



**COMMISSION
ON SCHOOL
REFORM**

Challenge Paper

Exams – do we need them?

August 2021

Membership of the re-convened Commission on School Reform is as follows:

- **Keir Bloomer (Chair):** Education Consultant and former Director of Education
- **John Barnett:** University Court Member, retired finance professional and former Parent Council Chairman
- **Helen Chambers:** Deputy Chief Executive of Inspiring Scotland
- **Jamie Cooke:** Head of RSA Scotland
- **Carole Ford:** Former Head teacher
- **Jim Goodall:** Former Head of Education and Community Services at Clackmannanshire Council and current Lib Dem councillor at East Dunbartonshire Council
- **Anna Hazel-Dunn:** Headteacher Royal High Primary
- **Frank Lennon:** Former Head of Dunblane High School
- **Ross Martin:** Economic Agitator with experience in a variety of roles in education-including school teacher, voluntary tutor, chair of local education authority, member of college board & university court, Chair of a charity, and, most importantly, a parent and former student.
- **Lindsay Paterson:** Professor of education policy in the School of Social and Political Science at Edinburgh University
- **Lesley Sutherland:** Board member, the Centre for Scottish Public Policy

Please note that all members of the commission participate in an individual capacity and that the views of the commission do not represent the views of any other organisation to which the individual members belong.

Purpose

The purpose of this report is to;

- set out the historical and other contexts in which decisions about the roles of coursework and examinations in the assessment and certification of young people in the Council’s secondary schools should be made;
- set out a sample of the research evidence and the arguments based on that evidence which can be made to support the use of coursework and examinations for these purposes;
- present a conclusion about how coursework and examinations could best be used in the assessment and certification of young people in the Senior Phase of their schooling.

Summary

Assessment by means of common standards has helped to guarantee the quality of Scottish secondary schooling. It has thus been at the heart of widening opportunities. For most of the past century, that assessment was by means of exams.

Assessment of course-work has partly replaced exams in the past three decades. Course-work can test skills and understanding that exams miss, such as deep understanding, the ability to collate information from a variety of sources, and solving problems that require lengthy thought.

Course-work has the further advantage of encouraging discussion and collaborative working.

Examinations have unique strengths that cannot be assessed by coursework. They test focus, memory and thinking quickly.

Studying for exams can be an effective way of consolidating learning because it trains the memory and helps pupils’ motivation and focus.

The growth of course-work has required a shift to marking by students’ own teachers. This cannot

be anonymous, and so risks bias against social groups who typically have low attainment.

Anonymous marking of exams by markers who are external to the student’s school is much less subject to bias.

Both exams and assessed course-work have an important role in any system of assessment that is valid and fair.

NOTE: coursework is the generic term used by the SQA to refer to assessed work which contributes to final grade but is not an examination.

Background and Main Issues

The problem

Scottish school examinations have been cancelled for two years in a row. The decision to cancel the examinations has provoked a wider debate about their place in schooling. The Scottish Government’s International Council of Education Advisers has [called](#) examinations ‘essentially an out-of-date 19th and 20th century technology operating in a 21st century environment of teaching and learning’. In their critique, they claim that “the capacity to apply learning creatively in unfamiliar contexts is increasingly the kind of high-value skill demanded by the workplace of the future. Traditional examinations are not capable of making such assessments on their own.” This position was, to all intents and purposes, stated in the OECD Report published in late-June 2021 and it is expected to be reiterated in a second OECD Report which will be published in August of this year and which will focus on assessment and certification.

It is alleged and asserted that examinations:

- lack validity, in the sense that they do not test learning that really matters;
- cannot assess the capacity to think creatively, to explore ideas in an open-ended way, to find out ‘things’ independently, or to engage in debate;

- require the wrong kind of motivation, in the sense that instead of encouraging students to develop an intrinsic interest in a subject, they are claimed to focus students only on getting a certificate.

Scottish examinations have increasingly been exercises in merely reproducing in the examination room material that has been learnt solely for that purpose – for example, essays learnt by heart, or basic summaries of routine scientific procedures. Thus Scottish examinations no longer assess what examinations have the potential to assess.

It is asserted and alleged that the assessment of course-work addresses all of these concerns and that certification derived from the assessment of course-work is more valid because, for example, it tests learning that really matters.

The complex of issues which arise in considering how best to go about the assessment and certification of young people is often presented as a straightforward choice between examinations versus the assessment of course-work. The choice is NOT straightforward. Issues of validity and reliability arise in any consideration of the relative merits of different approaches to assessment and certification.

