a	a
Finland		a	a	a	a	a	a	a	a	a	a	a	a	a
France		a	a	a	a	a	a	a	a	a	a	a	a	a
Germany	1	a	a	a	a	a	a	a	a	a	a	a	a	a
	2	a	a	a	a	a	a	a	a	a	a	a	a	a
Greece		a	a	a	a	a	a	a	a	a	a	a	a	a
Hungary		No	No	Yes	No	No	No	No	Yes	Yes	Yes	Yes	No	No
Iceland		a	a	a	a	a	a	a	a	a	a	a	a	a
Ireland		a	a	a	a	a	a	a	a	a	a	a	a	a
Israel		a	a	a	a	a	a	a	a	a	a	a	a	a
Italy		Yes	No	Yes	No	No	No	Yes	No	No	Yes	Yes	No	No
Japan		a	a	a	a	a	a	a	a	a	a	a	a	a
Korea		Yes	No	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No
Luxembourg		a	a	a	a	a	a	a	a	a	a	a	a	a
Mexico		No	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No	No
Netherlands		a	a	a	a	a	a	a	a	a	a	a	a	a
New Zealand		No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Norway		No	No	Yes	Yes	No	No	m	a	a	a	a	a	a
Poland		a	a	a	a	a	a	a	a	a	a	a	a	a
Portugal		a	a	a	a	a	a	a	a	a	a	a	a	a
Scotland		a	a	a	a	a	a	a	a	a	a	a	a	a
Slovak Republic		a	a	a	a	a	a	a	a	a	a	a	a	a
Slovenia		a	a	a	a	a	a	a	a	a	a	a	a	a
Spain		a	a	a	a	a	a	a	a	a	a	a	a	a
Sweden		Yes	No	No	No	No	No	Yes	Yes	No	No	Yes	Yes	No
Switzerland		a	a	a	a	a	a	a	a	a	a	a	a	a
Turkey		a	a	a	a	a	a	a	a	a	a	a	a	a
United States		No	No	No	No	No	Yes	Yes	No	Yes	Yes	No	No	No
Partners														
Brazil		Yes	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	No	No
Colombia		a	a	a	a	a	a	a	a	a	a	a	a	a
Latvia		a	a	a	a	a	a	a	a	a	a	a	a	a

Note: Federal states or countries with highly decentralised school systems may have different regulations in states, provinces or regions. Please refer to Annex 3 for additional information. Please refer to the Reader's Guide for information concerning symbols for missing data and abbreviations.

1. National Assessment Study (Ländervergleich).

2. State-wide comparison tests (VERA: Vergleichsarbeit).

Source: OECD (2015⁽⁶⁰⁾) *Education at a glance 2015*, <https://doi.org/10.1787/eag-2015-en>. See Annex 3 for notes (www.oecd.org/education/education-at-a-glance-19991487.htm).

4.3. Inhibiting factors in assessment reform

Assessment systems usually change gradually. The brake on any radical change may, in part, result from a social reluctance to change an established system. When a system has been in place for generations, parents, policy makers and teachers are familiar with it and value it – even when it may no longer be fit-for-purpose.

4.3.1. Teacher resistance

Teachers' co-operation in education reforms is a major factor of their success. Alternatively, their resistance can significantly delay their implementation (Viennet and Pont, 2017^[61]). In 2012, proposed policy-driven changes to the externally examined Irish *Junior Certificate* involved a substantial role for teacher assessment. These met with fierce teacher union resistance on the grounds that their role was to teach, not to examine, their students. The Irish junior cycle reforms are described in some detail because they bear direct relevance to Scotland's introduction of the new National Qualifications in 2012 (see Box 4.1). In contrast, the introduction of national curriculum tests in Norway in 2005 led to strikes by both teachers and students at concerns over how the results would be used, leading to their temporary withdrawal. Concerns about reform may lead to delayed, or revised, implementation. In Hong Kong, China it took ten years to agree *The New Academic Structure* (2005; 2013 see below) with its changes to the academic and assessment systems.

Box 4.1. Reforming the Junior Certificate in Ireland

The Irish Junior Certificate, formerly the Junior Leaving Certificate, is taken by 15-year-olds at the end of the three-year junior secondary school cycle. The results are published mid-September – after the students have already progressed to upper-secondary education (referred to as “the senior cycle” in Ireland). It is, therefore, ‘low-stakes’ for the students as there are limited consequences for them. However, it is still perceived as ‘high-stakes’ by the public, even though over 90% of students now stay in school until 18. It no longer has a ‘gateway’ function and so is largely treated as ‘a dry run, for the Senior Leaving Certificate two years later’ (Looney, 2006^[62]).

The Junior Certificate was introduced in 1989 to achieve greater breadth and balance in the curriculum. A review in 1999 found the curricular goals were not being achieved due to a mismatch between the re-designed and outward looking curriculum and the conventional terminal examination. The teaching-to-the-test and rote learning throughout the junior cycle were highlighted as contributing factors in the disengagement of students. The need for reform was recognised by the then minister of education and a development process set in train. A Framework for Junior Cycle was published in 2012 which included a classroom-based assessment component weighted at 40% of the award, with 60% based on an externally set, but teacher marked, examinations.

This reform met strong opposition from teacher unions. Their opposition took the form of non-cooperation with the introduction of the new junior cycle. This led to negotiations which in 2015 produced a revised Framework. This re-introduced state set and administered examinations. While the Junior Certificate was not re-introduced, the examination results became part of a Profile of Achievement. Reported alongside these were the teacher-assessed components, by this stage reduced to two structured classroom-based assessments (CBA) for each subject, one in the second year and one in the third. The tasks are externally set but assessed by the classroom teacher. After the

second CBA, students complete a written Assessment Task which reflects on their coursework. This is externally marked and weighted at 10%, the examination contributing 90%.

This retrenchment to externally set and marked examinations is an example of the resistance that assessment reform can generate. In this case it was how teachers in Ireland viewed their professional identity and roles. Summatively assessing their own students was not part of this.

Source: Looney (2006_[62]) Assessment in the Republic of Ireland, *Assessment in Education: Principles, Policy and Practice*, Vol. 13/3, pp. 345-353, <https://doi.org/10.1080/09695940601035544> and McPhail, A., Halbert, J., and O'Neill, H., (2018_[35]) The development of assessment policy in Ireland: A story of junior cycle reform, *Assessment in Education: Principles, Policy and Practice*, Vol. 25/3, pp. 310-26.

Scotland's reforms also met resistance, largely because of the assessment load placed on teachers, including the demands of the quality assurance system. The revised 2017 version of Scottish qualifications saw a retrenchment towards a more traditional examination system, with teacher assessment of individual units abandoned, on a phased basis, from National 5, Higher and Advanced Higher, leading to an overall award in which the role of the final examinations was expanded.

4.3.2. Comparability and selection pressures

Lack of change in assessment practices can also come from concerns about the comparability of assessment results between schools, and across exams that award the same levels of qualifications or give access to similar pathways, for instance, into higher education institutions. Hong Kong, China is a Confucian-heritage society that values examinations as central to selection and progress. Examination results are high-stakes for students as there is intense competition for government-funded university places. Both the curriculum and its assessment have seen reform. A significant element of this is *The New Academic Structure for Senior Secondary Educations and Higher Education* (2005; 2013), which changed the secondary education to three years of junior secondary schooling, three years of upper-secondary schooling and then four years of tertiary education. This was a break with the previous five, two (secondary) and three-year (tertiary) approach modelled on England's GCSE and A Level system. This reform took 'over ten years of patient discussion and lobbying to resolve the controversies' (Tong, Lee and Luo, 2020, p. 234_[63]).

For a strongly examination-focused culture such as Hong Kong, China, the introduction of school-based assessment contributing to final grades has been a contentious process. The aim has been to encourage broader study, but the need for comparable results between both schools and subjects has led to demanding, and disputed, moderation processes. This in turn has led to teacher resistance, particularly at the workload involved and concern about the moderation processes which include complex statistical moderation of coursework marks.

In England, with its three examination boards, there has been a longstanding concern with comparability, given university entrance is largely dependent on A Level grades. To ensure grades are comparable between the three awarders, extensive regulatory procedures have been progressively introduced to guarantee they involve similar demand and that marking and grading procedures are the same. The official regulator, the Office of Qualifications (Ofqual), now monitors every aspect of the examination process.

The same issues can arise within an awarding body when more than one syllabus operates in a particular subject. The regulatory response to a potential lack of comparability has been to limit the number of syllabuses an awarding body can offer in a single subject. This has

left no incentive for experimental syllabuses that may align better with recent curricular and educational approaches.

4.3.3. *The influence of accountability systems*

A key element in resistance to change is the role of examination results in the school accountability system, particularly in education systems where government rely heavily on these results for monitoring schools' progress and setting attainment targets (Eggen and Stobart, 2014^[64]). Examinations in high-stakes accountability systems, such as England's, have a dual purpose: they are high stakes for students in terms of progression and selection, and high stakes for schools in terms of being judged by their examination results. This may lead to distortions in the system, which reduce the validity of the assessments as measures of student learning. As such, the jurisdiction's accountability system can be a particularly difficult obstacle to incorporate teacher assessment into students' final results.

Where school accountability is largely predicated on examination results, there is a perception that teachers will inflate their students' coursework results in order to improve a school's results and therefore, its standing. The political logic was that teacher-assessed coursework 'allows teachers to reward themselves'. This has led to a reduction in the teacher assessment components in those 'British' systems with high-stakes accountability.

