

Rapid Review of National Qualifications Experience 2020

Professor Mark Priestley, Dr Marina Shapira, Dr Andrea Priestley,
Michelle Ritchie and Dr Camilla Barnett

Faculty of Social Sciences, University of Stirling

Final Report,
September 2020

Contents

List of acronyms and initialisations	3
Executive summary.....	4
Summary of findings	4
Summary of recommendations	5
The Review	6
Context	6
Remit	7
Methodology.....	7
Findings	10
General overview of findings	10
Estimation and local moderation.....	13
SQA Guidance.....	15
Local Authority support	16
Coursework	17
National Moderation	17
The approach to moderation	19
Issues arising from the moderation process	23
Post Certification Review and Appeals	28
Equalities.....	31
Communication and transparency	34
Impact on young people	37
Events following cancellation of the exam diet.....	37
Equity	38
Wellbeing	38
Transitions	39
Future exam arrangements	39
Involvement of young people in meaningful engagement.....	40
Longer term impact of this experience on young people.....	40
Discussion	41
Recommendations	45
Summary of recommendations	45
Rationale for recommendations	45

References	49
Appendices.....	51
Appendix A: Position papers submitted to the review	51

List of acronyms and initialisations

ACM	Alternative Certification Model
ADES	Association of Directors of Education
BOCSH	Building Our Curriculum Self Help Group
COSLA	Convention of Scottish Local Authorities
CRIA	Children's Right Impact Assessment
CYPCS	Children and Young People's Commissioner Scotland
EIS	Educational Institute of Scotland
EQHC	Equalities and Human Rights Commission
EQIA	Equalities Impact Assessment
LA	Local Authority
MSA	Modern Studies Association
NASUWT	National Association of Schoolmasters/Union of Women Teachers
NPFS	National Parent Forum Scotland
PCR	Post Certification Review
SAGT	Scottish Association of Geography Teachers
SATE	Scottish Association of Teachers of English
SLS	School Leaders Scotland
SQA	Scottish Qualifications Authority
SPD	Starting Point Distribution
SSTA	Scottish Secondary Teachers' Union

Executive summary

Summary of findings

- This report draws upon a range of evidence, including stakeholder testimony (generated in panel and individual interviews) and analysis of relevant documentation (including government and SQA emails).
- SQA, the government, local authorities and schools faced an extremely difficult set of circumstances, within which there were no easy solutions. In this context, a workable system for qualifications, the Alternative Certification Model (ACM), was developed. This was based on three core principles and four stages.
- All parties involved in the process were found to have acted with integrity, with the best interests of students in mind.
- Respondents (teachers, lecturers, head teachers and local authority officials) generally found that SQA guidance was clear and useful.
- The generation of estimated grades, while clearly undertaken with integrity in the majority of centres, has been subject to variation (in the types of evidence available, the processes followed for internal moderation and the support given by local authorities), which has impacted on reliability and consistency of assessment at this stage.
- The statistical approach to moderation could have been more transparent earlier in the process, and moreover it has led to anomalies in grade adjustment, especially at the level of subject cohorts within centres and individuals.
- There is widespread criticism by respondents of SQA for a perceived lack of transparency and a failure to engage in participative development of solutions with stakeholders.
- While the application of the appeals process offered an in-principle technical solution to address these anomalies, it paid insufficient attention to the severe impact on those students obliged to undergo it (in terms of mental health and wellbeing, missed opportunities to transition into Higher Education, etc.).
- Principles relating to what data is appropriate to be held by certain organisations at certain points in time.(i.e. SQA, the Scottish Government), which make perfect sense in normal times (e.g. arrangements around data sharing), appear to have impeded the development of actions that might have led to an earlier anticipation and mitigation of subsequent problems.
- The equalities implications of an over-reliance on a statistical approach, premised on comparison with historical cohort data, had been raised repeatedly from April onwards, but seem to have been under-emphasised by both the government and SQA until late in the process.
- Many stakeholders believe that, subsequently, opportunities were missed (or dismissed) to engage in qualitative moderation of the statistical process (e.g. sense-checking of anomalous cohort patterns by local authorities).
- There has been an erosion of trust/confidence in SQA amongst teachers and young people, and damaged relations in some cases between young people and their teachers.
- Communications (with professionals and with young people and their families) has been a constant source of criticism.
- Our overall assessment is that, despite the extremely difficult environment for decision making, there are points in the process where different decisions may

have led to better outcomes and at least partially avoided the controversy that ensued in August 2020. Of course, we are making this observation with the benefit of hindsight, thus our primary intention is to illustrate how the system can benefit from lessons learned in 2020 to avoid a similar predicament in 2021.

Summary of recommendations

1. Suspension of the National 5 examinations diet in 2021, with qualifications awarded on the basis of centre estimation based upon validated assessments.
2. The development of a nationally recognised, fully transparent and proportionate system for moderation of centre-based assessment.
3. The development of more extensive approaches to collaborative decision making and co-construction by professional stakeholders of assessment practices related to National Qualifications.
4. A commitment to embedding equalities in all aspects of the development of qualifications systems.
5. The development of more systematic processes for working with and engaging young people, as stakeholders and rights holders in education.
6. The development of a clear communications strategy, co-constructed with stakeholders, to ensure that the extraordinary arrangements for 2021 are as fully as possible understood by all parties.
7. A review of qualification appeals systems, including consideration of the rights and roles of young people, in the context of the incorporation of the UNCRC into Scottish law.
8. The commissioning of independent research into the development and application of the 2020 ACM, involving full access to anonymised attainment data and the statistical algorithms used to moderate grades.
9. The development by SQA and partners of digital materials and systems for producing, assessing and moderating assessment evidence, to ensure that operational processes for gathering candidate evidence for appeals is less reliant on paper-based systems.

The Review

Context

In March 2020, in the face of the disruption caused by the COVID-19 pandemic, and following the cancellation of the 2020 examinations diet, the Scottish Qualifications Authority (SQA) was commissioned by the Deputy First Minister and Cabinet Secretary for Education and Skills, John Swinney, to provide alternative means for awarding qualifications, based on three principles.

- fairness to all learners;
- safe and secure certification of qualifications, while following the latest public health advice;
- maintaining the integrity and credibility of the qualifications system, ensuring that standards are maintained over time, in the interest of learners.

The SQA subsequently developed the Alternative Certification Model (ACM), comprising the following steps:

- Step 1 — Estimates
- Step 2 — Awarding
- Step 3 — Results and certification
- Step 4 — Appeals

The release of results on 4 August, 2020, was accompanied by controversies and considerable media attention, centred around issues of equity. Subsequently, Professor Mark Priestley of the University of Stirling was commissioned by the Scottish Government to lead an independent review of the processes through which qualifications were awarded.

Professor Priestley established the following research team to undertake the review:

- Professor Mark Priestley – Principal Investigator
- Dr Marina Shapira – Co-Investigator (with responsibility for the statistical aspects of the review)
- Dr Andrea Priestley – Co-Investigator (leader of the strand investigating the experiences and perspectives of young people)
- Michelle Ritchie – Research Assistant
- Dr Camilla Barnett – Research Assistant

Additionally, the Review employed two independent external reviewers, to provide advice on process and preliminary findings and to review the final report.

- Professor Robert Davis – Professor of Religious and Cultural Education, and Director of the Robert Owen Centre for Educational Change at the University of Glasgow
- Associate Professor Gill Wyness – Associate Professor of Economics, and Deputy Director of the Centre for Education Policy and Equalising Opportunities (CEPEO) at the UCL Institute of Education.

Work on the review commenced on 17th August, with an interim report to the Deputy First Minister on 15th September and completion of a final report on 30th September.

Remit

The following remit was agreed with the Scottish Government:

The review will include considering evidence, providing commentary and recommendations around the following themes. A focus on those issues which are most pertinent to consideration of awarding methodology in 2021 if there is further significant disruption to learning and teaching and/or the cancellation of exams is key:

1. Events following the cancellation of the 2020 examination diet.
2. Advice and support given by SQA and Local Authorities to awarding centres on determining and quality assuring of estimates
3. Approaches to the gathering and quality assurance of teacher/lecturer estimates, including where possible feedback from teachers/lecturers/Directors of Education, prior to submission to SQA about the perceived rigour in the evidence base for making estimates, e.g. prelim marks, classwork, summative and formative assessment until the schools closed in March. This will include consideration of local quality assurance approaches taken by centres and Local Authorities to aid estimation; the conclusions reached by centres about estimated grades; and decisions about whether or not to share estimates with learners at that time.
4. Exploration of alternative approaches to grading and moderating national qualifications in the context of the disruption caused by Covid-19, that would maintain standards and the credibility of qualifications in Scotland and deliver public confidence.
5. Impact on young people (and their families) who did not receive what they believed their estimated grade submitted to be.
6. Feedback from teachers/lecturers on the estimation process and the moderated grades which were awarded on 4th August
7. Consideration of the post certification review process as a means to address the issues in 5 and 6 above.
8. Confidential draft report to ministers on findings by 15th September 2020.
9. Final report published by 30th September 2020.

Methodology

The Review has been treated as a research project, involving the collection of primary data and review of secondary data, as well as due consideration of ethical issues. The following approaches were used to generate a wide range of data to inform the review.

1. Review of documentation, including published materials, emails and other communication between Government officials, SQA, local authorities and other stakeholders.

2. Panel discussions with key stakeholders, including young people and parents/carers, teachers, senior school leaders, local authorities, SQA and government officials.
3. Interviews with key individuals (e.g. SQA, academics with specialism in assessment/statistics).
4. Analysis of short position papers, submitted by stakeholder organisations. These are listed in Appendix A of the report.
5. Consideration of how moderation was applied to centres' estimated grades. This included examination of processes set in place by local authorities for supporting and moderating grades at the estimation process, and consideration of the national moderation processes applied by SQA. The time scale and resources available for the Review have not permitted an in-depth analysis of the statistical approach used for moderation, and we have not had access to the algorithms or anonymised datasets necessary to undertake such a review¹.

The primary source for recruitment of panel members was national stakeholder groups, with a focus on people with direct experience of the awarding process. These groups nominated people for the panel in question. With the exception of one group of teachers (see below), and a small group of parents/carers, where contact was facilitated by the parents' advocacy group *Connect*, we have not been directly involved in selecting participants for panel discussions. We note here that the views expressed by panel members may not always agree with one another, or with the SQA view of events. We report stakeholder views presented in our evidence as perceptions of the process. It is important to do so, as these perceptions provide a clear indication of how the process was experienced by different people, thus providing insights into how the system might be operated differently in the coming year, when COVID-19 is likely to remain a factor. The following illustrates the range of stakeholders engaging with the review through panel discussions.

Discussion Panel	Number of participants
Children in Scotland	7
Scottish Youth Parliament	5
Children & Young People's Commissioner Scotland	6
Student Partnerships in Quality (Sparqs)	3
SQA: Where's Our Say?	2
Parents (Connect and the National Parent Forum of Scotland)	9
Head teachers	9 (+1 written response)
Independent Sector Teachers	5
Non-affiliated Teachers Group	9

¹ We initially expressed, to the government, an interest in conducting analysis of the dataset using the algorithm employed for grading. It was made clear to us that, while the government would make a request of SQA for the data, the request was unlikely to be granted in the short time scales of a rapid review. Subsequently, and due to challenges in resourcing and gaining university ethical clearance for such an analysis within the timeframe, we did not pursue this option. We do, however, believe that such an analysis is necessary to gain a full understanding of the processes undertaken, hence our recommendation 8.

College Lecturers	7
Teaching Unions	8
Subject Associations	8
Academics	8
Local Authorities	5
Scottish Government	4
SQA Technical	4
SQA Policy	6
SQA Practitioner	7
	Total 112 (109 individuals, accounting for SQA participants who took part in more than one interview panel)

All panels and interviews were conducted via Microsoft Teams and recorded with the permission of participants, who underwent a formal process of informed consent. The research was conducted in accordance with the 2018 BERA Ethics guidelines², with due regard for the human dignity and safety of all participants, following approval by the General University Ethics Panel at the University of Stirling.

Participants were guaranteed confidentiality as far as is possible in group interviews. We have not attributed any statements made in the interviews to individuals and/or particular schools and local authorities. Participants in group interviews were asked to refrain from identifying co-participants or divulging details of others' testimonies. Scottish Government and national agency staff were not present at panel discussions and interviews, and they will not have access to primary data (e.g. interview recordings and transcripts) or details about participants other than that which is public knowledge (i.e. named individuals publicly representing organisations). Different stakeholders were interviewed in discrete groups, avoiding, for example, a situation where teachers are nervous to testify frankly in the presence of local authority officers, or young people in the presence of teachers. All interviews and panel discussions were led by university researchers, who are independent of the qualifications system and processes.