External assessment in Scottish secondary education: a guarantee of common standards

External assessment and certification has played a key role in defining the central purpose of Scottish schooling.

The history of secondary schooling in Scotland illustrates how examinations can widen opportunity. Scotland adopted a system of national assessment, in 1888, when school inspectors persuaded the government that externally set and externally marked assessment was needed to ensure the quality of the secondary schools that were emerging at that time. The development of a full secondary system over the following four decades then depended crucially on the standards set by this assessment.

The number of examination candidates rose rapidly, reaching a temporary [peak](#) of about 20% of each age group in 1924. Numbers fell back during the depression and the Second World War but growth [resumed](#) after that reaching 10% in the early 1950s, 25% in the 1960s, 66% in the 1970s, and more than 97% in the late-1990s and after.

In essence, and reflecting their original purpose, externally assessed examinations defined the meaning and purpose of secondary education in Scotland, e.g. in relation to breadth and balance. The nature of the examinations has been subject over the years to a number of revisions; Ordinary Grades were introduced in the 1960s, Standard Grades in the 1980s, Intermediate Courses in the 1990s and National 5 Courses most recently.

The constant throughout these many decades has been the Higher Grade, as the main route into Higher Education, signalling Scotland's general preference for a coherent, national standard in its secondary courses.

Equity

External assessment has had also a key role to play in extending opportunities, for example, by enabling young women to demonstrate their achievements as never before. By the eve of the First World War, a [majority](#) (57%) of the candidates for the leaving certificate was female. This never fell below 40% in the next half century, and equalled the male proportion from the 1970s onwards. These assessments thus became the means by which female students' potential could begin to be fairly measured for the first time, well ahead of the legislation in the 1970s that sought to guarantee equal opportunities.

The same was true for Scotland's largest minority ethnic group historically, Catholics whose family origins were in Ireland. Through legislation in the aftermath of the First World War, the state took over the responsibility for almost all Catholic schools. In 1911, the proportion of leaving-certificate candidates who were in Catholic schools was only 5%, even though Catholics were about 11% of the population. By 1935, the

corresponding proportions were nearly at parity, 11% and 13%. The decades following the introduction of comprehensive schooling allowed attainment by children from Catholic schools to equal that in non-denominational schools. [For example](#), in 1998, the proportion passing at least one Standard Grade was 85% in non-denominational schools and 84% in Catholic schools.

Although socio-economic inequalities in attainment and progress have not vanished in the way that sex and religious differences have, opportunities have become more equal as a result of the extension of access to standardised exams which was opened up by successive waves of reform to the structure of schooling.

This very brief, reflection on the history of external assessment and certification in Scotland's schools may serve to indicate how it has been used to establish the very meaning of secondary schooling, guaranteeing that it is provided to a common standard throughout the country, for all types of school and all social groups. It has been a central part of the gradual overcoming of invidious discrimination, a central part of the process by which a young person is judged by what they can do rather than by who they are.

Reliability and validity

There are two key [ideas](#) in the evaluation of any mode of assessment – reliability and validity.

Reliability refers to whether a candidate would get the same result if the assessment was repeated. Validity is whether the assessment tests the objectives of the course. It is generally agreed that the strongest reliability is obtained by the most tightly standardised tests, such as in the use of tests of cognitive ability in research on intelligence. The problem is then that tightly standardised tests tend not to be valid in an educationally useful way.

For example, a highly reliable test of students' capacity to solve a quadratic equation in algebra can be designed. There is only one correct answer, and so all mathematically competent markers will

agree whether a particular student has got it right. Facts of this kind are important, and knowledge of them can readily be assessed by examinations of a fairly traditional kind. Understanding mathematics needs a knowledge of multiplication, algebraic rules, and the laws of geometry. Understanding science needs knowledge of the laws of physics, the basis of genetics, or the periodic table. Understanding history needs a knowledge of dates and of the basic story of what happened. Understanding a language needs knowledge of grammar, spelling and intonation. Being able to play a musical instrument requires practical knowledge of scales, rhythm and harmony.

All these kinds of knowledge can be reliably and validly assessed by an examination. Indeed, most of them can now be easily assessed electronically online, and marked completely reliably in the same way. Facts of this kind are the necessary basis of any higher-order types of understanding, but judging whether a student knows them is no more than the beginning of a fully valid assessment.

As soon as validity is sought, one is drawn into the case against exams. Judging whether a student has understood how to apply a specific bit of algebra either to a higher-level field of mathematics, or to a real-world problem in physics or economics, is not straightforward. The time available in a typical exam may not allow the complexities of genuine understanding here to be explored.