In England, there has also been a narrowing both of curriculum choice through what 'counts' in the accountability system, the *Progress 8* measure for GCSE which specifies which subjects 'count' in the calculations of school performance. At A Level the prestigious Russell Group of universities lists nine 'facilitating subjects' (for example, English literature, modern languages and mathematics) 'that keep your options open when choosing a degree' (Success at School, 2020, p. 1^[65]). Critical thinking, citizenship studies and general studies A Levels should only be taken as 'extras'. The format of examinations in England has also become more restricted as the result of the 2013 removal of modular examinations and a reversion to final one-off 'linear' examinations.

Scotland has not followed the English model of centralised high-stakes accountability. Policy makers have avoided examination based 'league tables' and opted for a more complex mix of local and central accountability such as the *Insight* comparator. The OECD's 2015 review of Broad General Education in Scotland (OECD, 2015^[66]) argued that governance in school education has two aspects, a vertical hierarchy running from the "minister downwards" and a "horizontal logic" which is about professionals' networks of sharing knowledge and collaborative support. One outcome of this is that teacher-assessed coursework still contributes to final qualification grades, though to a lesser extent.

In other jurisdictions, such as Ontario, Queensland and Norway, where examination results play only a limited role in the accountability systems, there is even more reliance on teacher assessment. These 'high-trust' jurisdictions use teachers' assessments to determine results which directly contribute to a student's qualifications, for example their school graduation award. Assessment is largely decentralised and therefore can be more flexible. In Queensland teachers act as examiners for their own students with their assessments locally moderated. In Ontario (and other Canadian provinces) teachers determine a student's course marks and make summative subject assessments which are the basis of graduation. There is only limited external testing, and this is not used to moderate teachers' assessment. The culture of these jurisdictions is that teachers can be trusted to assess fairly. The lack of checks for teacher bias and unreliability is offset by the fact that student progression is based on a wider range of measures (Ontario), on moderation processes (Queensland) or on local comparisons with examination results of a sample of students (Norway).

5. The COVID-19 ‘stress test’ revealed the fragility of assessment systems

The 2020 and 2021 COVID-19 pandemic has placed unique stresses on examinations systems in those countries where schools closed. An estimated 1.58 billion learners were off school in 2020, which represent 91.3% of total enrolled learners in the world and an unprecedented situation in the history of education. About 63 million primary and secondary teachers were also affected. Where IT infrastructure and stable internet connection existed, teachers then sought to transfer teaching and learning online (UNESCO, 2020_[67]). Those jurisdictions that relied solely on examination results experienced crises as procedures had to be hastily developed to determine grades. This was necessary for university and selection for the start of the university year. The COVID-19 related crises seem to reveal the vulnerability of systems that rely exclusively on the results of examinations taken under standardised conditions and scheduled at a fixed time. By comparison, those systems that draw on multiple forms of evidence were indeed able to adapt more flexibly.

5.1. Approaches to adapt upper-secondary examinations across countries

Examinations in Scotland were cancelled on 19 March 2020, a day before the closure of schools. The initial response in Scotland and England was to use statistical algorithms based on predictions of centre performance to generate results that were comparable to the grade distribution patterns in 2019. Both countries experienced a strong student and media backlash against the perceived unfairness of this quantitative approach for individual students and for disadvantaged schools which were perceived as penalised by the algorithms (Priestley et al., 2020_[12]).

In England, after student protests against the algorithm made media headlines, alternative ways of grading were rapidly sought. Teachers were then asked to award grades to their students. While teachers routinely provide predicted grades as part of university selection, they do not have to provide evidence for them. As there was no longer a coursework component in most GCSE and A Level subjects, the final grades were based on teachers own judgements of students’ classroom work. There was no system in place for any moderation of these grading decisions in England.

In Ireland, the initial response to the lockdown was to move the examinations to 29 July, with students returning to school two weeks before this. This led to a strong student protest, with a poll of 24 000 final-year students showing that 80% favoured the cancellation of the examinations (CNBC, 2020_[68]). In a change of policy, the final-year examinations were then cancelled with grades being calculated by schools and then moderated through a national standardisation process. A chance to take the Leaving Certificate examination was offered in November 2020, with over 2 000 students taking at least one subject.

In France, high school students did not sit the *Baccalauréat* examination in 2020. It was the first time since its introduction under Napoleon in 1808 that the exam did not take place in its traditional form. Even the student protests of May 1968 did not prevent the exam going ahead. In 2020, students received an average score in each subject, calculated from marks given for tests and homework throughout the year (Ofqual, 2020_[69]).

After Scotland’s grading crisis, the government commissioned a review of the 2020 awarding processes. The *Rapid Review of National Qualifications Experience 2020*, published in September 2020, detailed the events around the award of grades. It was critical of the statistical procedures that were initially used, of ignoring warnings about them from April onwards, and of the lack of transparency about them. The failure to engage in

qualitative moderation of the statistical findings at local level was also highlighted (Priestley et al., 2020_[12]).

There was also concern from some respondents that, instead of cancelling coursework which would have contributed to results, greater efforts to assess it could have been made, thus contributing to the final grade and providing a more robust evidence base for estimation (Priestley et al., 2020, p. 17_[12]). A broader finding was:

“This review has found consistent support from all stakeholders (including young people and parents), for a reduced emphasis on terminal examinations as the basis for qualifications. There is widespread support for continuous assessment and its benefits (including the potential for assessments to be used in a more formative way than at present), when teaching to the test – often in highly formulaic ways – seems to be the norm.” (Priestley et al., 2020, p. 44_[12])

In 2020, anticipation of more school closures in 2020/21 and the expected differential degrees of learning loss across different geographical areas and schools, the governments of Scotland and Wales announced examinations would not take place in 2021. Based on the Priestley (2020_[12]) recommendations, the Scottish Government first announced in October that National 5 examinations would be cancelled in 2021. A similar decision was made in December 2020 for Higher and Advanced Higher qualifications. Grades would be based on teacher judgement supported by quality assurance. The Alternative Certification Model (ACM) involved SQA generated online examination questions which teachers marked using SQA mark schemes (SQA, 2020b). There were then local quality assurance and sample based SQA quality assurance processes. The governments of England and Northern Ireland announced that examinations would go ahead in 2021, but then they too had to cancel examinations, the announcements coming as a result of school closures in January 2021. In England, a similar process was set up of awarding body mini-examinations, selected and marked by teachers, which focused on syllabus content that has been covered.

The early signs in England and Scotland are that this has led to considerable workload issues in schools for both teachers and students. In Scotland, it has been claimed that this has led to many smaller exams for every subject, and forced schools to create complicated timetables covering all these fragmented examinations over a five-week period (Deerin, 2021_[70]).

Other education systems were able to adapt their examinations more flexibly. In Ontario, which also experienced school lockdowns, the assessment problems were less acute as the system relies largely on teacher assessment. Provincial testing is not used as exit exams or for high-stakes decisions. Courses typically involve a final teacher-developed examination worth 30% of the grade. When schools were closed teachers typically used grades from pre-closure periods. The pre-closure grades or pass/fail judgements were solely based on teacher classroom assessment (information gathered from direct communications with Christopher DeLuca).

This was much the same in Norway where secondary school examinations were cancelled in 2020. The consequences of the cancellation were not dramatic, as exams count for 20% of the final grade, teachers' assessment counting for 80%. Similarly, the International Baccalaureate examinations were cancelled and instead students were awarded a Diploma or a Course Certificate, which was reportedly “based on the student’s coursework and the established assessment expertise, rigor and quality control already built into the programmes” (International Baccalaureate Organization, 2020_[71]).

In the United States, the summer SAT test was cancelled, leading to many US universities adjusting their admission criteria to focus on the other school-based elements, which are

also used in the selection. In contrast, the Advanced Placement tests went online (see Section 3.2).

In New Zealand, where schools had also been locked down in March 2020, the examinations for the National Certificate of Educational Attainment (NCEA) were delayed for an extra ten days in November. Despite calls for a reduced credit requirement, the qualification was awarded as usual.

5.2. Opportunities for long term change

Adapting Thomas Kuhn's concept of "paradigm shift" in science, Isaacs and Gorgen (2018^[72]) have considered the conditions that may lead to major shifts in a country's examination system. Kuhn had defined a paradigm as something offering a "universally recognised scientific achievement that, for a time, provides model problems and solutions for a community of researchers" (Kuhn, 1962^[73]). Over time anomalies build up and an alternative explanation proves more compelling to enough people for paradigm shift to occur. Isaacs and Gorgen point out that in educational assessment, unlike science, there is no one dominant model. They identified three pre-conditions that are involved in any paradigm shift:

1. There must be dissatisfaction with the currently accepted model (paradigm);
2. There must be an alternative, agreed upon, paradigm that is a better fit;
3. The advocates of the new paradigm must outnumber or outweigh those supporting the old paradigm.

While these are necessary conditions, they may still not lead to radical change. Accommodation within the traditional approach is still more likely. In relation to the first pre-condition, the 'stress test' failures of 2020 have left a deep level of dissatisfaction with those assessment systems dependent on examinations. The discontent was apparent in both media reporting and criticisms of the political handling of the situation. Within England, Scotland and Ireland, active student protests led to extensive media coverage. In England, the various U-turns and short notice announcements have led to widespread media criticism of the secretary of state for education, including from his own regulator, Ofqual (Adams, 2020^[74]).