We were cognisant of the need for additional sensitivity in the case of some groups of young people regarding confidentiality and anonymity, and access to the technology required to participate in the discussions. We were also aware of the potential for this research to cause emotional distress for some participants, who have been disadvantaged in the granting of awards and subsequent destinations (e.g. missed university places). The researcher leading this strand, Dr Andrea Priestley, is highly experienced at working with young people, including those in care and other vulnerable situations, and was able to address these issues. As all young people were representing third party organisations, they could usually also receive support from those organisations. A representative from the young people stakeholder organisation was permitted to attend the applicable session with the

² <https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018>

permission of all participants, in order to provide support for the young people. Where applicable, we referred young people to CYPCS for additional support.

Findings

This section of the report commences with a brief overview of the findings, before engaging in a more detailed analysis of the data related to a number of key themes. These are:

- Estimation and local moderation
- National Moderation
- Appeals
- Equalities issues
- Communication
- Impact on young people and their families
- Impact on teachers and lecturers

Each of the abovementioned sections provides the following: 1] an overview of findings; and 2] some discussion of findings.

General overview of findings

There was a general acceptance amongst the majority of respondents, with which we concur, that the SQA and government were faced with an impossible situation – a ‘monumental task’ (Learned Societies position paper) of moving from a well-established system of awarding qualifications based on exams and formal coursework assessment, to a very different system based on teacher estimates. This was exacerbated by the huge difficulties associated with being required at short notice to work remotely from home. Respondents generally recognised the professionalism, hard work and dedication brought to the task by SQA, in the face of formidable issues to resolve in a pressured and rapidly emerging context over a limited timescale. The following extract is typical of sentiments widely expressed in interviews and position papers.

After the cancellation of the 2020 exam diet, announced in March, and given the time constraints, it should be noted that the SQA were put in an exceptionally challenging position. It was very unlikely that they would be able to develop a solution that could replicate the current assessment conditions and system. (ADES position paper)

Moreover, SQA was faced with considerable capacity issues in moving to a system very different to what had previously been offered. Panel interviews with SQA painted a picture of the challenges involved in bringing in external expertise in statistics (government secondments and private agencies) and developing a new system to receive estimate and rank information from centres. It is widely accepted that no system could be perfect under these circumstances. Respondents generally agreed that there was no feasible alternative to cancelling the exams diet (including parents’ groups (e.g. NFPS position paper), and were supportive of this decision. Evidence presented to the review indicates a rapidly changing situation, where

decision making was exceptionally difficult in challenging circumstances, and often undertaken as a reaction to unpredictable political and media commentary. The following brief timeline illustrates clearly how emergent events effectively turned decision-making into an extremely uncertain process. The examinations diet was still planned almost up to lockdown; on 18th March, the Scottish Government and SQA joint statement on the *Coronavirus, and impact on August 2020 certification* stated:

The Qualifications Contingency Group agreed that every effort should be made to ensure schools remain partially open to allow Senior Phase pupils to complete learning and be able to submit coursework, in addition to being able to open as examination centres during the diet, should medical and scientific advice allow.

On 19th March, the examinations were cancelled by the Government. On 20th March, schools closed and SQA called on schools to collate evidence, including getting coursework completed. On 23rd March, the First Minister announced young people should not attend school to complete coursework. These examples illustrate the difficulties in making decisions at this stage, when the COVID-19 pandemic had many unknown dimensions, when concerns about safety were paramount and when the situation was changing daily.

We have seen little criticism of the three principles underpinning the process:

- fairness to all learners;
- safe and secure certification of qualifications, while following the latest public health advice;
- maintaining the integrity and credibility of the qualifications system, ensuring that standards are maintained over time, in the interest of learners.

In general, the majority of stakeholders support the notion that SQA have acted with integrity to realise these principles laid out by the government at the outset, in the face of very challenging timelines in an unprecedented situation. Some respondents, however, have questioned the subsequent realisation of the principles in the ACM, and particularly whether the first principle was ultimately undermined by an emphasis on the third. We will return to this issue later in the report.

We have found more disagreement with the decision not to continue with marking and submission of coursework. Many respondents would like to have seen more consideration of how coursework could have been completed, marked and used to contribute to grading/estimation. Again, we will return to this issue later in the report.

Despite this broad in-principle support for the stance laid out by SQA and the government, the widespread view of most respondents in our review is that many of the subsequent problems encountered could have been mitigated had different decisions been made. We wish to emphasise here that many of these observations are made with the benefit of hindsight; it may not have been possible to act differently, given the circumstances, and it is also not always clear that different forms of action advocated would have made a huge difference. Nevertheless, one of the purposes of this review is to learn from the experience of 2020, given the high

likelihood of continued COVID-19 disruption in the coming year, and reflection on the issues that affected the 2020 qualifications is an important part of this learning.

In particular, the following issues have surfaced:

- the generation of estimates, while clearly undertaken with integrity in the majority of centres, has been subject to variation (in the types of evidence available, the processes followed for internal moderation and the support given by local authorities), which has impacted on reliability of assessment at this stage;
- the statistical approach to moderation could have been more transparent earlier in the process, and moreover it has led to anomalies in grade adjustment, especially at the level of subject cohorts within centres and individuals;
- there is widespread criticism by respondents of SQA for a perceived lack of transparency and a failure to engage in participative development of solutions with stakeholders;
- while the application of the Post Certification Review (PCR) process offered an in-principle technical solution to address these anomalies, it paid insufficient attention to the severe impact on those students obliged to undergo it (in terms of mental health and wellbeing, missed opportunities to transition into Higher Education³, etc.);
- principles relating to what data is appropriate to be held by certain organisations at certain points in time (i.e. SQA, the Scottish Government), which make perfect sense in normal times (e.g. arrangements around data sharing), appear to have impeded the development of actions that might have led to an earlier anticipation and mitigation of subsequent problems.
- the equity implications of an over-reliance on a statistical approach, premised on comparison with historical cohort data, had been raised repeatedly from April onwards (e.g. CYPCS and NASUWT position papers), but seem to have been under-emphasised by both the government and SQA until late in the process;
- many stakeholders believe that, subsequently, opportunities were missed (or dismissed) to engage in qualitative moderation of the statistical process;
- respondents reported an erosion of trust/confidence in SQA amongst teachers and young people, and damaged relations in some cases between young people and their teachers.

We note here that SQA has stated to us that there is no regret in respect of the moderation approach used this year (in terms of its technical application), but that the regret lies in the fact that the PCR process was not allowed to run its course, as this component was designed to deal with the sorts of problematic results that generated such an intense political and media focus after results day on 4th August. SQA has stated that the case for moderation was clear and unequivocal – and should be seen in the context of commission from Ministers and the unprecedented position faced by the system, including the time constraints within which they were working. Evidence from discussions with SQA indicates that the organisation accepts that the statistical approach to moderation used in 2020 would not be acceptable to

³ As reported by young people and the organisations representing them.

the public in future, and there should be more emphasis on a qualitative element to moderation, with a more active role for schools. We have also seen, in our discussions with SQA, some agreement that messaging is important, and that better communication around aspects of the ACM – in particular warning schools and students that estimates would need a high level of moderation that might result in individual and cohort level anomalies, and clearer messaging that the PCR stage was an integral rather than a bolt-on part of the process – might have obviated a great deal of the furore that erupted after results day. SQA had clearly debated the pros and cons of releasing this information, and told us that the decision not to share more details about the implications of the model was based on a perceived need to avoid undue stress for students, parents/carers and teachers.

These issues are addressed in more detail in the following sections.

Estimation and local moderation

Perceived strengths	Perceived weaknesses
<ul style="list-style-type: none"> • SQA established a system that obtained estimates from every centre for every candidate and subject by the specified deadline; • Clear guidance for centres from SQA (with caveats). • Dedicated approach by teachers and lecturers. • Some excellent practice in some local authorities to support and moderate estimation. 	<ul style="list-style-type: none"> • Difficulties accessing evidence. • Variation in local moderation contexts and practices, with some limited input from some local authorities. • Complexity of enhanced banding scale and ranking processes. • Over-estimation and/or inaccurate estimation in some centres. •
Overall assessment	
<p>Estimation and/or centre-based assessment would be greatly enhanced by the development of systematic and consistent local moderation processes. While this moderation is applied locally, it requires national development by SQA working collaboratively with stakeholder groups such as local authorities. Moderation should extend to the development of validated sources of evidence, and internal and external verification of assessment.</p>	

Estimation by centres is the linchpin of the ACM. In this section we address some key aspects of this, including guidance, support for local moderation and the place of evidence in the process, including coursework. The evidence from our review suggests that the estimation process was taken very seriously by schools and colleges, and involved a great deal of professional integrity, dedication and hard work by practitioners, working remotely from their usual workplaces, and experiencing formidable difficulties in relation to evidencing estimation. Teachers and head teachers have reported two sets of difficulties: 1] different approaches to progression from subject to subject made a consistent approach across centres problematic; 2] difficulties in accessing evidence, particularly coursework (either in cupboards in school or already sent to SQA). According to local authority evidence presented to the review (ADES position paper), some centres over-estimated; this was not due to teachers deliberately inflating grades, but was instead to some extent

a consequence of an inability to do robust moderation (citing workload concerns, lack of LA capacity/expertise, lack of evidence) and a desire to assess how each individual would perform on the day of examination, given that all went well. We note here that we have seen no evidence of accountability systems leading to grade inflation grades – for example teachers experiencing pressure to enhance their estimates. Indeed, we have seen evidence of the converse, as schools were cautious in their allocations, and as local authorities in many cases moderated estimates downwards. This is encouraging given previous research indicating that cultures of performativity may lead to grade inflation in school-based assessment (e.g. Cowie, Taylor & Croxford, 2007; Priestley & Adey, 2010).

Local authorities, head teacher and teachers have pointed to a sense of grievance in many schools that teacher estimates are not trusted, exacerbated in the view of ADES by a lack of consistency in communications regarding the balance in the ACM between estimation and moderation. It is likely that stronger messages about the need for some form of national moderation would have been helpful at the outset. Existing research (e.g. Everett & Papageorgiou, 2011; UCU, 2015; Wilson, 2015; Wyness, 2016; Anders, et al. 2020; Murphy & Wyness, 2020) indicates that estimates (or predicted grades) have tended to be historically inaccurate (or at least different from eventual exam results), something backed up by SQA's own data (SQA 2020). This literature indicates clear patterns of over/under-estimation associated with particular demographic characteristics (e.g. students from disadvantaged backgrounds and state schools are more likely to be over-predicted whilst those in independent schools receive more accurate predictions). Significant patterns of divergence – between estimation in 2020 and historical patterns of attainment – should have come as no surprise, and yet we were told by SQA that, until the teachers' estimates were analysed after submission on 29 May, there were 'hopes' that teachers' estimates might be close to historical grades and therefore no (extensive) moderation would be needed⁴.

We saw some grievance in LAs that higher estimates were not necessarily the result of over-estimation, but rather a more accurate picture of student achievement than that provided by exams – an evidenced-based approach, which focuses on more than just exam performance, and ensures that the achievements of those pupils, for whom an examination is a barrier, are recognised. Many students felt frustrated that their wider achievement and contribution to the school was not recognised in their awarded grades. They would like to have seen more diverse forms of assessment, which captured their efforts. Students who did not agree with their estimated grade and who weren't supported in the appeals process by their school felt particularly aggrieved and betrayed by their school, when they had contributed to wider school life (e.g. charity work, sports teams, prefect duties). The SQA Future Report 2018 (Young Scot Observatory/SQA, 2018) committed the organisation to working with young people to co-design 'a new approach to assessing competence in the skills highlighted in the report, particularly in the area of life skills'. In this vein, young people would have liked a more holistic approach to the ACM.

⁴ We also note here that SQA provided additional guidance to centres and historical estimate and results data, for the purpose of allowing centres to review at dept level whether they got it about right or that they had a tendency to over or under estimate.

SQA Guidance

With some strongly expressed exceptions (notably teachers in the independent sector), the majority view of our respondents is that the SQA guidance for centres on estimations was clear and helpful. One subject association stated that the guidance was clear, but would have been useful earlier⁵ (MSA position paper). In our view, the SQA guidance on estimation provided clear and concise advice that identified key issues – evidence, past centre performance, et cetera. It was clear that additional prelims should not be set (although we note that the parents panel claimed that some schools allowed pupils to sit second prelims) and there was no need to mark coursework normally externally assessed (although this introduced some ambiguity as to how this could be then used to inform estimation). The online training provided by SQA to address unconscious bias was well-received on the whole.

According to some respondents and our own reading of the guidance, it had some shortcomings, perhaps understandable given the timing and circumstances of its production. First, while the paper suggested a wide range of evidence, it did not explicitly preclude limiting estimation to the prelim grade (which some schools seem to have done). The sign off system provided only a limited form of moderation, and a more comprehensive set of guidance around local moderation would have improved school-based processes for estimation. A subject association, reflecting a general sentiment that teachers would like more engagement with SQA in the development of processes for awarding qualifications, stated:

It was extremely disappointing, but not unexpected, that the SQA chose not to engage with any professional organisations during the development of the estimate process⁶. (SAGT position paper)

Moreover, it was noted by some (e.g. the independent schools panel) that the subsequent Post Certification Review documentation was more comprehensive – and more specific on what constitutes evidence, including coursework. Some respondents believed that the guidance had changed over time, creating difficulties; in the words of one respondent, ‘moving the goal posts’ (head teacher interview).