Out of these concerns came the growth of course-work assessment. Originally, that was confined to subjects where an inescapable practical core had to be included – for example, art, music, and metalwork. But it then was extended to laboratory work and fieldwork in science, and to similar tasks in social subjects and literature. Now the only main subjects in Scotland that have no course-work assessment are the various versions of mathematics. The case for this kind of extension of assessment is strong. Worthwhile understanding of any subject in the curriculum requires not only a sound basis in facts, but also the higher-order understanding that puts the facts

together. This can best be assessed by projects, experiments, and essays.

Examinations or course-work?

Is it one or another? The second major issue in the present debate is between different ways of assessing students.

A valuable summary of the evidence was provided as part of the UK government's [review](#) in 2013 of assessment in school exams in England, Wales and Northern Ireland. Because that context is more similar to the Scottish context than assessment systems anywhere else, the conclusions it reached are relevant to the current Scottish debate.¹

None of the arguments in the review in favour of course-work shows, of itself, that examinations are redundant. Both modes have strengths.

In an extensive review of research on course-work and exams in higher education, [Richardson \(2015\)](#) notes the advantages of each.²

The research abstract says;

In the UK and other countries, the use of end-of-module assessment by coursework in higher education has increased over the last 40 years. This has been justified by various pedagogical arguments. In addition, students themselves prefer to be assessed either by coursework alone or by a mixture of coursework and examinations than by examinations alone. Assessment by coursework alone or by a mixture of coursework and examinations tends to yield higher marks than assessment by examinations alone. The increased adoption of assessment by coursework has contributed to an increase over time in the marks on individual modules and in the proportion of good degrees across entire programmes. Assessment by coursework appears to attenuate the negative effect of class size on student attainment. The difference between coursework marks and examination marks tends

to be greater in some disciplines than others, but it appears to be similar in men and women and in students from different ethnic groups. Collusion, plagiarism and personation (especially 'contract cheating' through the use of bespoke essays) are potential problems with coursework assessment. Nevertheless, the increased use of assessment by coursework has generally been seen as uncontroversial, with only isolated voices expressing concerns regarding possible risks to academic standards.

All of the findings of the research would be as relevant to the senior years of secondary school as to universities.

Marking and bias

Although the use of course-work may solve some problems of validity, it exacerbates the difficulty of reliability, and also potentially introduces a new version of invalidity, the bias of markers. Fortunately, the ways of dealing with both of these are similar.

Reliability in marking essay-type questions is a notorious [problem](#), but three [procedures](#) can help to mitigate it. One is to have several people mark independently, and then hold a meeting to resolve differences. Essentially this is what is done for marking research dissertations in UK universities, but it would be wholly unaffordable to be used routinely for marking course-work in schools. A version of this is what has been called a 'community of practice', as might gradually develop from a group of teachers who have worked together for years, and who have well-developed ways of inducting new teachers into the group. The problem is that, although the community might agree with each other, it might drift away from practice in other communities. So, there have to be clear marking schemes, careful training of markers in them, and re-marking of a sample of scripts anonymously by external inspectors.

¹ [A focus on teacher assessment reliability in GCSE and GCE: Summary - GOV.UK \(www.gov.uk\)](#)

² [Coursework versus examinations in end-of-module assessment: a literature review: Assessment & Evaluation in Higher Education: Vol 40, No 3 \(tandfonline.com\)](#)

All of this is standard practice by the Scottish Qualifications Authority, and has been since the beginning of school exams.

Bias does not arise where the marking is anonymous, but the sheer scale of course-work assessment has required that most of it is marked by students' own teachers. This is evidenced by the Alternative Certification model which has been operating in Scottish schools in 2020/2021. It is then almost impossible for the marking to be anonymous. Even if names are removed from scripts, hand-writing and other idiosyncrasies are recognisable to a teacher who has known a young person for many years. Few subject departments in schools would be large enough for marking always to be done by teachers who do not know the students individually.

There is strong [evidence](#) of bias in marking that is not anonymous, which takes two forms.³

One is bias against or in favour of social groups. These patterns vary over time, but at present would be likely to be harmful to young people in deprived social circumstances.

There is also bias in favour of certain groups. [Stereotypical views](#) such as that Chinese pupils are good at maths, girls are better at foreign languages, or boys are more interested in physics, exist amongst teachers as well as in the general population.⁴

The other main kind of bias in non-anonymous marking is what has been called the 'halo effect', in other words expectations based on the individual candidate's past performance. Students who did well in the past are expected to do well again. The halo effect may interact with other kinds of bias. For example, teachers' beliefs about the mathematical aptitude of boys and girls has been [found](#) to be one factor in explaining higher mathematical achievement by boys.