The alternative that proved a better fit during the pandemic is found in those systems that include a substantial teacher assessment component in their qualifications. Jurisdictions such as Ontario, Queensland and Norway were able to adapt relatively easily to school closures and examination cancellation. This was also the case for the *International Baccalaureate* (IB), which used evidence available from teachers on their students' achievement; externally marked coursework, in each subject already completed; predicted grades from teachers; and evidence on how these two pieces of data usually related to final grades for each school (International Baccalaureate Organization, 2020^[71]).

At this stage, in mid-2021, it is difficult to judge whether the impact of the 2020 crisis, and the preparations for 2021 without examinations, and possible disruptions to the examination cycle in 2022, have created sufficient momentum to develop more resilient assessment systems in the United Kingdom. These systems would be likely to include substantial school-based assessments as part of qualifications.

In England, there has been increasing media advocacy of a switch from examinations to alternative forms of assessment. A *Sunday Times* headline in England on 31 January 2021 read '*Schools seize on 'perfect time' to explore exam-free future*'. The article cited high profile schools that were already dropping examinations at 16 in favour of digital passports

of teenagers' progress. Evidence of changing attitudes can be found in the current advocacy by Lord Baker of Dorking for the GCSEs removal, when it was he, as education minister, who oversaw the introduction of the GCSE in 1986. The *Rethinking Assessment* movement in England is campaigning for the reform of the “mutant exam system” which “is not turning out youngsters who succeed in the 21st century” (The Times, 2020^[75]).

The alternatives developed in the wake of COVID-19 disruptions opened many possibilities to rethink upper-secondary assessment systems. Whether they gain sufficient support and weight during the current examination crises, along with the recognition that Britain is the last country in Europe in which pupils take national examinations at 16, is unclear.

6. What is needed for a dependable assessment system that supports learning?

6.1. Dependable assessments: validity, reliability and manageability

Any assessment system that commands public support has to balance *validity*, *reliability*, and *manageability*. The OECD defines manageability in terms of *transparency* and *usability* (OECD, 2013^[4]). Public confidence is vital if an assessment system is to be effective, while a loss of faith in a system will undermine the status and value of qualifications. If they are perceived as unfair, a validity and reliability issue, or unmanageable (by students and teachers), results will not be trusted. As presented in Section 3, different education systems use different formats and technologies, each with their own curricular and technical trade-offs and each with particular threats to trustworthiness.

Scotland has a cultural tradition which values fairness and inclusivity (Arnott and Ozga, 2016^[42]) and this is reflected in equity concerns in relation to qualifications. These values are reflected in the progression levels through National Qualifications and Highers, in which students can choose to exit the examination system, and in the differential rates at which students can progress through it.

John Bigg's conceptualisation of *constructive alignment* between curricular intention and what happens in practice suggests that any system has to be evaluated for the *fitness-for-purpose* of its instruments – are they assessing what they claim to assess? Frederiksen and Collins' *principles of systemically valid testing* (1989^[31]) is a useful theoretical framework for such evaluation, especially when gauging the validity of examination papers and assessment by teachers which contribute to qualifications, as well as extended assignments and portfolio-based assessment. At the heart of this approach is how the desired curricular skills and understanding can be actively encouraged and reinforced by the format and demands of the test.

Operating within the American system, their concern was that the reliance on multiple-choice testing was not fit-for-purpose because it encouraged inappropriate forms of teaching and learning. They drew on performance-based forms of assessment and proposed standards for assessment, which included:

1. *Directness*. This involves assessing the cognitive skill of interest – the emphasis being on the authenticity of tasks, which are representative of the way knowledge and skills are used in real-world contexts;
2. *Scope*. This considers the range of skills needed to do well in the tasks;
3. *Reliability*. This seeks effective ways of assessing which, at the same time, fosters learning;

4. *Transparency*. This is the concern that those being assessed are clear about how they are being judged. This should enable learners to “assess themselves and others with almost the same reliability as the actual test evaluators achieve” (Frederiksen and Collins, 1989^[31]).

In the American system the trade-offs involve the benefits of manageability and reliability at the cost of the limited validity of what is assessed. When there are no, or few, open-ended (‘constructed’) responses which allow students to demonstrate their own thinking, only limited inferences can be made about students’ thinking and skills. In ‘high trust’ systems such as Norway there is the potential for a broader range of skills to be assessed, giving them more construct validity. There may be concerns, however, about the comparability of assessments between schools, a reliability issue.

In order to broaden the curriculum and students’ skills, Hong Kong, China’s examination reforms introduced more teacher assessment. This was intended to improve the examination’s validity and encourage broader learning. Because of the high-stakes nature of the qualifications, maintaining reliability was seen as essential. This led to complex moderation procedures, which threatened the manageability of the qualifications and teacher goodwill. Queensland also addressed reliability issues by extensive consensus moderation.

Manageability was also an issue for Scotland’s 2012 introduction of the new modular National Qualifications, which led to an assessment overload for teachers. The 2017 revisions of the National Qualifications could be seen as a move to make them more manageable by reducing the teacher assessment demands for each module and increasing the examination weighting. The trade-off in Scotland was a reduction in the range of what would be assessed, a restriction of construct validity.

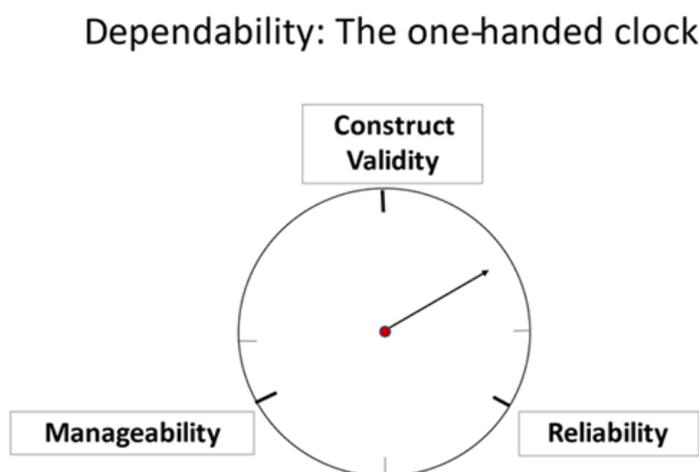
This interplay of validity, reliability and manageability is illustrated in the assessment of English for 11-year-olds in England in the National Curriculum Key Stage 2 tests. In 2011, there was a ‘longer writing’ paper as part of the test which required pupils to demonstrate their ability to construct a piece of writing on a given topic. This aligned with the demands of the national curriculum (construct validity). In 2011, there were protests about the fairness of the external marking (reliability) of the writing paper with 30 000 appeals made by schools. This led to marking being handed over to the teachers in 2012 (increased manageability). This, in turn, led to an increase in scores (a reliability issue) and the paper was dropped from 2013 onwards (a validity issue). English is now tested through a Reading and a Spelling, Punctuation and Grammar (SPaG) paper (easier and more reliable to assess). This reduction of the domain has led to it no longer being called English, because of the lack of construct validity in relation to the national curriculum. While teacher assessment incorporates the fuller curriculum, these assessments are not incorporated into the performance tables by which schools are judged.

Dependability is a useful concept with which to evaluate assessments. A dependable assessment is one we can rely on to give us a trustworthy estimate of students’ capabilities. It involves an optimal trade-off between construct validity, reliability, and manageability. Construct validity involves effectively sampling the domain being assessed. To be dependable, an assessment also needs reliability; grading a skill with no agreed assessment scheme reduces confidence judgements of students’ performances, as do differences in the administration of an assessment. We can achieve a valid and reliable assessment, but it may be at the (unaffordable) cost of manageability.

In Figure 6.1 the hand can only be in one place, illustrating the inevitability of trade-offs. At 10 minutes we have a valid and reliable assessment that may require costly and expensive processes, for example, training an airline pilot; at 50 minutes we may have a valid and manageable teacher assessment that has limited reliability. What is to be avoided

is to be at 30 minutes – a manageable, machine marked (reliable) test which is easy and reliable to assess but bears little relation to the construct, for example, a multiple-choice test of creativity.

Figure 6.1. Dependability: the one-handed clock



Source: Stobart (2008^[2]) *Testing Times: The Uses and Abuses of Assessment* (1st ed.). Routledge. <https://doi.org/10.4324/9780203930502>.

These trade-offs also apply to assessment systems. Where there are a variety of strands contributing to students' outcomes, then these trade-offs will vary and may lead to a more dependable, trustworthy and resilient overall system.

The National Qualifications in Scotland raise a trade-off issue in relation to construct validity: is the 'construct' that the examinations are sampling the 3-18 *Curriculum for Excellence* (CfE)? If so, how validly are the examination specifications sampling this? Is there evidence of the examinations under-representing the curriculum by only partially covering CfE? This is a curriculum debate that needs to precede any examination reform. The OECD report on *Curriculum for Excellence* (2021) offers the stimulus for this and may, in turn, raise the question of whether the current examination format is fit-for-purpose.

6.2. What do students want?

Students' voices, although legitimate and often constructive, can be overlooked in the midst of expert, public, and political deliberations about assessment systems. Some evidence collected for a number of studies, such as Priestley et al. (2020^[12]), the SQA and Young Scot report (SQA and YoungScot Observatory, 2018^[13]) and the OECD report (OECD, 2021^[5]), suggests that students would prefer a system which includes some form of continuous assessment by their teachers. Priestley et al. (2020^[12]) reported from discussions with students that:

“Young people would like to see achievement captured throughout the year, rather than the ‘two term’ dash towards examinations (in particular for Higher)”
(Priestley et al., 2020, p. 39^[12]).