The enhanced banding scale and ranking processes were found to be complex and stressful by many teachers, including the subject associations (e.g. SATE) and the teacher unions.

The process was made more complicated, in our view, by the SQA’s insistence on the sub-dividing of existing bandings and the creation of rank orderings. (EIS position paper)

The refined grade and ranking system, however, was quite complex and was often difficult for staff to quantify. (Colleges Scotland position paper)

⁵ SQA provided communications to centres on estimation on the 2nd April and 20th April.

⁶ SQA states that they consulted stakeholders on this, including subject associations.

We note here that some potential problems with the estimation process do not appear to have been thought through in detail. Some were addressed by inter-school collaboration, and local authority support, but this seems to have been variable.

1. Difficulties in accessing evidence (e.g. reported in the SSTA and SAGT position papers, head teacher panel and several teacher panels), which in turn made estimation difficult.
2. School size: 1] in small schools, not enough subject teachers to moderate each other's work or a lack of teachers with a specific expertise (these issues are exacerbated where staff are inexperienced, e.g. a new member of staff as the sole subject teacher in a department); 2] in large schools with many classes (e.g. maths), teachers do not know all students, and it is difficult to rank them (reported in several of the teacher panels)
3. College sector specific problems (e.g. one course could be spread across different campuses; lack of previous knowledge about students; lack of previous attainment data for adult students – reported in the college lecturer panel).

Again, more developed guidance on local moderation, a greater recourse by SQA to local expertise in schools, colleges and local authorities and clearer messaging about the necessity of national moderation may have mitigated these issues.

Local Authority support

The role of the local authorities appears to be crucial in respect of local moderation of the estimation process⁷. We have found evidence of highly variable approaches to local moderation (e.g. SLS position paper, analysis of LA documentation) – in some cases exemplary, in other more minimal.

In some LAs, we have seen rigorous approaches to supporting estimation, including guidance on evidence and cohort historical comparison, follow-up processes to query high estimates, and use of data to account for previous concordance between estimates and grades. In some LAs, analysis of results was undertaken post-award. In at least two of the examples we examined, this analysis quickly allowed anomalies in grading at a cohort level to be quickly identified. One Director of Education told us that an analysis of results in the LA took only one hour and forty minutes, with the implication that a national analysis of results, pre-award, would have been a straightforward exercise that would quickly have identified anomalous results, making qualitative moderation subsequently possible. Some LAs provided direct support to schools (e.g. those with low capacity, such as one teacher departments) and supplementary data on historical attainment and concordance patterns. Oversight allowed errors to be corrected at the local level, prior to estimates being submitted. In at least one LA, grades were adjusted by the LA prior to submission. Some LAs established a common process of estimation/moderation for schools to follow. In some cases, systems were developed in collaboration with schools, with occasional evidence of parental consultation. In one case, an estimation tool was

⁷ We note that nearly 100 centres are not within LA jurisdiction.

produced, which facilitated estimation and allowed analysis of post-estimation trends in the data by schools.

In other LAs, guidance was more limited (e.g. supplementary guidance on processes or even simply reiterating SQA guidance). In these LAs we saw little or no evidence of checking results patterns prior to submission. Even in the best practice cases, LA moderation could be limited in its effects; in one LA with extensive provision for supporting and moderating estimation, it was reported to us that schools were able to disregard LA advice press on with estimations (conducted by teachers and signed off by HTs).

In some cases, LAs stated that they submitted rationales for variance to SQA. Others collected data, and waited to be contacted by SQA – being concerned that moderated grades would be subject to arbitrary moderation by the national moderation process. According to one Director of Education, “The additional step of asking the SQA to contact Directors [of Education] to discuss any anomalies would have helped prevent this.”

We note that variance in approaches to moderation by LAs does not seem to be exclusively linked to size/capacity – some of the most thorough systems were evident in small LAs.

Coursework

Cancellation of coursework, albeit discussed and agreed with key stakeholders, has been contentious, with many stakeholders suggesting that a greater effort could have been made to assess it, to both contribute to final grades and to form a more robust evidence base for estimation (e.g. ADES position paper, NPFS position paper). For example:

There was potential for further discussion and thought around the use of coursework and assessments, much of which SQA already had. Reasoning for not using centred around the confidence of a carrier being able to distribute to markers and return. Should this have been investigated further? (ADES position paper)

Having considered the evidence, we accept that this was a pragmatic decision made for a combination of good reasons. These include: equity (while some students had completed coursework, in many cases it was not complete); logistics (getting coursework from schools to markers in face of disruption to courier services); and safety concerns (due to fears about spreading the virus through distributing and handling packages).

National Moderation

Perceived strengths	Perceived weaknesses
1. SQA designed a moderation system to adjust the centre’s estimates on centre/course/grade level, taking into	1. The moderation was primarily based on a quantitative

<p>account historical patterns of attainment for each centre</p>	<p>approach⁸. There was no engagement in a qualitative discussion with centres and/or local authorities in order to understand and cases where there was variance from historical attainment. We note that centres and LAs expected this to occur; the subsequent failure to meet expectations contributed to the later sense of grievance.</p> <ol style="list-style-type: none"> 2. Equity issues that might result from the application of a statistical moderation process could have been also considered more fully at this stage 3. Despite the early warning about potential equality impacts, there was little evidence of systematic data analysis to identify anomalies, drawing on government and local government expertise in statistics⁹. 4. Although the PCR system was in place to address anomalies, SQA do not appear to have fully appreciated the impact that the moderated results would have on individual learners, their families, teachers, public opinion, et cetera.
<p>Overall assessment</p>	
<p>After examining this evidence, we believe that more systematic engagement between SQA and different stakeholders in a process of co-construction of the moderation system and a better dialogue between the SQA, Local authorities and centres might have resulted in developing a moderation system that was more equitable to individual candidates. Creating a better understanding about the moderation process could have mitigated the impact that the publication of the results had on young people, their families, teachers and general public. We appreciate that significant pressures caused by time constraints significantly limited possibilities for such engagement – but, in line with stakeholders such as ADES, we do not believe that this was impossible.</p>	

⁸ There was qualitative input from SQA’s subject Principal Assessors, Qualifications Managers and Heads of Service into defining the Starting Point Distributions and reviewing the model outcomes.

⁹ SQA’s position is that ‘to include any considerations of socio-economic status into the model and/or seek to validate with Local Authorities, would have made the approach subjective and introduced ‘bias’ and perceptions of bias into the process’.

The approach to moderation

The moderation of centre estimates was a part of the Alternative Certification Model (ACM) developed by the SQA and is described in its Technical Report (SQA, 2020). We note here that estimates were produced by teachers and lecturers, using both the normal band scale 1-9 and the 'refined' band scale 1-19. Additionally, centres provided a rank order of candidates within each refined band. SQA argued that they requested more granular estimate scale and rank order to support more nuanced decision making and to address two important aspects of teachers estimates: absolute accuracy (where the grade is estimated against national standards) and relative accuracy (a rank order of the candidate among other candidate who achieved the same grade).

As we observed in a previous section of this report, existing literature on the accuracy of teachers' predictions highlights issues of accuracy. This, combined with the fact that many centres had a limited amount of evidence upon which to base their estimation (e.g. limited information about prior attainment and limited access to coursework) suggests that the accuracy of the estimates could have been problematic. Some form of moderation of estimates was therefore necessary.

SQA considered and evaluated several technical options for the moderation of centres' estimates and the awarding model. Full description of the options listed below is a summary of the information provided in the SQA Technical Report (SQA, 2020), where detailed discussions of advantages and disadvantages of each one of these options can be found. The possible approaches are as follows:

1. Directly awarding centre estimates.
2. Linear regression modelling.
3. Awarding using national moderation only.
4. Centre level moderation.
5. Awarding using centre-supplied rank order.

The SQA used the following assurance framework to develop their ACM.

- The application of extant existing policies and procedures whenever possible, the application of the SQA risk management framework and review by heads of services, directors and the Chief examiners.
- Oversight and approval by internal governance groups, including relevant project boards and oversight by the Code of Practice Governing Group and the SQA Board, supported by the Qualifications Committee and Advisory Council.
- Independent review using appropriate sources of technical assurance.

Expertise in educational assessment and statistics was provided by private contractors, AlphaPlus and SAS, who supported SQA in formulating a robust and deliverable approach for moderating estimates. SQA used key members of its Qualifications Committee and Advisory Council to provide professional expertise at key steps in the process. SQA also sought the advice of the Scottish Government's Qualifications Contingency Group, which involves key system stakeholders, at key points in the process.

The moderation approach is outlined below (SQA 2020).

1. A centre's estimates (per grade per course) were assessed against that centre's own historical attainment on the same grade on that course with allowance for variability beyond the previous years' historic attainment;
2. The approach allowed for variability in attainment relative to historical attainment through making wider the tolerable attainment range for attainment at each grade.
3. The approach allowed for a historical variability in attainment at course level, through undertaking assessment at each grade for each course (rather than using total estimated attainment for each grade at the centre compared to historical total attainment for the same grade at the centre)
4. Estimates were only adjusted when a centre's estimated 2020 attainment for a grade were outwith the tolerable ranges, including the allowances for variability on historic attainment.
5. To ensure that the cumulative result of centre moderation was broadly consistent with historical attainment by grade for each course nationally starting point distributions (SPD) were used. SPDs were created, based on: 1) proportional national attainment level for each grade in 2019 (with some adjustments) for Higher and Advanced Higher qualifications; and 2) taking averages of attainment data per course for years 2018 and 2019 for National 5 qualifications.

The ACM has been repeatedly stated (by the government and SQA) to be a mixture of both quantitative and qualitative approaches and said not to rely wholly or even mainly on historical comparisons at the level of whole cohorts. For example, the SQA said:

The data we will be working with includes school and college estimates, rank orders, historical results and estimates for all National Courses as well as learners' prior attainment data for many Highers and Advanced Highers. This will allow us to explore the reasons for any apparent changes in the pattern of attainment (compared with previous years) that are reflected in the estimates submitted by schools and colleges. Such an approach needs to incorporate multiple checks and decision rules to identify where adjustment may be necessary. (Latest SQA statement to schools and colleges – Wednesday 3 June 2020)

On 6 June it was stated that:

After completion of the initial check, SQA will ... carry out a centre level moderation exercise Based on the above centre-level moderation exercise, SQA will explore if it is feasible, within the time available, to engage with schools, colleges and/or local authorities to discuss any reasons for the change in estimated attainment'. (Qualifications Committee 6 May 2020 Alternative Certification Model for Diet 2020).

In fact, the developing of SPDs was the only part of the moderation process where the SQA Technical Report mentioned a qualitative phase. Thus the report says:

This initial SPD was supplemented by a qualitative review by key SQA subject expert staff and appointees including Qualifications Development heads of service, qualifications managers and principal assessors. In some cases, this review resulted in adjustment to the initial quantitatively-derived SPD based on insight provided or trends highlighted by these subject experts... Accordingly, the subject experts might advise that a slightly different national distribution would be expected for 2020, relative to previous years. (SQA, 2020, p.29)

Subsequently, after analysing centre estimates, a decision was made not to enter dialogue with centres and use a purely quantitative approach to the moderation.

Statements from SQA in panel interviews suggest that the decision to move entirely to a quantitative approach was taken once the scale of what was seen as 'over-estimation' became apparent in early June – given the short timescales and the sheer volume of work/limited capacity, qualitative checking as part of the moderation as abandoned at this point. As one SQA official told us, 'The sledge hammer was because of the estimates and how different they were from historic distributions.' (SQA panel). The main reason for using this approach was that there were not enough data in Scotland about previous attainment at an individual level. Thus, a pragmatic approach was taken with some tolerances built in to account for year on year cohort variation; SQA maintains that this was the best approach in the circumstances and that any candidate-level anomalies would be resolved through the PCR process.

Some questions of equity were taken into the consideration at the outset of the ACM. Thus, SQA acknowledged that not all young people have conditions at home to continue to work on their coursework. These assertions are difficult to square with the fact that the subsequent key process – the national moderation phase – was entirely quantitative, based on a mathematical optimisation procedure, Mixed Integer Linear Programme (MILP; see below), using prior data of cohorts on subject/level for past four years in the same centres (except in the cases of first presentation by a centre or very small cohorts of 5 or fewer students). We would argue that equality and equity issues should have been also considered more fully at this stage, and reflected in the methodology, not least because the research literature questions the accuracy of the prediction of attainment, which varies not just between different types of schools, but also by students' prior attainment, socio-economic background and other characteristic (gender and ethnicity). For example, after controlling for prior attainment and socio-economic background, students from state schools are actually less likely to be over-predicted than those in independent and grammar schools (Wyness, 2016¹⁰). We believe that the government could have run some statistical analysis of the data at the immediate post-submission stage to identify patterns in the data, and as requested by ADES.

¹⁰ Although this is based on A-levels we do not have reasons to think that it operates different in Scotland.