3

https://journals.sagepub.com/doi/abs/10.1177/0004944116664618?casa_token=VZaYSecRGA4AAAAA%3AhG-TFyfG8unHldpBbyWt2vrUYaS9cUYga7PLPRpjefBoGaZlv3HknX9-SxiY5ITLk9AtW9WL4gee&

4

<https://www.tandfonline.com/doi/abs/10.1080/13803611.2018.1548798?c>

There are further problems. Appeals against grades are much more difficult to arrange when a large component of the assessment is by teachers' marks. To enable students to appeal against these, there has to be a mechanism by which external assessors re-mark the work, because in most schools there is no-one sufficiently independent of the original marking process. But having to invoke external assessors in this way means that students who appeal are assessed by potentially different standards to those who did not appeal. One partial safeguard against such distortion is to allow appeals only where the teacher marks are not wildly different from the externally assessed exam, but that then has the incidental effect of placing greater emphasis on the exam.

There is also the problem that teachers can be put under pressure to award high grades, a risk that has been [shown](#) to be acute in the USA where teacher grading plays a much greater role than in European systems. Anonymous marking is thus a good defence of teachers' professionalism.⁵

There have been reports this year of this happening in England; the Labour Party's Education Spokesperson has commented on this.

The growth of course-work threatens the integrity of school assessment because marking so much course-work externally and anonymously is simply not feasible. Externally marked exams are the best way of making sure that an A in one school means the same as an A in another. Basing assessment predominantly on course-work that is marked by teachers would be unavoidably unfair.

Conclusions

Examinations are not good at everything. Course-work can assess skills and understanding that examinations simply cannot assess.

[asa_token=YLtdNozk_fAAAAAA:aPUAcYd2nihexSZM6Rr8zwJCTRJu naOIMIBgqD9E0CBUvy_zE2Fb0UozDxc-YdPx-Yha9etYaDal](https://books.google.co.uk/books?hl=en&lr=&id=8e5IDwAAQBAJ&oi=fnd&pg=PP1&dq=Negotiating+Opportunities&ots=fGLCayGDQ&sig=tkd2iLY6ndhuk3e8uF8ZcCZrFFw#v=onepage&q=Negotiating%20Opportunities&f=false)
5

<https://books.google.co.uk/books?hl=en&lr=&id=8e5IDwAAQBAJ&oi=fnd&pg=PP1&dq=Negotiating+Opportunities&ots=fGLCayGDQ&sig=tkd2iLY6ndhuk3e8uF8ZcCZrFFw#v=onepage&q=Negotiating%20Opportunities&f=false>

Examinations, however, can assess important skills too, especially relating to memory and focus, and the process of preparing for them can be an important way in which knowledge can be consolidated.

Moreover, the greater the use of course-work the greater will be the need for internal marking of that course-work, i.e. that is, marking by students' own teachers. This dependence on internal marking, even when associated with elaborate systems of moderation runs the risk of introducing new forms of bias into the assessment and certification of young people. That bias can be mitigated by, for example, clear marking criteria, the mentoring of new markers by experienced teachers and by re-marking a sample of work anonymously. Bias, however, arising from internal marking probably cannot be eliminated.

External assessment at the point when students leave school has been at the heart of universal secondary schooling for decades; it is the [norm](#) in school systems across the world, including in countries, such as Finland, which generally place less emphasis on formal assessment than Scotland.

Of course, in Scotland, pupils do not sit exams at the point at which they leave school, but continually in S4, 5 and 6. As the leaving age is 16, there has to be some assessment offered in S4 to people who choose to leave at that point, which has led to the three years of different exams and a situation where pupils are potentially over assessed. There are different routes that could be used to address this issue:

- The school leaving age could be raised, perhaps – as in England – with the option of extra years being taken a college rather than in school;
- A High School Certificate could be offered to those choosing to leave at S4, with a combination of external and internal assessment;
- Greater integration of vocational qualifications, perhaps building on the

success of the Foundation Apprenticeships.

Scotland's systems of assessment have grown gradually, evolved, often subject to little systematic planning. The debate occasioned by the pandemic and reports such as the OECD's is valuable. Raising doubts about examinations ought not to be a reason for abandoning them. The debate can show that, alongside a range of other kinds of assessment, examinations marked anonymously remain the best way of guaranteeing that standards are consistent. They are thus a necessary part of ensuring that opportunities are equal for every student.