This observation came off the back of the 2020 cancellation of examinations and the use of an algorithm that, in seeking similar overall grade distributions to previous years, ignored individual student performance, something that was widely perceived as unfair. The same happened in England, leading to student protests. In Ireland, there were student objections to delaying the examinations until late July.

These were reminders that it is those being assessed who are most directly affected by an assessment. To attempt to preserve year-on-year patterns of grade distributions ('standards') was a corporate response from the United Kingdom's nationally controlled systems. It was the perceived unfairness to individuals that became the toxic political issue leading to rapid policy changes in how qualifications were awarded in 2020.

Even if the result is low-stakes for the students, it will still impact on their identity as learners. The *Junior Certificate* in Ireland was a low-stakes examination, there were few direct consequences for the student, but its public legacy meant it was perceived as high stakes. This resulted in pressure being placed on the students about their performance. This is also the case with the Key Stage 2 tests at the end of primary school in England. They have no direct impact on secondary school selection, which has already happened, and are therefore low-stakes for the pupils. However, because the accountability system judges schools by their KS2 results, they can have serious consequences for schools. This in turn leads to teachers spending much of Year 6 teaching to the test and pressurising pupils to take them very seriously (Reay and Wiliam, 1999^[76]).

7. Options from Scotland and beyond for the future of student assessment in upper-secondary education

7.1. International examples of alternatives to the British examination legacy

Scotland's senior phase assessment system sits firmly with the British legacy tradition with its focus on curriculum-based national examinations which are externally set and marked. Like Wales, Ireland and Hong Kong, China, but unlike England, most qualifications involve an element of coursework. The questions are not pre-tested and the grading process relies on examiner, and statistical, judgements about the relative difficulty of a paper. The long history of external examinations in the United Kingdom carries with it the legacy that both the public and policy makers may believe that frequent external examinations are necessary and the fairest way to assess learners. The discussions around the teacher-assessed National 4 qualifications in Scotland, in which there was pressure from some teachers and employers to introduce examinations, reflect these social pressures (Assessment and National Qualifications Working Group, 2018^[77]).

One of the distinctive features of this tradition is the continued use, at the end of compulsory education, of large-scale national examinations for 16 year-olds. In England, where education and training are compulsory until 18, the GCSE still dominates secondary schooling up to 16. Now in Scotland only around 12% of students leave at the end of compulsory schooling (S4), in 2009/10 the leaving rate was around 45% (European Commission, 2021^[78]). In Wales and Ireland, staying in education beyond 16 is now the norm.

Other educational systems in which compulsory education ends at 16 years rely less on a diet of external examinations:

France. The *brevet (Diplôme national du brevet)* is awarded at the end of lower-secondary education (at 15-16 years) and is based on equal contributions from teacher assessment

and examinations. The examination component incorporates an oral test based on an inter-disciplinary project by the student which accounts for 100 of the brevet's 800 marks. The marks in the brevet, which 90% of students pass, are not a requirement for progress into upper-secondary but contribute, along with teacher recommendations and parent and student preferences, to the type of education followed in upper-secondary education.

Norway. The assessment of 16-year-olds, prior to moving on to upper-secondary, is based on teacher assessment of classwork (88%), with local (5%) and central (7%) examinations making up the rest. It is a decentralised model, with recent curriculum changes based on a new core curriculum implemented from 2020 and *Subject Renewal* which seeks to form a better link between the core curriculum and each subject curriculum by emphasising key elements of the core curriculum within each subject (European Commission, 2021^[79]).

New Zealand, which itself broke away from the British tradition, is also undergoing reform as a result of concerns about assessment overload, similar to those Scotland experienced after the introduction of the National Qualifications from 2014. The reformed National Certificate of Educational Attainment (NCEA) will have 'fewer, larger standards' of what has to be achieved, and require fewer credits within each subject (60 instead of 80). The NCEA Level 1 qualification in Year 11 (age 15-16) is optional and intended for school leavers. Most students start Level 2 qualifications which they complete in Year 12. Teacher assessment and external assessments are equally weighted, with the Government announcing that 'Perhaps the biggest assessment opportunity presented by the change package is the chance to reconsider what is at the heart of our learning programmes and to design approaches to assessment that recognise this' (Education Central, 2019^[80]).

Internationally, the Middle Years Baccalaureate (MYB), the junior secondary course offered in many IB international schools, offers a MYB certificate at 16. Candidates for this must take eight e-Assessments, five on-screen examinations in five subjects and three e-Portfolios of coursework, including a personal project. They must also meet school requirements for 'service in action'.

In Scotland the narrower diet of examinations appears to dominate teaching and curriculum in S4 and S5. There is also a backwash into S3, since a nominal 160 hours are needed for each subject of the typical 6-9 subjects, so preparation in schools begins before S4, casting 'a long shadow' back to the initial experiences of secondary school. The subject specifications for the National 5 examinations have become the *de facto* curriculum and can lead to a narrower 'teaching-to-the-test' pedagogy in the subjects that students will generally take. At this point, the 3-18 *Curriculum for Excellence* loses strength (OECD, 2021^[5]).

The question that this comparative study raises is whether the National 5 diet of subject specific examinations is still needed at age 16. CfE already provides a framework for schools' curricula, which calls for a wider range of future-oriented capacities to be developed. Over 80% of students continue in school beyond this point, most taking Higher and Advanced Higher qualifications which have a clearer progression and selection function (Skills Development Scotland, 2015).

7.2. Options for moving beyond legacy models

This section offers some options that may aid the alignment of the current senior phase assessment system with *Curriculum for Excellence* and with more general aspirations for 21st century learning. It is based on the analysis of the Scottish student assessment system in a comparative perspective:

7.2.1. Exploring the replacement of examinations at age 16 by a school graduation certificate

There is a historical ambiguity in Scotland about whether the assessments at S4, S5 and S6 should be seen as a step-by-step ‘ladder’ of qualifications up which students progress; or whether students should simply take a single ‘exit’ examination at the appropriate level (for example National 4/5 for school leavers; Highers for Higher Education; Advanced Highers for university entrance both inside and outside Scotland). The original intention appears to have been for students going on to higher levels of qualifications to bypass lower level qualifications. This did not materialise and the anecdotal evidence is that most schools follow the step-by-step route. An immediate option here is to clarify the intentions behind having three diets of national examinations in three years and to establish the choices that students have in this. Priestley et al. (2020_[12]) have questioned ‘the continued viability of a ladder of qualifications approach, characterised by the ‘two term dash’, and a competency-based ‘mastery approach to assessment’ (p.44).

In comparative terms, Scottish upper-secondary school students are more frequently examined than those in other jurisdictions. This is a consequence of the tradition of offering three suites of examinations (National 5; Highers; Advanced Highers) during secondary years S4, S5, and S6. The yearly demands of the examination system attract criticism from both students and educationalists. After experiencing a Broad General Education (BGE) based on the *Curriculum for Excellence* in primary school, transfer to secondary school is perceived as an experience dominated by examination preparation. Many students are involved in traditional examinations for three consecutive years – a continuation of the historic ‘two term dash’. This diet of examinations may limit the depth and breadth of teaching and learning. The system could be simplified by substituting other forms of certification at S4 that capture more of the students’ capabilities than the current National Qualifications, which for many students will involve a narrowing of the curriculum studied. This is an examination loading not found in the other jurisdictions.

Scotland may consider ‘de-cluttering’ the historical diets of examinations during upper-secondary years S4-S6, and reflect on when and why Scottish students should take examinations, and consider alternative ways to acknowledge the end of compulsory schooling. One example of such alternative would be to remove National 5 examinations at 16 (S4), and to move to a school graduation certificate or diploma. There are now relatively few jurisdictions outside the British tradition with national examinations at the end of compulsory schooling. Internationally, the majority of students now stay on in education or training beyond 16 years of age and upper-secondary school assessments focus on students at 18.

As it is the end of compulsory education, some form of school graduation certificate may indeed be more appropriate. This may include some external components combined with school-based assessments and other contributions to the community. It would build on the teacher assessment that has been central to Primary to S3 assessment. This would allow for clearer recognition of the development of the four capacities of the CfE and offer better alignment with the 3-18 curriculum. The certificate would not simply be an extension of the S3 report. It would incorporate activities inside and outside school and broader attainments and might be organised around the four capacities of CfE. It could draw on the broader-based graduation diplomas found in North America, such as the Ontario Secondary School Diploma (Box 7.1). Other examples to draw from include certification in Nordic systems, and the IB Middle Years certificate.

Box 7.1. The Ontario Secondary School Diploma (OSSD)

The Ontario Secondary School Diploma (OSSD) is awarded at age 17/18. Students must earn 18 school assessed credits from ten compulsory weighted areas: English (four credits), one per senior grade; Mathematics three credits (across two grades); Science (two credits), History, Geography, Arts, Health and Physical Education; French (each one credit); Career Studies and Civics (0.5 credit each), and one credit from each of three subject groups (e.g. Language). 12 optional credits from other subject areas are required; as well as 40 hours of community involvement; and the provincial literacy requirement.

For those who do not meet these requirements, the Ontario Secondary School Certificate (OSSC) can be awarded. It requires seven credits (two in English) and seven optional credits. Students who do not meet the OSSC minimum may be awarded a Certificate of Accomplishment which recognises students' achievements.

Source: Ottawa-Carleton District school board, "Ontario Secondary School Diploma requirements", https://ocdsb.ca/secondary/programs/ontario_secondary_school_diploma_requirements [accessed on 15 July 2021].