Many respondents have suggested that it would have been possible to undertake qualitative moderation to complement the quantitative approach used, for example dialogue with centres, and this was initially considered by SQA, before being rejected on two stated grounds: 1] the sheer scale of the task would be impossible given limited resources and short time scales; and 2] to attempt to do so would create inequity if not all centres could be involved in dialogue. A decision to moderate centre estimates using a purely quantitative moderation procedure created, according to many respondents, a huge gap. Teachers, head teachers and local authorities we have spoken to, felt very strongly that there was a need to have a system in place for verifying evidence used for producing estimates, at least for those cases where the centre estimates were in a stark contrast with historical attainment trend, prior to moving to a national moderation phase. Although many respondents agreed that this might not have been feasible for the SQA, given the time constraints, to engage in a dialogue with every centre, they felt that the SQA should have engaged in dialogue with local authorities. For example,

In their position paper submitted to this review ADES said:

ADES continued to communicate with SQA over a willingness to support the moderation process. They offered that every local authority would make themselves available to discuss a 'first draft' of grades where patterns at departmental level, school level or authority level were not in line with previous trends. It was accepted that SQA could not be expected to work with individual centres but could have worked with 32 local authorities. Despite a series of conversations, SQA declined this offer giving reasons of potential unfairness. It is our belief [sic] that this could have had a major bearing on the outcomes.'

Indeed, we have seen evidence that local authorities were concerned that centre estimates would be subject to arbitrary moderation by the national moderation process. According to one LA, 'The additional step of asking the SQA to contact Directors of Education in LAs to discuss any anomalies would have helped prevent this.' As we have already described in previous sections, some local authorities (although there was a considerable variation in these practices) told us that their centres submitted rationales for variances between the 2020 centre estimates and the centre's historical attainment to SQA. Other local authorities collected such data from the centres and expected to be contacted by SQA.

Based on the stages described above, the following procedure was applied (this is a simplified description of the procedure; see the SQA technical report for a detailed description):

- Historical attainment data were used to calculate an upper and lower tolerance for estimates for each centre, course and grade.
- For each one of years 2016, 2017, 2018 and 2019, centres were ranked by proportion of entries achieved each grade (per course)
- These rankings were split into ventiles (20 bands).
- A representative attainment percentage was derived for each ventile, by taking the four-year mean percentage for each ventile.

- The acceptable tolerance for each school/course/ grade combination was two ventiles higher than its historical best and two ventiles lower than its historical worst performance¹¹.
- Moderation took place if the estimate was outside the tolerance range
- In addition to centre moderation to ensure consistency with that centre's historic attainment, this approach also ensures that the cumulative moderated outcomes across centres for a course are within pre-defined national tolerances using the SPDs.

To implement this moderation procedure the optimisation technique based on mixed integer linear program (or programming) (MILP) was used (SQA, 2020, p.40). MILP is part of a family of Mathematical Programming techniques that optimise (by maximising or minimising) a (linear) objective function subject to a number of constraints. Mixed integer programming adds an additional condition that some of the variables are integers. MILP has many applications such as production planning, scheduling, et cetera. (Williams, 2013). SQA defined the optimisation problem as follows: When adjustment was needed the primary linear objective function was to minimise the number of candidates moved between the grades to meet the centre constraints for each grade and A-C rate (SQA Technical Report, p. 40).

As explained previously, we have not had access to the student datasets and detailed methodology and the detailed algorithm/computer code used by SQA (nor the resources/time to undertake such an analysis in the context of a rapid review). These would be needed to conduct a comprehensive analysis of the working of the ACM, and/or examining in detail the overall suitability of using the MILP approach to moderation, as well as exploring whether some changes in the definition of the optimisation problem, including the formulation of the primary objective function, could have produced better moderation results. The datasets and codes would also be required to conduct modelling and evaluate alternative approaches. Such an analysis would be necessary to address various questions raised by our review, for example relating to evidence of unexplained variance in moderation between different schools (with some centre moderation results for some subjects being lower than they should, based on the centre's historical attainment trends), between subjects in the same schools (e.g. MSA position paper) and (anecdotally) between candidates within the same cohort. We have, however, seen local authority and school level analysis of trends in grade adjustment, suggesting a number of problems highlighted below.

Issues arising from the moderation process

The first issue is the one that has received lots of media attention: the schools in areas with higher level of socio-economic disadvantage have been downgraded more than schools in more advantageous areas. Concerns about the impact of statistical moderation on the outcomes of pupils from disadvantaged schools were voiced repeatedly before the publication of the results on 4th August. For example, a

¹¹ If any centre had only one or two years' attainment history on a course for which they had entries in 2020, then the historic range for that centre on that grade was extended in each direction, to provide a range of five ventile bands. The additional allowance of two ventiles in each direction is then further applied to this extended ventile range. Centres with no history (i.e. presenting entries for a course for the first time) were therefore awarded the original estimates submitted by their centres (SQA, 2020, p.36)

letter sent to the DFM in July by Johann Lamont MSP, detailing comments made by constituents – made the following points:

The SQA is going to change pupils' grades to ensure attainment is in line with "prior attainment" of that centre. This will disproportionately punish schools in more deprived communities whilst simultaneously over rewarding schools in more affluent communities. This is because the pass rate in the former is historically lower than that of the latter. (letter from Johann Lamont MSP, copied to the DFM, 15th July, 2020)

This outcome might have been anticipated. Existing research shows that there is a large variation in the accuracy of the predicted grades between different types of schools and by student socio-economic background (Wyness, 2016). There are two reasons why the schools in areas with higher level of socio-economic disadvantage were downgraded more than schools in more advantageous areas:

1. Schools in socially and economically disadvantaged areas historically have on average lower levels of attainment than schools in advantaged areas. Therefore, standardizing in line with prior attainment of the centre disproportionately affects schools in more deprived areas. As a result, high performers at historically low attaining schools would be disproportionately affected by moderation based on historical record of the school because their grades are out with the aggregate level historical performance.
2. Pupils in poor schools are more likely to be lower attaining. Lower attaining pupils are harder to predict, and more likely to be over-predicted. Hence, moderating grades based on the actual performance of their schools would inevitably result in more downgrading for these pupils: students from the most disadvantaged backgrounds are more likely to experience moderate to severe over-prediction (from 2 to 5 grade points) than those from the most advantaged background (ibid).

Therefore, if acceptable tolerance for each school/course/ grade combination is based on the school's historical performance, then given the tendency of over-predicting grades in these schools the estimates would need to be adjusted (downgraded) more to meet the acceptable tolerances. This approach, which at a centre level managed to produce plausible distributions in line with and often better than a centre's historical patterns, seemed to fare far worse at the level of subjects¹² and worse still at the level of individual pupils. Although, this year, the results of schools in areas of socio-economic deprivation were overall better compared to previous years, emerging evidence suggests that individual level injustices have happened, with 'outliers', such as high performing pupils in these low performing schools, who were arbitrarily downgraded. The evidence of the narrowing of the attainment gap between the students from the least and the most disadvantaged socio-economic backgrounds in 2020 has been praised, yet this feels like over-focusing on a wrong metric, since this aggregate trend hides the fact that high attaining students from lower socio-economic backgrounds and improving schools in disadvantaged areas were downgraded more by the moderation procedure, than their more socially and economically advantaged peers in historically better

¹² Based on the information provided by head teachers and LAs. A full analysis of the datasets would be needed to confirm this.

performing schools. More research is needed to gauge the nature and extent of these patterns.

Some schools presented data for this review (based on the analyses of the adjusted grades in relation to the teachers' estimates and historical trends) that shows that, although the 2020 grade distribution at the level of schools broadly resembled historical grade distributions, there were huge variations between the 2020 results and the historical trends for some subjects, and from the evidence presented by the head teachers, there were many candidates whose grades were moderated down in an apparently arbitrary way. Conversely, mediocre students in high performing schools may be unduly rewarded with higher than their estimated grades. While the latter problem was far less discussed in the media than the former one, we saw that many teachers felt very strongly, not only where their estimates were downgraded, but also when their estimates were upgraded in an unjustified way.

The second problem suggested by emerging evidence was that some centres were extensively moderated and ended up with attainment levels lower than they expected or had achieved in the previous four years¹³. From our conversation with the SQA technical panel, it seems that this inadvertently resulted from trying to prevent a creation of centre constraints that are too rigid and do not allow some degree of variability for centres that might perform better/worse in 2020 than in the past four years. To achieve this greater variability, the tolerance intervals estimated for every centre (for each grade/course/level), first based on the centre's performance over the years 2016-2019, were expanded both upwards and downwards. The rationale for allowing a tolerance both ways, rather than upwards only, was to avoid unnecessary upward adjustment of estimates which were lower than the historical performance. Yet, in some cases, where a centre was found to have 'overestimated' compared to the historical attainment, it was adjusted downwards towards the point lower than their historical attainment (although still within their tolerance interval). We think it plausible that in addition to what was mentioned, there were cases where centres had estimated better grades than in previous years, yet still within the tolerance range, but might have been downgraded anyway, because the national level corrections for the tolerance range were added, which might have been lower than the centres' historical attainment¹⁴. Of course, these are only hypotheses which cannot be tested without having access to the computer code, used for the moderation algorithms, and the data. Yet, it seems that introducing more rigid restrictions on the lower boundary of tolerances, which would not allow the centres to be moderated below their historical averages, would have solved these problems.

¹³ Based on the information provided by head teachers and LAs. A full analysis of the datasets would be needed to confirm this.

¹⁴ Distributions (SPD) to introduce adjustments to the tolerances, based on centre-level constraints. The SPDs were estimated as a proportional national attainment level for each grade on a given course. Although the SQA (2020) Technical Report says that they "sought to take the average of as many recent comparable years of attainment data as was available for the course", it subsequently clarified the SPDs were based on two year (2018 and 2019) averages only for National 5 courses, while the SPDs produced for Higher and Advanced Higher courses were predominantly based on 2019 data (p. 29). The latter means that, if the 2019 national result were particularly low for some courses (which we know has been the case for some Higher courses), that could have affected the acceptable tolerances for these courses and downgraded the centre results more than it should, based on the centre's historical attainment alone.

Another potential source of the problem might be the way in which the optimisation problem was defined. A linear program contains two elements: a cost/optimal function and a set of constraints. The constraints must be met at all costs. The cost function, on the other hand should be minimised provided that no constraint is violated. The solution tells us the optimal value of each unknown, such that the cost function is minimised and each constraint is met. If the priority is to prevent an extreme grade movement this should have been set as a constraint. The cost function, on the other hand, should depend on the difference from the historical attainment patterns, since this is what the approach sought to minimise. Yet, SQA did this the other way around. They set the cost function depending on grades movement and assumed that giving high penalties to particular types of movement (e.g. three or more grades) would prevent these movements; but this did not always seem to be the case. High cost is unlikely, but it is still possible. Only if it was given as a constraint could it never happen.

The third problem lies in small numbers of entries in many courses¹⁵ and the resulting problem of over-moderation of some courses¹⁶, which were big enough to be included in the moderation procedure, but still far too small to obtain reliable statistical estimates. Thus, relatively small numbers of candidates distributed across many centres means it is challenging to make statistically significant decisions across centres and nationally in some low-uptake subjects (ibid). Yet, the SQA states that, for the approach adopted in the moderation process for setting centre constraints, sample sizes are not critical (SQA, 2020). SQA believed that the problem of year on year variability of the outcomes for small centres was solved by setting the tolerance range for each grade/course/centre as the minimum to maximum attainment of the centre on this grade for this course, for years 2016-2019, plus additional tolerances to allow for year-on-year changes in the centre performance. Yet, it seems that the latter still did not solve the problem of year-on year variability. This is to a large extent because teachers' estimates for small uptake courses are less based on the historical patterns and more on teachers' knowledge of the pupils whose grades were estimated this year (after all teachers know that six students who they taught last year might be very different to six students they teach this year).

The fourth problem – downgrades of more than one grade or from pass to fail – has been referred to as the waterfall effect (which was downplayed by SQA in reporting of the national trends). What we have seen in the local authority data analysis looked more like an avalanche effect – the smallest number of entries moved from A to B, then larger numbers from B to C and still larger from C to D. One local authority specifically mentioned that the largest number of moderations were for grade C¹⁷. We posit various reasons for this. First, when adjustment was required, entire bands were moved up or down.

¹⁵ In 2019 for National 5, at least half of class entry sizes were made up of 19 or fewer candidates; for Higher, at least half of class entry sizes were made up of 14 or fewer candidates; and for Advanced Higher, at least half of class entry sizes were made up of four or fewer candidates (SQA 2020).

¹⁶ Based on the moderation vs result data provided by teachers.

¹⁷ Based on the information provided by head teachers and LAs. A full analysis of the datasets would be needed to confirm this.

... where it was necessary for entries in a refined band to be moved into another refined band in another grade, those entries previously in the recipient refined band were displaced, rather than the two groups of entries merging. (SQA, 2020, p.34).