In a culture like Scotland's, which is historically steeped in examinations, this option may be dismissed as a 'non-runner'. However, it has been observed that at the outset of designing the new National Qualifications, there was active consideration of making National 5, along with National 4, a teacher-assessed qualification. Previous attempts to introduce profile-type assessment have had limited impact, such as the Howie Report published in 1977 (Scottish Government, 1992^[81]), which sought to encourage this at S5/6, and the UK Technical and Vocational Educational Initiative (TVEI) in the 1990s. The S3 report at the end of Broad General Education encourages recording of a wide range of both academic and broader achievements. However, there is little evidence of them being taken seriously. The wide recognition of the lack of alignment of National 5 with *Curriculum for Excellence* (OECD, 2021^[5]) and the system failures during the COVID-19 pandemic, could provide a new impetus for change.

If the intention of the National Qualifications is alignment with *Curriculum for Excellence*, further steps are needed to encourage a more expansive pedagogy in S4-6. The critical question is whether the preparation for the National Qualifications is still closer to the historical preparation for the Standard Grades than to the aspirations of *Curriculum for Excellence*.

The case for a certificate at 16 years incorporating a broader range of attainments than SQA subject certificates is strengthened by the examination statistics from 2019. The teacher-assessed National 4 had 46 544 learners entered. Of these, 43.2% (20 107 learners out of 46 544) registered only one pass while 4.9% (2 278 out of 46 544) had no passes (Scottish Qualifications Authority, 2019^[82]). These results provide minimal information about the students, a substantial proportion of whom will leave school. This is also the case for National 5 in which, of the 80 046 learners entered, 30% (23 994 out of 80 046) achieved only one A-C pass and 15% (12 052 out of 80 046) had no passes (Scottish Qualifications Authority, 2019^[83]). Of the students who left secondary school at this point in 2019-20, 57% went into Further Education and 18% into employment (Scottish Government, 2021^[84]). A fuller profile of achievement might provide richer information for users and serve leavers better at this stage, for example by identifying social contributions and attainments such as vocational, music, or Duke of Edinburgh awards. For

the majority who stay on into post-compulsory education, the Higher and Advanced Higher examination results are the basis for selection to university, training and employment. National 4 and 5 qualifications are not critical to these selection processes. The removal of National 5 examinations could allow more in-depth preparation for the Higher examinations.

These options are not new, they were discussed both in the design of CfE for secondary schools and at the formulation of the new National Qualifications. It appears that each time the embedded examination tradition has prevailed.

7.2.2. Developing a more resilient upper-secondary assessment system

The COVID-19 experience of 2020-2021 could stimulate the option of further developing more localised and resilient models of assessment as schooling moves back to normal levels. The lockdowns caused by the COVID-19 pandemic exposed the fragility of many qualification systems, particularly those in the United Kingdom. The cancellation of external examinations created an awarding crisis, especially in Scotland and England. The use of statistical measures at a national level worsened the crisis by alienating both students and the general public. This led to a return to the original teacher estimated grades.

The October and December cancellations of the 2021 examinations in Scotland meant that teacher assessments, based in part on SQA-issued mini-examinations and mark schemes, and on classwork, were used in the awards, as they were in Wales. While examinations were later cancelled in England, results were based on teacher assessments with ‘mini-exams’ made available by awarding bodies for optional use by teachers (Roberts and Danechi, 2021^[85]).

What has been brought into sharper focus during lockdown in the United Kingdom, with its school closures and online learning, are the disparities experienced by students, in terms of the difference in resources available for the socially advantaged and the socially disadvantaged. Examinations are only meritocratic and fair when candidates have the same opportunities to access the curriculum and examination resources. The differential learning losses in 2020 and 2021 raise the issue of who is in the best position to judge the relative attainments of students. It may well be that local solutions are more dependable.

Awards were less affected by lockdown in those ‘mixed economy’ systems where they are based on a combination of teachers’ continuous assessment, on school-based examinations as well as external examinations. The United Kingdom at present is particularly vulnerable to any disruption to the national examination system or by differential learning loss as students prepare for examinations. Schools, rather than the central agencies, have become the fallback during the present crisis and proved they could cope. For example, grades were issued without major disruption in Canada, Norway the IB, whose ‘mixed economy’ example could be the foundation of a more dependable system.

7.2.3. Seeking better alignment of assessment with curriculum and pedagogy through broadening the forms of assessment

John Biggs’ concept of constructive alignment seeks to strengthen the relationship of curriculum, pedagogy and assessment by ensuring they all pull in the same direction (Biggs and Collis, 1982^[1]). Misalignment occurs, for example, when the curriculum encourages analysis and reflection, but the examination only asks for naming or describing. Frederiksen and Collins’ systemic validity (1989^[31]) asks a similar question: does the test itself encourage the very skills it is assessing?

The policy intention for the new National Qualification was to:

“reflect the values, purposes and principles of Curriculum for Excellence, and to develop skills for learning, life and work. They are supposed to have an increased emphasis on skills but still have appropriate knowledge and skills for that course area. They aim to be less prescriptive than previous qualifications as they offer flexibility, provide time for learning and have scope for personalisation and choice, especially within the coursework” (Assessment and National Qualifications Working Group, 2018, p. 2_[77]).

The National 5 and Higher and Advanced Higher examinations replaced the Standard, Intermediate 1 and 2, Higher and Advanced Higher qualifications from 2014 onwards with this intention to better align the new examinations with the approaches to teaching and learning encouraged by *Curriculum for Excellence*. Concerns remain however that the National Qualifications have done little to move away from the dominance of examination preparation, with its emphasis on memory and past paper drills, which leads to more didactic secondary school teaching. The Scottish examination system remains a relatively traditional and cautious one (Bhattacharya, 2021_[86]), which is in contrast to the pedagogy encouraged by *Curriculum for Excellence*.

Scotland has a highly centralised and regulated upper-secondary assessment regime. Given the role of Highers and Advanced Highers in the selection for university and occupational selection, this process ensures the necessary comparability of standards in assessment. The format of these relies largely on traditional pen-and-paper timed examinations under standardised conditions. As other jurisdictions, and some SQA qualifications, demonstrate, external assessments can be broader and more creative than simply traditional examinations. SQA could further develop a range of options:

- More use of Information Technology to provide online examination resources and more interactive approaches, with opportunities for candidates to use computers to respond. Relevant international examples from Norway, New Zealand, Finland, and Israel were developed in Section 3.2;
- Incorporation of e-Portfolio and personal projects for external marking (see the International Baccalaureate example in Section 3.2);
- More use of oral presentations and practicals as a way of broadening the assessment formats (see the French and International Baccalaureate examples in Section 4.2).

International evidence points to the potential of digital, AI-powered technologies to expand what skills, knowledge and attitudes assessments can measure, thus offering opportunities to bring 21st century learning and assessment closer together. Where paper-and-pencil, and even computer-based tests fail to measure higher-order skills such as creativity, or emotional and behavioural skills such as collaboration, newer technologies offer great possibilities for effective, equitable and efficient assessments. As education systems explore this potential, they must keep in mind that these technologies can only be beneficial under certain conditions, i.e. when human end users are involved and in control of decision making, when the technologies’ architecture remain transparent, and when decisions on their use are made based on reliable data (OECD, 2021_[22]).

7.2.4. Reconfiguring and increasing the role of school-based assessment and adapting the central moderation system

The more ambitious and future-oriented the curriculum and teaching, the more challenges these present to examinations, especially traditional pen-and-paper ones. In Section 5 issues about the dependability and trustworthiness of assessments were raised, particularly those

of how effectively a curriculum is sampled. The format of traditional examinations means that only a limited range of knowledge and skills can be assessed, while more complex and less tangible skills, for example collaboration and creativity, maybe excluded.

While teacher assessment plays a part in the Scottish examination system, it is generally narrowly conceived as specific pieces of coursework that are submitted and moderated by SQA. Teacher assessment can take other forms, particularly the continuous assessment of regular classroom work, which may include oral and practical work. These are central to upper-secondary student assessment in jurisdictions such as Norway, Ontario, New Zealand, and Queensland. Tests developed by teachers, possibly drawing on a central questions bank as in the reformed Baccalauréat in France, may also be part of this wider range of assessments.

Teachers in Scotland are trusted to make ongoing assessments of their pupils throughout primary and lower-secondary schooling, as are lecturers in Further and Higher Education. Upper-secondary teachers do provide coursework marks, which contribute to examination grades, though the weighting of these is relatively low, and moderation requirements are administratively demanding. This system is highly centralised and controlled in comparison to systems such as Ontario, Queensland and Norway. Even the highly centralised *Baccalauréat* is regionally assessed.

In ‘high-trust’ jurisdictions such as Norway and Canada, assessment is far more decentralised and left to schools and teachers. A source of dissatisfaction with the original National Qualifications, similar to the assessment reforms in Hong Kong, China, were the complex recording and moderation procedures. Much simpler models are used in other jurisdictions, incorporating teacher’s continuous assessment of their students based on performance over time. This allows a wider range of attainments to be incorporated. SQA’s role would then be far more ‘light touch’ in relation to National Qualifications.

Anecdotal comments collected from students during this study suggest they see exam preparation as a narrowing experience, with rote learning and memorisation a regular feature in their classes. This points to a misalignment with the curricular intentions of CfE. Providing schools with more freedom to assess the wider curriculum may increase the validity of the assessments.

In many jurisdictions, more is devolved to the school level which then generates additional assessment options:

- A more central role for continuous teacher assessment during the course based on classwork and school-based tests (see the examples of Ontario, Norway, France, and New Zealand in Section 3.3);
- Teacher set and marked work which is externally moderated by other teachers (see the example of Queensland, *ibid*);
- Oral and practical presentations which are locally set and moderated (France (Section 4.2.3), Norway Section 4.2.2).

Such approaches allow for fuller alignment with 21st century curriculum reform such as Scotland’s. As evidenced during the COVID-19 pandemic, these approaches also offer greater resilience where there is major disruption. Future evolutions may require Scotland to decentralise more some of its assessment procedures, while further developing teachers’ assessment literacy and the professional capacity of schools in assessment.

7.2.5. Systematically investigating students' perceptions and views of assessment arrangements

Students, as individuals, are the most affected by any assessment system and are therefore key stakeholders who need to be consulted. However, they may only have a limited voice in shaping it. The *Rapid Review of National Qualifications Experience 2020* (Priestley et al., 2020_[12]) observed:

“Young people, as stakeholder and rights holders, are at the heart of a qualifications system for schools and colleges, and need to be involved fully in decisions which affect them, in line with Scotland’s obligations to the UNCRC.” (Priestley et al., 2020, p. 47_[12]).

The authors found support from young people for achievement to be captured throughout the year, rather than the ‘two term’ dash towards examinations (in particular for Higher). The report also found:

“consistent support from all stakeholders (including young people and parents) for a reduced emphasis on terminal examinations as the basis for qualifications. There is widespread support for continuous assessment and its benefits.” (p44).

There appears to be limited systematic research evidence on how Scottish students themselves would prefer to be assessed. As Scotland is a signatory to the *United Nations Convention on the Rights of the Child* (UNCRC), fuller attention to the preferences of students would be anticipated. Evidence from consultation panels, for example SQA’s *Young Scot Vision Panel* (2018), and from interviews (OECD, 2021_[5]) also suggest young learners would prefer a greater emphasis on continuous assessments by their teachers. While these play a central role in other jurisdictions, many teachers, employers and parents in Scotland appear to be wedded to examinations and would even wish to see them at National 4.

The tensions in the system were illustrated in the 2018 *Curriculum and Assessment Board* report on National 4 Qualifications. These are not externally examined, but the report pointed that:

“a clear majority of teachers and senior managers expressed the opinion that National 4 learners needed an examination at the end of the course. Teachers commented on the need for an exam to motivate learners...” (Curriculum and Assessment Board, 2018, p. 5_[87]).

However,

“The majority of S4 learners took a different view, and this view was echoed by learners in S5 and S6. Learners judged they were working hard or very hard and did not require an external assessment at National 4” (Curriculum and Assessment Board, 2018, p. 5_[87]).

While the Scottish examination system offers students a choice of pathways and progression rates towards qualifications, it may be that, in practice, school timetables and resource constraints limit these options. If policy is to be responsive to student needs and aspirations, more systematic study of their perceptions needs support.

7.2.6. Further developing the role of vocational qualifications in broadening the curriculum

A 19th century legacy feature of the British examination system has been the privileging of academic forms of assessment, particularly the written examination, over the direct

assessment of practical or performance skills. Hanson (1993_[88]) points out that:

“Because tests act as gatekeepers to many educational and training programs ...the likelihood that someone will be able to do something, as determined by the tests, becomes more important than one’s actually doing it.” (Hanson, 1993, p. 288_[88]).

Hanson calls this the “fabricating quality of tests”. One consequence of this has been to treat the more applied vocational routes as less prestigious than the general academic routes. This may mean that students are discouraged from mixing applied and academic courses, thus narrowing their options and experiences.

With SQA being responsible for both general and vocational qualifications, Scotland is well placed to raise the profile of vocational qualifications as a way of broadening its curriculum offer and developing student capacities. There is also considerable permeability in the system for transfer within Higher Education, through such as Foundation Apprenticeships and National Progress Awards (NPAs) (Brown, 2019_[89]). One option is to further integrate them into the mainstream qualification offer. Some applied subjects can be taken as National and Higher Qualifications (for example Accounting; Care; Health and Food Technology) though these have relatively small entries. In 2019, 64 267 skill-based qualifications were achieved, easily double the figure in 2012. By 2020 this had risen to 71 723 (information provided by the SQA).

In countries such as Australia and New Zealand, vocational and academic subjects can be combined in a single mainstream qualification offer. In Ireland the Leaving Certificate Vocational Programme (LCVP) is integrated with the national Leaving Certificate and involves joint modules. It has the same status for university entrance and is taken by over a quarter of the Certificate cohort (O’Donnell, 2018_[90]). This is also the case in Norway, where 50% of students follow vocational routes at 16 which lead into either apprenticeships after two years or staying in education for three years.

New Zealand has continued with the modular system inspired by Scotland in order to provide vocational qualifications. It has gone further by encouraging the integration of general and vocational subjects in the National Certificate of Educational Achievement (NCEA). Around one third of the 85% of students who achieve NCEA Level 2 will have at least one subject from one of the six vocational pathways.

In France, the *Baccalauréat* was broadened in 1985 to include the *Baccalauréat professionnel* (*professional Baccalaureate*) with the intention of creating parity of esteem for vocational qualifications as it would qualify successful candidates for university entrance. This is now taken by just over a quarter of the cohort, with an 80% pass rate. However, the fact that fewer than 5% of the successful students go on to become university graduates suggests parity has not been fully achieved. The *professional Baccalaureate* is currently being reformed to modernise its format (15 groups of professions instead of around 100 options currently), and to include more opportunities for workplace experience. The assessment will also include an oral presentation on a practical project the student has worked on. This will provide a more valid and fit-for-purpose qualification. Whether it improves the public status of the *professional Baccalaureate* has yet to be seen.

The historical pattern of giving preference to the written examination has left vocational education and training struggling to establish parity of esteem. The more successful approaches appear to be in those systems which have integrated both in their qualifications but, in these too, society may still value the academic strand over the vocational. Scotland is in a position to further integrate the separate qualifications and to further raise the status of vocational qualifications. In its final report, the Commission for Developing Scotland’s Young Workforce observed:

“Curriculum for Excellence by its nature provides the opportunity for a more balanced and inclusive approach to academic and vocational education with the potential to blend the two to the needs of individual pupils.... Through ambitious partnership between our schools and colleges, many of our young people not inclined to pursue an academic pathway could leave school with high level vocational qualifications which have strong currency in the labour market. By significantly enhancing the vocational content of the offer to pupils, we would follow the example of the best performing European countries in terms of youth employment without splitting young people off into separate streams at school age” (Commission for Developing Scotland's Young Workforce, 2014, p. 20_[91]).

To fulfil these aspirations, options may be explored to increase both the integration and status of vocational awards. This would require changes in how they are viewed by schools and parents, Higher Education selectors, and employers. In the case of selectors and employers, endorsements and training and study offers from prestigious sources may help to change the traditional British mindset.

8. Conclusions

We live in a world where the changing educational expectations of an increasingly diverse student population are reflected in curriculum reform and changing pedagogy. Scotland's *Curriculum for Excellence* has been recognised as a pioneering example of one such future-facing curriculum. Assessment systems around the world have often struggled to ensure constructive alignment with these educational aspirations. In jurisdictions where upper-secondary assessments have high-stakes selective and accountability functions, national examinations have often inhibited changes to teaching and learning. This is partly because the examination syllabus becomes the *de facto* curriculum and teachers switch to narrower test preparation methods with secondary years students. The COVID-19 pandemic of 2020 illustrated the fragility of those systems that are largely dependent on terminal examinations for secondary school students.

Examination systems are products of specific cultures and each has its own historical legacy. Scotland's system sits within the British tradition, which had its origins in the Victorian enthusiasm for written academic examinations. One legacy feature of this is the diet of central examinations at 16, even when few students now leave education at that age, and something no longer seen in most other assessment systems.

Better alignment between curriculum and examinations may result from further technical innovation, particularly online interactive approaches and resources which allow for a wider range of question types and answer formats. A more decentralised approach, in which schools share more assessment responsibilities, may also align better with the curriculum and 21st century pedagogy. This may involve rethinking SQA's demanding quality assurance and moderation processes. The cancellation of central examinations in 2020 and 2021 and the reliance on teacher assessment has shown how this is possible, as have practices in such as Ontario, New Zealand and Queensland.

Whatever routes are chosen, at the heart of any such system should be those who are most impacted: the learners being assessed. Students should have opportunities to influence assessment policies. Educational equity requires a dependable and resilient assessment system which is perceived as trustworthy by the key stakeholders in society.