This had a knock on effect on the lower band's entries, which were respectively moved further down (when A grades become B grades, the lowest band of B grades may have to become C grade, etc.). Thus, for entries in a refined band (e.g. band 5 in grade A) to be moved into next refined band (e.g. band 6 in grade B) those entries previously in the band 6 were displaced, rather than the two groups of entries merging (SQA 2020). That might result in too many entries being moved down.

We think that one way to avoid this was to use ranking of students within the grade bands (submitted by centres) and to move a minimum amount of lowest ranked entries from the bottom of a higher band to the top of the next band when required, and merge them with the entries which are already within this band, with moving the lowest ranked entries from this band to the next one only if the total number of estimates within a refined lower band exceeded the centre's historical proportion with tolerances.

A related problem is that students in lower grade bands paid a price for overestimation in higher grade bands. The following example illustrates this. The total number of estimates within grade A exceeded the centre's historical proportion with tolerances, while the total number of estimates within the grade B corresponds with the centre's historical proportion with tolerances. Yet, when one moves entries out the grade A and down to grade B, then as the result of this is to move the lowest band(s) of B grades into C grade bands, and so on. As a result, although the original number of entries achieving grades B and C were within the tolerance interval, the students would be downgraded (including from pass to fail) because their teachers 'overestimated' their higher performing classmates.

The potential inequity here lies in the arbitrary nature of the approach; its inability to deal with cohort by cohort variation and particularly its effects on individuals. The use of an appeals system is a technical solution that fails to appreciate the impact on individuals and subsequently on public opinion. As stated by CYPSC in their position paper to the review:

However as a method it appears to have ignored the fact that each statistical point on the graph is an individual young person whose work, effort and attainment have been moderated based on factors entirely outwith their control and which have no bearing on their individual abilities. It succeeds in creating an overall perception of fairness but fails to deliver actual fairness for individuals. (CYPSC position paper)

Email correspondence between the SQA and the government suggests that this issue and its explosive implications for public opinion appear to have not been fully grasped by SQA, other than through its recourse to appeals, until the EQIA was finalised in July, nor by the government until after the results and EQIA were seen at the end of July. Even at this late stage, the focus seemed to rest on presenting a positive picture (the attainment gap had closed in general terms) rather than seeking a fuller understanding of the nuances in the data.

The DFM has asked that we do lots of digging in the stats to show how young people from deprived backgrounds have not been disadvantaged by the results. (Government email, 6th August)

We concur with SQA’s position that it was not possible, to engage in dialogue at a centre level. We do, however agree with many stakeholders that the following would have been possible:

- Analysis of data to identify anomalies, drawing on government and local government expertise in statistics.
- Dialogue with local authorities to discuss and moderate in a qualitative sense (for example engaging with the rationales for cohort variance collected by local authorities).

After examining this evidence, we believe that – despite the constraints of time and resources – more systematic engagement between SQA and different stakeholders in a process of co-construction might have resulted in developing a moderation system that was more equitable to individual candidates. This could have mitigated the impact that the publication of the results had on young people, their families, teachers and general public. It is a view reflected in the evidence submitted by stakeholders, for example:

A stronger commitment to genuine partnership working may well have headed off the subsequent debacle. It would certainly have eliminated the bulk of individual discrepancies (EIS position paper)

Post Certification Review and Appeals

Perceived strengths	Perceived weaknesses
<ul style="list-style-type: none"> • The original PCR process was technically appropriate with clear guidance, based on a review of individual candidate evidence. • PCR was free-of-charge and thus there were no cost disincentives for centres. • The priority ‘fast-track’ PCR process was designed to address the needs of students whose university offers were dependent on their grades. 	<ul style="list-style-type: none"> • PCR was perceived widely as an appeals process, rather than an integral part of the awarding process. • This was exacerbated by SQA not publishing details of the statistical moderation process and its likely implications. • While technically appropriate, the PCR took insufficient account of equity, especially the impact of the process on individuals. • The revised appeals process following the decision to revert to teacher grades narrowed the grounds for appeal, with subsequent problems for schools and young people. • Appeals can only be initiated by centres, with no right of appeal for young people.

Overall assessment

The likely impact of the PCR process, and its public reception in relation to equity issues, could have been thought through more carefully. Clearer messaging¹⁸ about the role of the appeals system, and discussion prior to results day about the ACM model and its implications would have helped mitigate the subsequent political furore. Use of qualitative moderation after the submission of estimates, to complement the statistical approach, may have greatly reduced the number of cases requiring recourse to appeal. In line with the recently announced incorporation of the UNCRC into Scottish law, consideration needs to be given to whether young people should be able to initiate appeals (as rights holders).

The processes outlined for appeals – Post Certification Review – and associated documentation, were clear and technically appropriate in the view of many respondents. Many teachers found, for instance, that the additional guidance on what constituted evidence to be helpful (e.g. independent schools panel). Nevertheless, the appeals process lies at the heart of the fundamental problem with the ACM, and is subject to a number of caveats raised by different stakeholders.

The view of many respondents, echoed to some extent in our discussions with SQA relates to the manner in which the appeals stage of the ACM was presented. Typically, appeals are a recourse available to small numbers of young people, for example to question a grade on the grounds of extenuating circumstances. In such a scenario, it is entirely correct to present the appeals system as a bolt-on part of the process. In the circumstances of 2020, when estimates might be unreliable, and when a statistical approach to moderation might even amplify this, and/or create inequity at a cohort or individual level, an appeals process serves a very different process. In this case, it is an integral part of the ACM, intended for large scale application to ‘fix’ problems that are a consequence of the system of awarding grades itself. In this scenario, the final appeals stage should, in the view of many respondents, have been more strongly emphasised this year as pre-award part of the awarding process, rather than its usual function as a separate post-award process affecting only small numbers of candidates.

Clear understanding highlighted to the country that the awarding of grades was only a step of the overall process. It should have been communicated that this was not the final step to determining grades and that the appeals process both at authority and school level was the final process. (ADES position paper)

This is an issue of messaging, but one that seems to have had profound consequences due to the expectations created. SQA communications did indeed

¹⁸ SQA guidance for centres (Post-Certification Review – Information for centres), released in June and revised in July stated: ‘The alternative certification model is based on teachers’ and lecturers’ estimates, which have been moderated by schools, colleges and SQA. The process may lead to a candidate or group of candidates being certificated with a grade that’s different from their estimated grade. To be as fair as possible to candidates, we are providing a post-certification review (appeals) service to allow centres to request a review of the grade awarded for a candidate or a group of candidates.’ While this clearly indicates the possible effects of moderation, it conveys a message that PCR is an appeals process rather than an integral stage in the process.

position the ACM as a four stage process – but the high numbers of respondents making the above points indicates clearly that the messaging could have been more effective. Moreover, the view of many respondents (local authorities and teachers) – and one we share – is that expectations could have been different, had there been publication in June of more detail about the national moderation process, as called for by the Scottish Parliament Education and Skills Committee.

This would have allowed an explicit acknowledgement that under the unique circumstances, such a process would not only be needed, but was to an extent unavoidable to deal with inevitable issues of students being penalised unfairly.

The second point relates to the likely number of appeals that would have been necessary had the original PCR system being carried to its conclusion – numbering in the tens of thousands. Head teachers perceived this to be a shifting of the burden of appeals from SQA to schools, with significant workload and capacity issues (head teacher panel). We note here the strong view of many respondents that a qualitative supplementary approach to national moderation may have mitigated this.

We share the view that addressing anomalies at the level of individuals was not possible given the pressures on the system, but agree with ADES and other respondents that the number of appeals could have been reduced greatly had there been more analysis of data trends in June, relating to anomalies and dialogue at local authority level (for example to explain variance at cohort and subject levels¹⁹).

Head teacher and local authorities have reported issues arising from the revised appeals system, introduced once the DFM announced the decision to honour centre estimates, in response to the controversy that erupted following results day. The decision to exclude academic judgment (e.g. where new evidence questions the original estimation) from the revised appeals process has removed recourse to students to pursue appeals where estimates were inaccurate, and placed large pressures on schools²⁰. Many respondents have stated that where schools accept the right to appeal on the grounds of bias/discrimination in the original decision, this places schools at risk (e.g. litigation). This in turn may create conditions where appeals are denied because they are not in the school's interest to pursue them:

In this situation young people are dependent upon the school or college agreeing that they have discriminated against the young person or have made an administrative or procedural error and submitting an appeal' (CYPCS position paper).

We have seen significant evidence that this situation is severely damaging relations between schools and parents. The decision to limit grounds for appeal seems to us to be both unnecessary and counter-productive. First, following the decision to revert to estimated grades appears to place only a small number of students – schools report typically 3-4 cases – at a disadvantage, and yet these small numbers have created a great deal of controversy, out of proportion to the number of cases.

¹⁹ Noting that not all centres sit within local authorities)

²⁰ SQA's position is that meant that any appeals process that did not award based on the original centre estimate was contrary to the Ministerial direction on the 11 August 2020.

Second, SQA has repeatedly emphasised to us that many centre estimates were inaccurate; and yet, the system put in place by SQA denies students an avenue to appeal against inaccurate estimates.

A related issue raised by some stakeholders, especially young people, is the view that the appeals process continues to deny young people the option to personally instigate appeals. Only a school can lodge an appeal. According to CYPCS,

Being denied a direct right of appeal, where they believe they have experienced discrimination, breaches not only the young person’s right to an effective remedy under Article 13 and the prohibition on discrimination in Article 14 of the ECHR, and Article 2 of the UNCRC and in the case of disabled young people Article 23 of the UNCRC. (CYPCS position paper).

We suggest that, following the announcement by the First Minister on 1st September 2020 that the UNCRC will be incorporated as far as possible into Scottish law, the time has come to review the rights and role of young people in the examinations appeals process.

Equalities

Perceived strengths	Perceived weaknesses
<ul style="list-style-type: none"> • The principle of ‘fairness to all learners’ was clearly stated as underpinning the ACM. • EQIA and CRIA documents were produced by SQA. • There was a clear focus on bias in assessment, and well-received training on unconscious bias. 	<ul style="list-style-type: none"> • EQIA and CRIA documents were produced very late in the process, with only limited evidence that equalities issues had been fully considered at the development stage of the ACM. • SQA does not routinely collect equality data about candidates. • SQA’s position that it does not have a sound legal basis for routinely collecting information about protected characteristics appeared to impede analysis of data in relation to equalities issues. • The nuanced impact of the ACM in relation to equalities seems to have been obscured by a debate as to whether the ACM advantaged or disadvantaged cohorts in low SES centres.
Overall assessment	
<p>There need for more systematic and robust systems in future to address equalities issues, particularly in relation to the collection and analysis of data, and in the central role of equalities impact assessment in the design and implementation of awarding systems.</p>	

It is clear that equalities issues were considered at various stages of the process of developing and implementing the ACM. We have, for example, seen evidence of discussion relating to bias in the estimation process as early as March (followed by the well-received unconscious bias training), and (following an offer of support from EHRC on 9th April), ongoing dialogue between SQA and various organisations such as the Scottish Youth Parliament and EHRC regarding equalities issues. A primary focus on equalities work seems to have been in the area of bias in assessment, with less focus on how the moderation process itself might produce inequity. For example, the following extract from a presentation to the SQA Board suggests that a focus on bias may even have prevented analyses related to identifying equalities issues, by anonymising data.

Measures built in to moderation and validation process e.g. all data anonymised for analysis, analysis at aggregate level. (presentation to the Board, 9th July)

Moreover, we found only limited evidence that equalities issues were systematically considered or built into the development of the ACM from the outset, other than the sorts of instances related above and through general commitments to and acknowledgement of equalities issues. Concerns about the absence of an Equalities Impact Assessment (EQIA) were raised as early as May by the Scottish Parliament Education and Skills Committee and Equalities and Human Rights Commission. At this point, the DFM stated it was a matter for the SQA (email correspondence). There is little evidence that this was undertaken comprehensively until July, after results were finalised. SQA (in its technical report published in August) described equality impact assessment as being developed ‘in parallel with’ the development of the ACM, rather than it being an integral part of the process. A meeting note on 11 July indicated that ‘SQA have committed to completing and publishing an EQIA to support the certification model, but have not given an indication of a likely date yet’ (Scottish Government 2020 Awarding Presentation to the Deputy First Minister, 11th July). The EQIA and accompanying Children’s Rights Impact Assessment (CRIA) documents and associated processes for their development attracted considerable criticism from interested stakeholders. .

The draft CRIA was not considered by the SQA Board until 30th July and the published document does not address the full range of rights engaged or properly assess the impact of decisions. This meant that the predictable negative impacts of the alternative certification model were not identified and no mitigations were put in place. In particular, the application of a statistical modelling approach at school level resulted in clear and obvious unfairness and disadvantage for many young people. The CRIA should have identified this. (CYPCS position paper)

From the start of this process the NASUWT also pressed for the SQA to publish the details of any equality impact assessment, particularly in respect of the extent to which equalities issues were taken into effective consideration throughout the design and implementation of the moderation process for 2019/20. It is very difficult to understand how decisions were being taken in the absence of any completed equality assessment and the late arrival of the EIA only served to further undermine teachers’ confidence in the process.