References

- ACT (2021), *About the ACT Test*, <http://www.act.org/content/act/en/products-and-services/the-act-educator/the-act-test.html> (accessed on 15 July 21). [47]
- Adams, R. (2020), *Gavin Williamson to blame for England exams fiasco, says Ofqual chair*, <https://www.theguardian.com/education/2020/sep/02/gavin-williamson-to-blame-for-england-exams-fiasco-says-ofqual-chair> (accessed on 15 July 2021). [74]
- Arnott, M. and J. Ozga (2016), “Education and nationalism in Scotland: governing a ‘learning nation’”, *Oxford Review of Education*, Vol. 42/3, pp. 253-265, <http://dx.doi.org/10.1080/03054985.2016.1184865>. [42]
- Assessment and National Qualifications Working Group (2018), *Assessment Overload*, <https://bit.ly/2UIjkrD> (accessed on 15 July 2021). [77]
- Baird, J. et al. (2017), “Assessment and learning: fields apart?”, *Assessment in Education: Principles, Policy & Practice*, Vol. 24/3, pp. 317-350, <http://dx.doi.org/10.1080/0969594x.2017.1319337>. [45]
- Baird, J. et al. (eds.) (2018), *Culture, context and controversy*, UCL IoE Press. [72]
- Baird, J. et al. (eds.) (2018), *Examination Standards: How measures and meanings differ around the world*, UCL IoE Press. [6]
- Baird, J. et al. (eds.) (2018), *Setting standards in the United States: The Advanced Placement Program*, UCL IoE Press. [50]
- Baird, J. et al. (eds.) (2018), *Standard setting in France: The baccalauréat*, UCL IoE Press. [53]
- Bhattacharya, A. (2021), *Encouraging innovation and experimentation in Scottish schools*, Social Market Foundation, <https://www.smf.co.uk/wp-content/uploads/2021/03/Innovation-and-experimentation-in-Scottish-schools-March-2021.pdf> (accessed on 15 July 2021). [86]
- Biggs, J. and K. Collis (1982), *Evaluating the quality of learning: The SOLO Taxonomy*, Elsevier, <https://doi.org/10.1016/C2013-0-10375-3>. [1]
- Black, P. et al. (2010), “Validity in teachers’ summative assessments”, *Assessment in Education: Principles, Policy & Practice*, Vol. 17/2, pp. 215-232, <http://dx.doi.org/10.1080/09695941003696016>. [33]
- Broadfoot, P. (1979), *Assessment, Schools and Society*, Routledge. [44]
- Brown, J. (2019), “Integrating vocational education and training for secondary school students”, *Australian Council for Educational Research (ACER)*, <https://research.acer.edu.au/nswcurriculumreview/2/>. [89]
- CEA, SQA and the Welsh Government (2014), *Centre readiness to use e-assessment*, Council for the Curriculum, Examinations and Assessment (CCEA), Scottish Qualifications Authority (SQA), Welsh Government, <https://ceea.org.uk/downloads/docs/regulation-asset/Guidance/Centre%20Readiness%20to%20use%20e-Assessment.pdf> (accessed on 15 July 2021). [25]

- CNBC (2020), “Irish students rejoice over government U-turn calling off high school tests due to the coronavirus”, *CNBC Make it*, <https://www.cnbc.com/2020/05/12/coronavirus-irish-students-rejoice-over-leaving-certificate-u-turn.html> (accessed on 15 July 2021). [68]
- College Board (2021), *Advanced Placement (AP) Central*, <https://apcentral.collegeboard.org/> (accessed on 29 July 2021). [28]
- CollegeBoard (2021), *College Readiness - SAT*, <http://www.collegeboard.com/student/testing/sat> (accessed on 15 July 2021). [49]
- Commission for Developing Scotland’s Young Workforce (2014), *Education working for all!*, <https://www.gov.scot/publications/education-working-commission-developing-scotlands-young-workforce-final-report/documents/> (accessed on 15 July 2021). [91]
- Compass Education Group (2020), *2020 AP Exams in Review*, <https://www.compassprep.com/2020-ap-exams-at-home/> (accessed on 29 July 2021). [29]
- Cumming, J. (2019), “Senior secondary school assessment and standard-setting in Queensland, Australia: social context and paradigmatic change”, *Assessment in Education: Principles, Policy & Practice*, Vol. 27/2, pp. 160-177, <http://dx.doi.org/10.1080/0969594x.2019.1684877>. [57]
- Curriculum and Assessment Board (2018), *National 4 - Way Forward: Update*, <https://bit.ly/3rsyQEh> (accessed on 15 July 2021). [87]
- Deerin, C. (2021), “How Scotland is heading for an even bigger exams debacle”, *NewStatesman*, <https://www.newstatesman.com/politics/staggers/2021/05/how-scotland-heading-even-bigger-exams-debacle> (accessed on 15 July 2021). [70]
- DeLuca, C. et al. (2017), “Grading policies and practices in Canada: a landscape study”, *Canadian Journal of Educational Administration and Policy*, Vol. 184, pp. 4-22, <https://files.eric.ed.gov/fulltext/EJ1165318.pdf>. [51]
- Education Central (2019), “Changes to NCEA announced”, *Education Central / Pokapu Matauranga*, <https://educationcentral.co.nz/changes-to-NCEA-announced/> (accessed on 15 July 2021). [80]
- EGG, T. and G. Stobart (eds.) (2014), *High-Stakes Testing in Education*, Routledge. [64]
- Emmertsen, S. (2020), *The future of digital assessment - Covid-19, short courses and beyond*, <https://www.timeshighereducation.com/hub/inspera/p/future-digital-assessment-covid-19-short-courses-and-beyond> (accessed on 15 July 2021). [23]
- European Commission (2021), *Comparative Reports*, https://eacea.ec.europa.eu/national-policies/eurydice/publications_en (accessed on 15 July 2021). [78]
- European Commission (2021), “Norway National Reforms in School Education”, *Eurydice*, https://eacea.ec.europa.eu/national-policies/eurydice/content/national-reforms-school-education-48_en (accessed on 15 July 2021). [79]
- Frederiksen, J. and A. Collins (1989), “A Systems Approach to Educational Testing”, *Educational Researcher*, Vol. 18/9, p. 27, <http://dx.doi.org/10.2307/1176716>. [31]
- Furlong, J. and I. Lunt (2016), “Education in a Federal UK”, *Oxford Review of Education*, Vol. 42/3, pp. 249-252, <http://dx.doi.org/10.1080/03054985.2016.1184867>. [43]

- Geisinger, K. (2016), “21st Century Skills: What Are They and How Do We Assess Them?”, *Applied Measurement in Education*, Vol. 29/4, pp. 245-249, <http://dx.doi.org/10.1080/08957347.2016.1209207>. [19]
- Gouédard, P. et al. (2020), “Curriculum reform: A literature review to support effective implementation”, *OECD Education Working Papers*, No. 239, OECD Publishing, Paris, <https://dx.doi.org/10.1787/efe8a48c-en>. [18]
- Hanson, A. (1993), *Testing, Testing: Social consequences of the examined life*, University of California Press. [88]
- Hong Kong Curriculum Development Council (2001), *Learning to Learn - The Way Forward in Curriculum Development*, <https://www.edb.gov.hk/en/curriculum-development/cs-curriculum-doc-report/wf-in-cur/index.html> (accessed on 15 July 2021). [15]
- Hopfenbeck, T. and G. Stobart (2015), “Large-scale implementation of Assessment for Learning”, *Assessment in Education: Principles, Policy & Practice*, Vol. 22/1, pp. 1-2, <http://dx.doi.org/10.1080/0969594x.2014.1001566>. [21]
- Hunter, S. (1963), “Scottish education: Changes in the examination structure in secondary schools”, *International Review of Education*, Vol. 9/3, pp. 310-324, <http://dx.doi.org/10.1007/bf01416157>. [9]
- Husen, T. (1967), *International study of achievement in mathematics: A comparison of twelve countries (Volume II)*, Almqvist & Wiksell. [40]
- International Baccalaureate Organization (2020), “May 2020 examinations will no longer be held”, *IBO.org*, <https://www.ibo.org/news/news-about-the-ib/may-2020-examinations-will-no-longer-be-held/> (accessed on 23 March 2021). [71]
- Jeantheau, J. and S. Johnson (2019), “Assessment innovation and reform in France”, *Assessment in Education: Principles, Policy & Practice*, Vol. 27/3, pp. 290-308, <http://dx.doi.org/10.1080/0969594x.2019.1577218>. [37]
- Kuhn, T. (1962), *The structure of scientific revolutions*, University of Chicago Press. [73]
- Lee, H., G. Lee and R. Openshaw (2013), “Radical Assessment and Qualification Reform in New Zealand: The Rocky Road to National Certificate of Educational Achievement”, *World Studies in Education*, Vol. 14/2, pp. 25-45, <http://dx.doi.org/10.7459/wse/14.2.03>. [58]
- Looney, A. (2006), “Assessment in the Republic of Ireland”, *Assessment in Education: Principles, Policy & Practice*, Vol. 13/3, pp. 345-353, <http://dx.doi.org/10.1080/09695940601035544>. [62]
- MacPhail, A., J. Halbert and H. O’Neill (2018), “The development of assessment policy in Ireland: a story of junior cycle reform”, *Assessment in Education: Principles, Policy & Practice*, Vol. 25/3, pp. 310-326, <http://dx.doi.org/10.1080/0969594x.2018.1441125>. [35]
- Martínez, J., B. Stecher and H. Borko (2009), “Classroom Assessment Practices, Teacher Judgments, and Student Achievement in Mathematics: Evidence from the ECLS”, *Educational Assessment*, Vol. 14/2, pp. 78-102, <http://dx.doi.org/10.1080/10627190903039429>. [34]
- Matters, G. and G. Masters (2014), *Redesigning the secondary-tertiary interface*, *Queensland Review of senior assessment and tertiary entrance*, Australian Council for Educational Research (ACER), https://research.acer.edu.au/qld_review/1/ (accessed on 15 July 2021). [56]