(NASUWT evidence to the Scottish Parliament Education and Skills Committee, 7th August 2020)

The EHRC has also been critical of SQA for shortcomings in its treatment of equalities issues, while acknowledging the constraints on this:

SQA did act upon much of the information we provided. However, their effectiveness in meeting their duties was hampered by a lack of embedded structures and practice, which would have allowed them to fully consider equality in the development of the ACM. They were constrained in what they could do not only because of the very tight timescales they were working to but because:

- There was limited existing knowledge and expertise in meeting the PSED, which meant awareness of equality and an understanding of their statutory equality duties were not built into their decision-making structures;
- They do not routinely collect equality evidence, including equality data about candidates and the views and lived experiences of people with protected characteristics; and
- There was no systematic process to ensure such equality evidence and data was used to inform decision-making. (EHRC position paper)

A lack of access to equalities data is evident in correspondence between SQA and the government in July 2020 – ‘a request to perform analysis to support an Equalities Impact Assessment they are performing on their Alternative Certification Model’ (email from government official to John Swinney, 24 July. SQA requested government assistance to analyse attainment patterns using protected characteristic data. SQA do not have any records of the individual data for pupils apart from grades and estimates (and postcode).

SQA do not hold equalities data and therefore cannot examine the 2020 approach for impact on protected characteristics. (Note attached to internal government email dated 3rd August)

Two alternative approaches to this analysis were not subsequently possible: SQA's view was that they could not take receipt of equalities data from government in the absence of a ‘legal basis on which to hold and process pupil characteristic data’; and the government deemed that it could not undertake the analysis prior to results day as this might be seen as unwarranted interference in the workings of an independent exams regulator.

This means that for the analysis to proceed we would have to take receipt of SQA grade data. We would not otherwise receive the pre-moderation data and there could be some concern about us having access to this given the independent role of the SQA in using this data to award qualifications. However these concerns are somewhat reduced as (i) we would not be in a position to take receipt of the data from SQA until after results day on 4th August and (ii) the relevant documentation would make it clear that the data was shared only for the purposes of this analysis and that it would be deleted

immediately upon its completion. We have also consulted with [Redact s30(c)] who have advised that there is no legal impediment to proceeding with the analysis. (email from government official to John Swinney, 24 July)

Our data (interviews with teachers and parents) suggested that some protected groups were disadvantaged more than others, for example children with learning difficulties, and yet the full extent of this was unknown at the time due to a lack of analysis by SQA and the government. More research is necessary to explore these patterns.

The circumstances outlined above seem to have led to a situation where some of the impacts of the moderation model were not fully anticipated or mitigated. We have, for example, found little evidence in email communications between or public statements by SQA and the government that the equity nuances had been anticipated or publicly acknowledged (even fully understood) prior to the furore that erupted after the publication of results. Emails (for example those sent internally on 4th August) suggest a government priority to defend the position that the system is fair on low SES students, in the face of accusations that low SES centres were more likely to have had awards downgraded (e.g. emails about suggested lines of argument to justify the position). Within this dichotomous argument, some implications were clearly grasped (e.g. general pattern of rising attainment in low SES schools²¹), but the focus on this, combined with a lack of systematic statistical analysis at a fine grained level, seems to have obscured other effects (e.g. reported negative effects on high performing students in low performing schools²²).

Another equity issue lies in variation in the evidence used to underpin estimation by centres. Although estimates were largely based on the evidence submitted prior to the closedown, there is evidence that, in some centres, later evidence was taken into account, which to cite one respondent was ‘incredibly unfair’ (local authority panel). Moreover, the evidence for appeals was considered up to 29 May (teacher panels) – this created an issue of inequity since there was a huge variation in the ability of young people to work from home and submit additional evidence (and there was a variation between schools in the amount of available support, virtual teaching, etc.). According to one Director of Education, there needed to be a clear statement that evidence should not be generated after lockdown – this caused ambiguity and unfairness – but neither SQA nor the government provided such a statement.

Communication and transparency

Perceived strengths	Perceived weaknesses
<ul style="list-style-type: none"> • Extensive approach to communication developed by SQA. • Some guidance was clear and well-received. • Some evidence that SQA is developing approaches to working with young people. 	<ul style="list-style-type: none"> • Unclear and inconsistent approaches to communication. • An apparent reluctance by SQA to share some information, widely seen as a lack of transparency.

²² As stated earlier in the report, more analysis of data is needed to explore these emerging patterns.

	<ul style="list-style-type: none"> • SQA did not take up some offers of partnership working.
Overall assessment	
<p>In the context of the pandemic, SQA should continue to develop its work with young people (as stakeholders and rights holders) and to develop greater partnership working with other stakeholders. There needs to be greater transparency in relation to processes for awarding qualifications.</p>	

While it is clear that SQA invested considerable resources in communicating key messages, and while guidance was in general welcomed as being clear, other aspects of communication were experienced in a less positive fashion.

There is a general perception by teachers that SQA communication throughout the process was not always clear or comprehensive (for example important updates being included in an FAQ). Some respondents (teacher and local authority panels) complained about a tendency to send out important updates on a Friday evening after schools had closed, especially when these generated high numbers of parent queries over the weekend.

Young people experienced SQA and school communications as ambiguous, unclear and inconsistent. Many young people and their families saw shortcomings in communication from schools and local authorities. This included: the decisions of LAs not to reveal estimates to children and parents, which due to lack of other communications added stress and anxieties; and young people and their families did not always understand what estimates mean (there was a conflation between the predicted grades, used for UCAS applications and estimates). All this added to the scale of the uproar after the publication of the results, since predicted grades could be more generous than the estimates. While we understand the decision (made by local authorities) to treat estimates as confidential, we are of the view that better communication with young people and their families from the start, including clearer communication about the implications of a statistical moderation system and the use of the appeals system to mitigate these, may have lessened the strong reaction to the published grades in August. We note here SQA's stated position of withholding some information to avoid causing undue confusion and stress, but emphasise that the majority view of young people and parents in our panels was that they wished for clearer and more comprehensive information on the awarding processes and their implications. For example, young people stated that they would have welcomed communication regarding the SQA timeline/development process; even if the SQA did not have the answers in a shifting landscape they would have appreciated being kept up-to-date with the thought process behind decision-making and ongoing developments.

Many respondents see SQA as lacking in transparency, and resistant to working with stakeholders in a genuinely collaborative manner.

Previous concerns about SQA lack of transparency, and perceived organisational resistance to open communication came to the fore – lack of clear communication on how grades would be determined, with the SQA publishing their methodology on results day in a technical way which was not

in clear language for young people or parents/carers. (Connect position paper)

Some respondents reported a perception of SQA as remote from, and lacking in trust in teachers. This feeling has been reinforced by an apparent reluctance to share the technical details of the moderation model and its effects on estimates, despite multiple calls for this to be done.

Had SQA provided stakeholders with early sight of its proposed methodologies as had been recommended by the Scottish Parliament's Education and Skills Committee, this would have provided an opportunity to consider the extent to which they were fit-for-purpose and to put in place measures to address any unintended consequences. (Learned Societies position paper)

SQA justified this approach through a desire to avoid causing uncertainty:

I wonder if we should have been more overt about the profile of estimates versus historical distributions. It would have been difficult and it would not have been popular, but it would have certainly managed expectations. But it could also have unsettled teachers and young people. (SQA panel interview)

We have some sympathy with SQA's position, which can easily be criticised with the benefit of hindsight; we are aware that the full technical aspects of the methodology were iteratively developed through the analysis of data, and that there were genuine concerns about causing undue anxiety for young people. Nevertheless, we are of the view that it would have been constructive, for the reasons already outlined in this report, to have published relevant information about the methodology and its impact on estimates as soon as the estimates had been submitted by schools. The fact that this was not done has contributed to a widespread view – expressed repeatedly by respondents in our panel interviews – that SQA lacks transparency and does not trust in expertise that resides outside of the organisation. We reiterate the point that effective communication is effective insofar as it is experienced as such by its recipients; the fact that so many stakeholders experienced it otherwise should send a clear message to SQA.

We suggest that, given that COVID-19 has created a situation, presumably continuing into the new academic year, where whole system approaches will be needed for the foreseeable future. This can be achieved through dialogue and co-construction of systems required to award qualifications in the coming year in the face of a continuing pandemic. Stakeholders expressed a view that final decisions regarding qualifications need to be made by SQA, as the body with the formal responsibility for awarding qualifications (e.g. local authority panel). SQA can quite rightly point to its well-developed networks of practitioners, who provide a consultative function for the organisation (although we note that many teachers perceive these to be an inaccessible and closed clique; e.g. SAGT position paper). Nevertheless, testimony presented to the review conveys strong perceptions that SQA is an organisation that is resistant to working with stakeholders.

A meeting was brokered by ADES at beginning of April attended by SQA, EIS, SLS and ADES representatives to discuss methodology for determining grades. Support was offered from experienced practitioners across the system to help determine an appropriate methodology. SQA listened to the offers being put forward but felt they had the expertise and knowledge required within their own organisation. (ADES position paper)

We also note that SQA had developed some dialogue with young people during the summer of 2020, building on earlier initiatives since 2018 to involve young people more in decision making and communication (e.g. SQA 2018), and recognise young people as stakeholders. SQA has acknowledged the need to develop a more systematic approach to working with and engaging young people. These early steps provide good foundations for further embedding engagement with young people in their organisational processes, including over the coming year in the likely eventuality of continued COVID-19 disruption to qualifications.

In general, we see considerable potential for a greater involvement of stakeholders, especially in the context of the unprecedented situation caused by the pandemic. We agree with the view expressed by some respondents, that no one organisation could possibly have developed the best set of responses in such an unusual situation, and that this necessitated greater degrees of participative planning and decision making, which would draw more effectively on the collective expertise and contextual knowledge of professionals and young people.

We will return to these issues in our recommendations.

Impact on young people

An important aspect of this review was to better understand the impact of the cancellation of the exam diet on young people. The perspectives of young people were gathered through online discussion panels and position papers submitted by key stakeholder groups. Young people were recruited through national stakeholder organisations including Children in Scotland, Scottish Youth Parliament, Children & Young People's Commissioner Scotland, Student Partnerships in Quality Scotland (SPARQS) and the 'SQA: Where's Our Say?' social media campaign. The young people were all sixteen and over and diverse in terms of geographic spread, level of qualification and type of centre. It should be noted some invited national stakeholder organisations were unable to participate due to the time constraints of the review.

We report on these experiences and perspectives in the following sections.

Events following cancellation of the exam diet

There was a visceral reaction to the cancellation of the exam diet. Young people described a 'meltdown' situation, with students crying and screaming when the announcement was made. There was uncertainty surrounding what counted as evidence and the amount of evidence required. Students reported they were confused by the method by which grades were to be awarded – then about the uncertainty of coursework. Some students, whose schools had submitted coursework to the SQA for marking, had no access to it for evidence.

Inconsistent approaches to applying the Alternative Certification Model were described at school level. Different approaches were noted between teachers within and across departments. Some students reported that approaches varied between subjects, with traditionally academic subjects such as STEM subjects being more rigorous in their estimates than Arts based subjects. Students felt more confident in subjects where their teachers had a comprehensive record of their coursework (e.g. folders of evidence, tracking). Some students reported that their estimated grades would be based solely on prelims, and others on a mixture of evidence that had been collected. Moreover, some students reported that they had been told their estimated grades or there was an intimation of a grade band, whereas others were told this was not permissible.

Overall, the young people reported that the messages they received about their self-worth are based on their school performance. In short, grades matter in their lives. The fact that young people experienced inconsistent estimation processes regarding their estimated grades matters, when they experience the pressure that grades matter. The ongoing stress emerging from the cancellation of the exam diet cannot be underestimated.

Equity

Many young people felt that extenuating circumstances were not taken into account during estimation. For example, students reported that extended periods of illness around the time of the prelims were not considered. Young people, who had experienced extenuating circumstances during the spring semester, such as bereavement, taking on caring responsibilities (young carers) and being care-experienced young people (whose home circumstances can be precarious due to their temporary nature), may not have generated much evidence for estimated grades, and hence were disadvantaged.

Students reported that the impact of poverty and the lack of funding in certain places for digital technology meant that often young people were working with mobile phones to write essays and access materials. Moreover, access to Wi-Fi is an issue within certain families. Young people will tend not to disclose these issues, because of the stigma surrounding poverty. The young people were aware that some private schools continued online teaching throughout lockdown, with fewer issues around technology. Young people reported being unable to hand in jotters with homework, or take jotters home. This disadvantaged students working on paper.

Wellbeing

The societal impact of the anxiety, confusion and ongoing uncertainty of the pandemic needs to be acknowledged, as young people reported it is a very challenging situation for them.

Parents also reported negative effects on wellbeing – especially widespread anxiety.