- Maxwell, G. (2010), “Moderation of Student Work by Teachers”, in *International Encyclopedia of Education*, Elsevier, <http://dx.doi.org/10.1016/b978-0-08-044894-7.00347-x>. [55]
- McMillan, J. (ed.) (2013), *Research on Classroom Summative Assessment*, SAGE Publications, Inc., <https://www.doi.org/10.4135/9781452218649.n14>. [32]
- McVittie, J. (2008), *National Qualifications: a short history*, Scottish National Agency (SQA), https://www.sqa.org.uk/files_ccc/PNP_ResearchReport3_NationalQualificationsAShortHistory.pdf (accessed on 15 July 2021). [41]
- New Zealand Government (2019), *Let’s talk about Education - Korero Matauranga*, <https://conversation.education.govt.nz/> (accessed on 15 July 2021). [59]
- Norwegian Directorate of Education and Training (2018), *Primary, Lower and Upper Secondary School*, <https://www.udir.no/in-english/quality-in-ecec-schools-and-vocational-education-and-training/kindergarden/> (accessed on 15 July 2021). [17]
- Norwegian Ministry of Education and Research (2017), *Education Act*, <https://lovdata.no/dokument/NLE/lov/1998-07-17-61> (accessed on 15 July 2021). [16]
- O’Donnell, S. (2018), *Upper Secondary Education in Nine Jurisdictions*, National Council for Curriculum and Assessment (NCCA), <https://ncca.ie/media/3337/scoping-report-online-2.pdf> (accessed on 15 July 2021). [90]
- OECD (2021), *OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/589b283f-en>. [22]
- OECD (2021), *Scotland’s Curriculum for Excellence: Into the Future*, Implementing Education Policies, OECD Publishing, Paris, <https://dx.doi.org/10.1787/bf624417-en>. [5]
- OECD (2020), *Education at a Glance 2020: OECD Indicators*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/69096873-en>. [7]
- OECD (2020), *PISA 2018 Results (Volume VI): Are Students Ready to Thrive in an Interconnected World?*, PISA, OECD Publishing, Paris, <https://dx.doi.org/10.1787/d5f68679-en>. [11]
- OECD (2020), *What Students Learn Matters: Towards a 21st Century Curriculum*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/d86d4d9a-en>. [20]
- OECD (2019), *OECD Future of Education and Skills 2030 Conceptual learning framework. Concept note: OECD Learning Compass 2030*, https://www.oecd.org/education/2030-project/teaching-and-learning/learning/learning-compass-2030/OECD_Learning_Compas_2030_concept_note.pdf (accessed on 15 July 2021). [3]
- OECD (2015), *Education at a Glance 2015: OECD Indicators*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/eag-2015-en>. [60]
- OECD (2015), *Improving schools in Scotland: an OECD perspective*, OECD Publishing, <http://www.oecd.org/education/school/Improving-Schools-in-Scotland-An-OECD-Perspective.pdf>. [66]
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264190658-en>. [4]

- OECD (2010), *Learning for Jobs*, OECD Reviews of Vocational Education and Training, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264087460-en>. [8]
- Ofqual (2020), *Online and on-screen assessment in high stakes, sessional qualifications. A review of the barriers to greater adoption and how these might be overcome*, Ofqual, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943382/Barriers_to_online_111220.pdf (accessed on 15 July 2021). [24]
- Ofqual (2020), “The impact of the coronavirus outbreak on exams around the world”, *The Ofqual blog*, <https://ofqual.blog.gov.uk/2020/05/22/the-impact-of-the-coronavirus-outbreak-on-exams-around-the-world/> (accessed on 15 July 2021). [69]
- Ontario Ministry of Education (2010), *Growing Success: Assessment, evaluation and reporting in Ontario Schools*, Ontario Ministry of Education, <http://www.edu.gov.on.ca/eng/policyfunding/growsuccess.pdf> (accessed on 15 July 2021). [52]
- Opposs, D. et al. (2020), “Governance structure and standard setting in educational assessment”, *Assessment in Education: Principles, Policy & Practice*, Vol. 27/2, pp. 192-214, <http://dx.doi.org/10.1080/0969594x.2020.1730766>. [48]
- Priestley, M. et al. (2020), *Rapid Review of National Qualifications Experience 2020*, <https://www.gov.scot/publications/rapid-review-national-qualifications-experience-2020/> (accessed on 29 March 2021). [12]
- Princeton Review (2021), *What is the SAT?*, <https://www.princetonreview.com/college/sat-information> (accessed on 15 July 2021). [46]
- Qualifications Wales (2021), *Qualified for the Future, the right choice for Wales*, Qualifications Wales, <https://www.qualificationswales.org/media/6997/qualified-for-the-future-the-right-choice-for-wales-consultation.pdf> (accessed on 15 July 2021). [30]
- Reay, D. and D. Wiliam (1999), “‘I’ ll be a nothing’: structure, agency and the construction of identity through assessment”, *British Educational Research Journal*, Vol. 25/3, pp. 343-354, <http://dx.doi.org/10.1080/0141192990250305>. [76]
- Roberts, N. and S. Danechi (2021), *Coronavirus: GCSEs, A Levels and equivalents in 2021*, House of Commons Library, <https://researchbriefings.files.parliament.uk/documents/CBP-9045/CBP-9045.pdf> (accessed on 15 July 2021). [85]
- Scottish Government (2021), *A changing nation: How Scotland will thrive in a digital world*, <https://www.gov.scot/publications/a-changing-nation-how-scotland-will-thrive-in-a-digital-world/> (accessed on 15 July 2021). [27]
- Scottish Government (2021), *OECD Independent review of Curriculum for Excellence 2020-2021 - Initial evidence pack*, <https://www.gov.scot/publications/oecd-independent-review-curriculum-excellence-2020-2021-initial-evidence-pack/> (accessed on 24 March 2021). [10]
- Scottish Government (2021), *Summary statistics for attainment and initial leaver destinations*, <https://bit.ly/3iVSTax> (accessed on 15 July 2021). [84]
- Scottish Government (2016), *Realising Scotland’s full potential in a digital world: a digital strategy for Scotland*, <https://www.gov.scot/binaries/content/documents/govscot/publications/strategy-plan/2017/03/realising-scotlands-full-potential-digital-world-digital-strategy-scotland/documents/00515583-pdf/00515583-pdf/govscot%3Adocument/00515583.pdf> (accessed on 15 July 2021). [26]

- Scottish Government (1992), *Scottish Government Yearbook*, [81]
http://www.scottishgovernmentyearbooks.ed.ac.uk/record/23104/1/1992_9_TheHowieCommitteeonpost-compulsoryschooling.pdf (accessed on 15 July 2021).
- Scottish Qualifications Authority (2019), *National 4 Annual statistical report*, [82]
https://www.sqa.org.uk/sqa/files_ccc/ASR2019_National4.xls (accessed on 29 July 2021).
- Scottish Qualifications Authority (2019), *National 5 Annual statistical report*, [83]
https://www.sqa.org.uk/sqa/files_ccc/ASR2019_National5.xls (accessed on 29 July 2021).
- SQA and YoungScot Observatory (2018), *Insight #SQAfutures*, Scottish Qualifications Authority (SQA), [13]
https://www.sqa.org.uk/files_ccc/221018_SQAAssessmentFutures_YoungScotReport.pdf (accessed on 15 July 2021).
- Stobart, G. (2008), *Testing Times*, Routledge, <http://dx.doi.org/10.4324/9780203930502>. [2]
- Stobart, G. (2005), “Fairness in multicultural assessment systems”, *Assessment in Education: Principles, Policy & Practice*, Vol. 12/3, pp. 275-287, <http://dx.doi.org/10.1080/09695940500337249>. [39]
- Success at School (2020), *Success at School*, <https://successatschool.org/> (accessed on 15 July 2021). [65]
- Suto, I. and T. Oates (2021), *High stakes testing after basic secondary education, how and why is it done in high performing education systems?*, Cambridge Assessment, [36]
<https://www.cambridgeassessment.org.uk/Images/610965-high-stakes-testing-after-basic-secondary-education-how-and-why-is-it-done-in-high-performing-education-systems-.pdf> (accessed on 15 July 2021).
- The Times (2020), *Rethinking assessment: mutant exam system is failing our children*, The Sunday Times, [75]
<https://www.thetimes.co.uk/article/rethinking-assessment-mutant-exam-system-is-failing-our-children-svwkzvx7n>.
- Tong, C., C. Lee and G. Luo (2020), “Assessment reform in Hong Kong: developing the HKDSE to align with the new academic structure”, *Assessment in Education: Principles, Policy & Practice*, Vol. 27/2, pp. 232-248, <http://dx.doi.org/10.1080/0969594x.2020.1732866>. [63]
- Tveit, S. (2013), “Educational assessment in Norway”, *Assessment in Education: Principles, Policy & Practice*, Vol. 21/2, pp. 221-237, <http://dx.doi.org/10.1080/0969594x.2013.830079>. [38]
- UK Department for Education (2019), *Guidance - English Baccalaureate (EBacc)*, [54]
<https://www.gov.uk/government/publications/english-baccalaureate-ebacc/english-baccalaureate-ebacc> (accessed on 15 July 2021).
- UNESCO (2020), *Review of high-stakes exams and assessments during Covid-19*, UNESCO, [67]
https://en.unesco.org/sites/default/files/unesco_review_of_high-stakes_exams_and_assessments_during_covid-19_en.pdf (accessed on 15 July 2021).
- Viennet, R. and B. Pont (2017), “Education policy implementation: A literature review and proposed framework”, *OECD Education Working Papers*, No. 162, OECD Publishing, Paris, [61]
<https://dx.doi.org/10.1787/fc467a64-en>.
- Welsh Government (2020), *Curriculum for Wales*, <https://hwb.gov.wales/curriculum-for-wales> (accessed on 15 July 2021). [14]