For young people with Additional Support Needs (ASN), these pressures have been amplified. Some parents have reported that during the school closures there was a lack of support, and this in turn created additional anxiety and

pressure for children with ASN, and has had a long-term impact on their confidence, mental health, and well-being. (NPFS position paper)

Some parents report that their children now lack confidence in the system, lack motivation, and some relationships with schools and teachers have been detrimentally affected by the estimation process. (NPFS position paper)

Transitions

There was a feeling that SQA had not considered the personal impact of the ACM on young people's lives, for example their school subject choices, university offers, college places, et cetera. Young people, who attained poorer than expected results, changed their university courses based on the results released on 4th August. When the decision to award estimates was made, and as A level results were released, they reported no communications from the Universities about their confirmation of the place, causing further stress and anxiety for them. It was reported that students who went through clearing, following poorer 4th August results, were not able to go back to their original course choices following the reversion to teacher estimates (i.e. it was too late to go back if their grades were upgraded). This has altered their study/career trajectory.

Students expressed concerns about the possibility of inflated university entrance grades for 2021, due to the number of students applying for places. Respondents urge flexibility (e.g. that offers are made based on two sittings because of the detriment they experienced in S5). It has also been reported that entrance grades have been inflated because of the increase in demand for places as a consequence of the number of students achieving high grades (e.g. we were told that Law at Glasgow has increased from 5As to 6As).

Future exam arrangements

There is support amongst our respondents for the following:

- **Direct appeals process.** Young people are frustrated by the limited nature of the of the SQA appeals process for 2020. Young people have expressed that they were unable to challenge the decisions of their presenting centre and that they would like to see a direct appeal process available to individuals in 2021. This would account for the extenuating circumstances mentioned above.
- **Continuous assessment.** Young people would like to see achievement captured throughout the year, rather than the 'two term' dash towards examinations (in particular for Higher).
- **A more consistent, transparent moderation process.** The reports from students regarding the variation in how grades were estimated in schools, the nature of coursework and prelims, and the internal deadline for coursework have led them calling for a clear, consistent and transparent process of moderation. This could address the variation in moderation processes and the potential for teacher bias. It is also more likely to engender trust in the system and avoid erosion of teacher-student relationships in schools.

- **Flexible plans, clearly communicated.** The young people suggested that we need flexible plans, that are clearly communicated beforehand and that these should be in place now for the coming year.

Involvement of young people in meaningful engagement

The young people participating in our review advocate a greater recourse to co-construction of policies and documentation. They see a need to be meaningfully involved in the process of policy development and enactment. This may have mitigated some of the issues which emerged in 2020. Relevant information, to guide young people through the process of awarding, could have been developed with young people and shared through media that they access. Keeping young people informed and connected seems key to building a system based on trust and mutual respect. Clear, consistent, transparent lines of communication are considered to be crucial by young people. The points mentioned above all feed into making this happen. Moreover, the young people were clear that telephone helplines do not suit all children and young people. The young people felt that instant messaging is often a less threatening medium rather than a telephone line²³.

Longer term impact of this experience on young people

Our review has highlighted a number of concerns raised by the young people, regarding the future:

- The ongoing impact of COVID-19 on courses, particularly practical subjects where social distancing and health & safety measures have impacted on course content (e.g. PE students reported that they are unable to play indoor sports);
- Mental health/wellbeing – this is and has been a period of prolonged anxiety, compounded by uncertainty relating to arrangements for 2021;
- Impact on relationships with teachers – students embarking on further study with teachers whose estimates they did not agree with;
- Mistrust in the qualifications system;
- Impact on the 2020 cohort – many young people expressed concern that their grades/achievements are devalued and would be looked upon unfavourably for entry to FE/HE and by future employers;
- Financial hardship – many young people have fallen into poverty as a result of the pandemic (e.g. parental job losses, increase in applications for free school meals and school clothing grants). It has been reported that many young people can no longer afford to go to university.

²³ We note here that SQA offers a range of ways in which candidates can make contact including phone, email, Facebook, Twitter, MySQA Sam, Candidate Enquiry Webform. The consistent perceptions of young people interviewed, that SQA communication is confusing, perhaps indicates the need for greater engagement with young people (building on existing work) to co-construct communication channels and promote their wider use.

Discussion

We commence this section of the report by reiterating the extremely challenging conditions under which the ACM was developed, implemented and subsequently received. This is the majority view of respondents, and such sentiments prefaced most panel discussions. Moreover, while this review has made critical observations about aspects of the process, it is not our intention to apportion blame; instead we see the review as an opportunity to offer constructive criticism which will inform future responses to the ongoing COVID-19 pandemic, and especially to ensure that young people undergo a consistent, rigorous and [above all] equitable approach to the award of qualifications in 2021. Award of qualifications in such circumstances is clearly a whole system issues, and requires whole system responses. We also preface our observations here with the following points:

- Young people, their families and teachers and lecturers deserve as much certainty as can be reasonably given in the face of an uncertain set of circumstances. This entails clear and transparent communication as soon as possible about the arrangements for 2021, and the rapid development of appropriate support and systems for making them happen.
- It has been communicated very clearly to us – by head teachers, teachers and local authorities – that it will not be possible to both prepare young people for examinations and work comprehensively to generate evidence to be used if they cannot run. This is a case of either/or – but not both. The general view that we need to prioritise a focus on the rigorous evidence base.

The development of the ACM required the establishment of quite different approaches to those employed normally – moving into unknown territory. It was clear that centre estimation would be needed as the foundation for awarding, and that some form of national moderation would be needed, given the historical issues of accuracy with predicted grades – exacerbated in this case by a lack of access to the full range of evidence and under-developed systems for local moderation that could not be easily set up in the available timeframes. This combination of factors created the ‘impossible situation’ described by many respondents. Within these constraints, a coherent approach was developed enabled the award of qualifications to proceed – and we note here that in 75% of cases, estimates submitted by centres were not adjusted.

That said, we believe that certain decisions could have been taken differently, and that this may have led to different outcomes, and prevented the subsequent negative reaction that led to this review. These decisions relate to the following:

- A greater recourse to partnership working in the early stages to develop the ACM. It is clear from our evidence that such working was on offer, but that it was not taken up by SQA.
- Greater transparency, as requested repeatedly (e.g. by the Scottish Parliament Education and Skills Committee) around the moderation system and its implications.
- A different presentation of the PCR as an integral part of the awarding process, rather than as a bolt-on appeals process (as is the case usually). It is worth reflecting here on how the use of different terminology might shape

perceptions of this phase of any future ACM. We also note here the potential for inequity in a system that intentionally puts large numbers of candidates through a post-award process with impacts (as noted in this report) on transitions.

- A greater level of embeddedness of equalities impact assessments in the development of awards systems, at the outset.
- Greater levels of cooperation between agencies, including between the SQA and Scottish/local government, for purposes including analysis of data and national moderation.
- More systemic engagement with young people, as stakeholders and as rights holders, to inform the development of systems.

In reviewing the 2020 award of National Qualifications, we have engaged with a very wide range of respondents, offering their perceptions of the process and sharing their experiences. We have also reviewed a wide range of written evidence. This has allowed us to form views on the development and application of the ACM, and has informed the recommendations we make in the final section of the report. We conclude this Discussion section with some observations.

First, we see a lack of appreciation, by key bodies throughout the process, that the issue of perceived fairness to individuals might become a toxic political issue if not handled with sensitivity and forethought. This has been exacerbated by the lack of clear processes for: 1] embedding thinking about equalities into the initial design of the ACM; 2] limited engagement in collaborative decision making and co-construction at the outset in the development of the model; and 3] a lack of targeted analysis of emerging data trends at key points in the process (compounded by a lack of equalities data at SQA and data-sharing agreements to permit closer working between the government and SQA).

One of the core issues emerging from this review is the apparent focus on the primacy of preserving previous years' distributions. A statement from SQA in one of our panel discussions would seem to reinforce this view:

At the end of the day the bigger picture is preserving the value of the certificate (SQA panel interview).

This concern seemed, in the view of many respondents, to override the other two principles (Fairness to all learners and Safe and secure certification), meaning that, once the estimates arrived at SQA at the end of May, insufficient attention was paid to the impact on individuals. For example, one head teacher stated that of the three principles, the focus was more on system integrity, and less on young people – and that this is wrong (head teacher panel).

We are not arguing here against the idea that national moderation necessary; quite the converse, in fact. However, in our view, the main problem with the specific approach to the moderation was that the task of maintaining integrity and credibility of the qualification system was treated as largely technical exercise that aimed to fit the shape of this year's estimates' distribution into the shape of the historical grade distribution. To achieve this, the procedure was developed that moved 'entries' (neglecting the fact that 'entries' weren't just figures but represented real people)

down the grade scale until the optimal distribution was achieved. Therefore, we would like to shift here the attention from 'how suitable the algorithm was for the task' to whether the task was operationalised in a valid way. Does a shape of the distribution that follows the historical patterns deliver fairness to individual learners and ensures that their grades reflect their effort and achievement? We do not think so. In fact, there was no way to achieve this task, because the statistical procedure did not use any information whatsoever about the individual candidates. So the main question here relates to what the moderation algorithm was supposed to do; and to do what it needs to do the algorithm needs adequate input (data). The algorithm does not 'care' that the data are individuals; it would move the data around until the 'optimal' distribution was achieved. But the solution is 'optimal' only in terms of total distribution, not because it reflects any attributes of individual learners; therefore, the solution could be unstable in terms of individuals. This is why there should have been adequate procedures for sense-checking of data at the level of centres and even subjects (e.g. analysis of data to identify outliers and anomalies), and manual adjustments based upon the qualitative information in the system (e.g. local authority rationales for variance). The appeals process provided a technical solution to this, but one limited by the resources needed to undertake massive numbers of appeals; more especially, it failed to account for the very real impacts on those large numbers of young people, including impact on mental health and wellbeing, and negative outcomes in relation to transitions to Higher Education.

We welcome the action by SQA to provide mitigation for the 2021 qualifications diet, and suggest strongly that arrangements are published as quickly as is possible, to obviate concerns in schools, where teaching of courses is already well-developed. This is essential to remove uncertainty and restore teacher and student confidence in the system. We also note that what is necessary this year will not be the same as long term consideration of the future of qualifications in Scotland. Nevertheless, we have some concerns about the draft proposals published in August, and the revised document due for publication at the time of writing.

- The proposals appear to be premised on an assumption that the examinations in 2021 will proceed as planned (and therefore seek to reduce the assessment burden to compensate for missed teaching and learning). This is by no means a given. There seems to be little consideration of the need to create a robust evidence base in the event that exams are not possible, and estimation once more becomes necessary. The removal of coursework components in many subjects will further erode the existing evidence base.
- The review has uncovered concerns that the proposals will lead to a narrowing of courses, with significant implications for education. Related to this, it has been communicated to us that the proposals may impact negatively on attainment, particularly for disadvantaged students who might perform better in coursework. Several respondents have suggested that the issue with the divergence of estimates and historical performance this year may not be due entirely to inaccurate estimation by centres (as SQA have consistently stated), but instead may also be influenced by a combination of recent policies to close the attainment gap and a possibility that teacher estimation actually provides a more accurate assessment of achievement than exams (which are said to disadvantage some learners).

- The SQA proposals have been criticised for offering a piecemeal approach, which differs from subject to subject. While there is some merit in addressing the contextual nuances of different subjects, there is also considerable merit in a set of proposals that offer a more holistic approach across the system. The BOCOSH position paper, submitted to SQA in response to the consultation, and outlining a set of radical proposals – including suspending exams for N5 to allow more space for the arguably more important Higher exams diet – should be seriously considered. We have found widespread support for this sort of action from teachers, head teachers and local authorities, as we took evidence for the review.

In the longer term, and beyond the remit of this review, we wish to offer some observations about the future of qualifications. There is widespread support across all of the stakeholder groups, with whom we engaged during the review, for a fundamental rethink of the long-term approach to awarding qualifications. Many spoke of the ‘opportunity’ presented by the current disruption. The 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 final test – often in highly formulaic ways – seems to be the norm. We do not hold with a prevalent discourse which frames this debate as an either/or-ism – e.g. either exams or coursework. Exams have their place in any qualifications system, as a valid method of assessment, albeit (as is the case with other methods) with particular strengths and weaknesses. We do, however, advocate a mature debate about the future of qualifications that involves enhancing assessment literacy amongst education professionals, as well as challenging stereotypical attitudes amongst the wider population about what constitutes valid assessment. This debate needs to be balanced against the literature that points to the potential unreliability of teacher assessment and variable levels of assessment literacy amongst teachers, particularly in highly performative cultures that can encourage grade inflation (e.g. Priestley & Adey, 2010; Willis et al., 2013; DeLuca et al., 2016). This in turn raises broader questions about the governance of education systems and particularly the place of accountability mechanisms in creating perverse incentives that might distort educational decision making (e.g. see: Cowie et al., 2007; Biesta, 2010; Priestley et al., 2015).

This, in turn, sheds light on the continued viability of a ladder of qualifications approach, characterised by the ‘two term dash’ and a competency-based ‘mastery’ approach to assessment. We would argue, on the basis of the evidence from our review, that the Covid-19 crisis has stimulated some valuable debate in this area, including amongst young people and their parents, and that the time is ripe for meaningful debate about larger scale reform. We note here that many young people want the opportunity to sit exams next year and said that physical measures should be put in place to allow this to happen (i.e. socially distanced exams/perspex screens). We suggest that these discussions are taken up by the OECD review and subsequently through a national conversation.

Recommendations

Summary of recommendations

1. Suspension of the National 5 examinations diet in 2021, with qualifications awarded on the basis of centre estimation based upon validated assessments.
2. The development of a nationally recognised, fully transparent and proportionate system for moderation of centre-based assessment.
3. The development of more extensive approaches to collaborative decision making and co-construction by professional stakeholders of assessment practices related to National Qualifications.
4. A commitment to embedding equalities in all aspects of the development of qualifications systems.
5. The development of more systematic processes for working with and engaging young people, as stakeholders and rights holders in education.
6. The development of a clear communications strategy, co-constructed with stakeholders, to ensure that the extraordinary arrangements for 2021 are as fully as possible understood by all parties.
7. A review of qualification appeals systems, including consideration of the rights and roles of young people, in the context of the incorporation of the UNCRC into Scottish law.
8. The commissioning of independent research into the development and application of the 2020 ACM, involving full access to anonymised attainment data and the statistical algorithms used to moderate grades.
9. The development by SQA and partners of digital materials and systems for producing, assessing and moderating assessment evidence, to ensure that operational processes for gathering candidate evidence for appeals is less reliant on paper-based systems.

Rationale for recommendations

1. Suspension of the National 5 examinations diet in 2021, with qualifications awarded on the basis of centre estimation based upon validated assessments.

This recommendation draws on the oft-repeated statement in our panel discussions that schools lack the capacity, especially in the context of an already disrupted school year, to both prepare students for exams and develop a robust evidence base for centre estimations, should they be needed. It reflects to some extent the BOCSH position paper submitted in response to the August SQA consultation on arrangements for 2021. It is based on the following rationales:

- National 5 is not a leaving qualification for the majority of candidates²⁴, and therefore less high-stakes for most.
- National 5 involves large numbers of candidates – cancelling the examinations diet would enable considerable space to be freed for the arguably more important Higher and Advanced Higher examinations (both of which involve smaller numbers of students, and can be dispersed more readily across school building).

²⁴ For the 2018/19 school leaver cohort, 24.5% of school leavers' highest qualification was at SCQF Level 5 (Nat 5 level) (source Scottish Government).

- The cohorts currently entering school year S3 and S4 have already experienced considerable disruption to teaching time; cancelling the National 5 examinations would allow for additional teaching time in the summer of 2021 for both cohorts.

2. The development of a nationally recognised, fully transparent and proportionate system for moderation of centre-based assessment.

There is a need for a nationally recognised, fully transparent and proportionate system for moderation of estimation (and teacher assessment in general if required), with a number of stages and co-constructed by stakeholders with SQA. This will develop validated assessments for use in National 5 (see recommendation 1), address the current high variance in the capacity and expertise of local authorities, and draw upon the expertise that clearly resides in many LAs. We stress here that **moderation is much more than simply adjusting grades**. It should include: 1] **clear identification and validation of evidence sources**, along with development of protocols for their use (including protocols for using historical data showing both individual prior attainment and cohort variance); 2] **proportional internal verification procedures** (e.g. sampling of decisions and underpinning evidence); 3] **nationally agreed external verification procedures**, based on sampling of decisions and underpinning evidence; and 4] **statistical moderation to identify variance from trends**, accompanied by further qualitative verification (with clear messaging that this will focus on candidates not the system). We note here that such a system needs to be ‘owned’ by teachers, including teacher unions (due to concerns about workload and professional trust in teachers). Its development could be enhanced by fully utilising the experience and expertise of continuous assessment and moderation that resides with the FE college sector. Effective professional education to enhance assessment literacy is essential, and should be developed by SQA working with local authorities and the Regional Improvement Collaboratives as a matter of priority.

3. The development of more extensive approaches to collaborative decision making and co-construction by professional stakeholders of assessment practices related to National Qualifications.

A clear message from the review is that because COVID-19 is an unprecedented threat, normal processes are inadequate to deal with this. We heard strongly expressed sentiments that no single organisation can solve this issue. We have seen evidence that normal protocols and ways of working, including a perceived tendency for SQA to eschew external involvement in its technical processes, have actively hindered actions which might have mitigated the problems experienced this year. For the coming year, there needs to be proactive approach developed which enables a greater degree of collaborative decision making by actors across the system, including where necessary the establishment of systems for data sharing and analysis. Such an approach will ensure higher degrees of ownership of solutions developed in the coming months, and moreover ensure that knowledge – different ways of thinking, local contextual knowledge and expertise, et cetera. – can be pooled more readily.

4. A commitment to embedding equalities in all aspects of the development of qualifications systems.

Considerations of principles of equity, ensuring just decisions for all individual students, should inform all processes for the award of qualifications. This means ensuring that Equalities and Children’s Rights Impact Assessments are

conducted at the outset of developing arrangements for the coming year, and that the principles contained therein should be at the heart of all planning and development. We would strongly endorse the idea of the government and SQA continuing to work closely with CYPCS and EHRC to realise these principles

5. The development of more systematic processes for working with and engaging young people, as stakeholders and rights holders in education.

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 that affect them, in line with Scotland's obligations to the UNCRC. This entails thorough consultation with young people as both a precursor for development and as a sense-checking mechanism during development. We recommend that young people are involved in the co-construction of a more effective communications system, whereby SQA, Local Authorities and schools might better disseminate key messages about qualifications. We also recommend that the final report of this review is accompanied by a young people/family-friendly briefing that clearly communicates the key messages in the review. Young people should be actively involved in the publication of this briefing.

6. The development of a clear communications strategy, co-constructed with stakeholders, to ensure that the extraordinary arrangements for 2021 are as fully as possible understood by all parties.

The effectiveness of communication is ultimately measured by how it is experienced by its recipients. Despite extensive channels of communication developed by SQA, many respondents reported that they did not experience this communication as effective. We therefore recommend that SQA engages in dialogue with practitioners and young people to develop their channels of communication further, to highlight the key priorities for communication in the coming year (i.e. what recipients would like to know) and to identify key points in the year when messaging is needed.

7. A review of qualification appeals systems, including consideration of the rights and roles of young people, in the context of the incorporation of the UNCRC into Scottish law.

Current appeals processes should be reviewed to consider the following: 1] appeals which allow estimation to be revisited where new evidence becomes available; 2] the possibility of direct appeals by young people, in line with the principles of the UNCRC, particularly Article 12. According to CYPCS (position paper), 'Many of the negative impacts relating to the cancellation of the 2020 exam diet are the result of a failure to recognise young people as rights holders and as the key stakeholders of the Scottish education system. Similar failings must be avoided in the future through transparent and pro-active consideration of children's human rights at all stages of SQA and Scottish Government decision-making in future models for assessment and certification of young people's achievements.' We endorse this view.

8. The commissioning of independent research into the development and application of the 2020 ACM, involving full access to anonymised attainment data and the statistical algorithms used to moderate grades.

As stated in the Learned Societies position paper, 'While we appreciate that it would be very challenging to undertake such an evaluation within the timeframe of the review, there is a need to generate a detailed understanding of the methodological approach used in order to plan improvements for the future.' We

endorse this view. There needs to be a thorough independent analysis of the application of the ACM, using the estimation/attainment data and the specific codes/algorithms employed by SQA in 2020. This should include the following: 1] modelling of the algorithms used and alternatives; and 2] analysis of patterns of attainment, linked to demographic characteristics. This will enable fuller understanding of the issues relating to the use of statistical approaches to moderation (strengths and limitations), avoiding problems in future cohorts, and especially its impacts on the cohort of 2020. Analysis should include the effects of the 2020 awards on student transitions and destinations. This process should be achieved through the commissioning of independent fully-funded, accompanied by the publication of the full and transparent technical details of the approach employed, alongside the anonymised dataset, which would also facilitate analysis by independent researchers. We believe that the transparency involved in such an exercise would go a long way to restoring trust in the system.

9. The development by SQA and partners of digital materials and systems for producing, assessing and moderating assessment evidence, to ensure that operational processes for gathering candidate evidence for appeals is less reliant on paper-based systems.

Young people and other stakeholder groups (e.g. College Scotland, University of Glasgow Educational Assessment Network) have called for digitisation of coursework and other centre-based assessment materials, which would mitigate the problems caused by hard copies being inaccessible for estimation in 2020. This development would need to be accompanied by an evaluation of whether and how a digital divide in terms of access to hardware and software by young people might impact on disadvantaged young people.

References

- Anders, J., Dilnot, C., Macmillan, L., & Wyness, G. (2020). *Grade Expectations: How well can we predict future grades based on past performance?* (No. 20-14). Centre for Education Policy and Equalising Opportunities, UCL Institute of Education.
- Biesta, G.J.J. (2010). *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Boulder, Co: Paradigm Publishers.
- Cowie, M., Taylor, D. & Croxford, L. (2007). 'Tough, Intelligent Accountability' in Scottish Secondary Schools and the role of Standard Tables and Charts (STACS): A Critical Appraisal. *Scottish Educational Review*, 39[1], 29-50.
- DeLuca, C., LaPointe-McEwan, D. & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28, 251–272.
- Everett, N. & Papageorgiou, J. (2011). *Investigating the Accuracy of Predicted A Level Grades as part of 2009 UCAS Admission Process*.
<http://dera.ioe.ac.uk/id/eprint/9868> (accessed 23 September 2016).
- Murphy, R. & Wyness, G. (2020). Minority Report: the impact of predicted grades on university admissions of disadvantaged groups. *Education Economics*, 1-18.
- Paterson, L. (2018). *Scottish Education Policy: Why Surveys Matter*. CES Briefing, Number 66. Edinburgh: University of Edinburgh.
- Priestley, M. & Adey, C. (2010). *Supporting teacher assessment in National 4/5 Qualifications: a report for the Scottish Qualifications Authority*. Glasgow: SQA (unpublished).
- Priestley, M., Biesta, G.J.J. & Robinson, S. (2015). *Teacher Agency: An Ecological Approach*. London: Bloomsbury Academic.
- SQA (2018). 'Young people are experts of their own experience'. SQA Futures Blog.
<https://blogs.sqa.org.uk/sqafutures/2018/10/22/young-people-are-experts-of-their-own-experience/> (accessed 23 September 2016).
- SQA (2020). *Technical Report: National Qualifications 2020 Awarding — Methodology Report*. Glasgow: Scottish Qualifications Authority.
- UCU (2015). *Undergraduate application and admission survey*. University and College Union. https://www.ucu.org.uk/media/7330/UCU-undergraduate-application-and-admission-survey-Jun-15/pdf/ucu_undergraduateapplicationandadmissionsurvey_jun15rev1.pdf (accessed 23 September 2016).
- Williams, H. P. (2013). *Model building in mathematical programming*. Oxford: John Wiley & Sons.
- Willis, J., Adie, L. & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40, 241–256.
- Wilson, J. (2015). *Students should apply to university once they've received their grades, not before*, The Guardian Higher Education Network, blog post.
<https://www.theguardian.com/higher-education-network/2015/jul/23/students-should-apply-to-university-once-theyve-received-their-grades-not-before> (accessed 23 September 2016).
- Wyness, G. (2016). *Predicted grades: accuracy and impact. A report for University and College Union*. London: University College London.

Young Scot Observatory/SQA (2018). Insight#SQAfutures.
https://www.sqa.org.uk/sqa/files_ccc/221018_SQAAssessmentFutures_YoungScotReport.pdf

Appendices

Appendix A: Position papers submitted to the review

Organisation	Paper submitted	Publicly available?
ADES	Position paper	No
COSLA	Position paper	No
BOCSH	Position paper	Published online
SLS	Position paper	Published online
NASUWT	Submission to the Education and Skills Committee	No
EIS	Position paper	No
SSTA	Position paper	Published online
Connect	Position paper	No
National Parent Forum Scotland	Position paper	No
Colleges Scotland	Position paper	No
University of Glasgow, School of Education	Position paper	No
Learned Societies Group on STEM Education	Position paper	Published online
CYPCS	Position paper	Published online
Scottish Youth Parliament	Scottish Youth Parliament SQA Equality Impact Assessment and Alternative Certification Model Project Report- June 2020	Published online
Modern Studies Association (MSA)	Position paper	No
Scottish Association of Teachers of English (SATE)	Position paper	No
Scottish Association of Geography Teachers (SAGT)	Position paper	No



Scottish Government
Riaghaltas na h-Alba
gov.scot

© Crown copyright 2020

OGL

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3 or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at www.gov.scot

Any enquiries regarding this publication should be sent to us at

The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

ISBN: 978-1-80004-166-0 (web only)

Published by The Scottish Government, October 2020

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA
PPDAS771406 (10/20)

W W W . g o v . s c o t