

INDEX OF ENCLOSURES

No.	Description	Date	Explanation of redactions	Clause
1	Information document – SQA results	21 August 2020	Information not in scope	-
2	Draft Rapid Review of National Qualifications Experience 2020 report	25 September 2020	Information not in scope	-
3	SQA factual corrections for Priestley report	30 September 2020	Information not in scope	-
4	Email: Results 2020 - Submission - EqiA and CRWIA	02 December 2020	Personal information	38(1)(b)

Enclosure 1 – Information document – SQA results

SQA results and pupil background

[Redacted information not in scope]

4.2 Impact on pupils from the most deprived areas overall

[Redacted information not in scope]

The gap between pass rates of pupils from the most deprived and least deprived areas is narrower in 2020 (based on the Alternative Certification Model) than the gap in 2019 or indeed the average gap over the last 4 years. This is true for National 5, Higher, and Advanced Higher.

Teachers want their young people to achieve the best they possibly can do, and being aware of the challenges and stresses that young people have faced may have meant that teachers who are aspirational for their young people may have ‘given the benefit of the doubt’ to those young people who are close to grade boundaries.

We can examine teacher estimates compared to attainment (2019 exam diet) to check for any patterns in estimation accuracy by deprivation group.

4.2.1 Teacher estimates compared to attainment – 2019 exam diet

In 2019, both teacher estimates and achieved grades show a gap in performance between pupils from the most and least deprived areas at National 5 and Higher. In general, those estimates were also much closer to the final result attained than is the case in 2020, and the proportion of over- and under-estimated grades is more balanced. It is worth bearing in mind that teacher estimates were made under very different circumstances in 2019 than in 2020.

Across all levels, the gap between pass rates from the most deprived and least deprived areas is wider in achieved grades than for teacher estimates. This wider gap in attainment, in comparison to estimates, appears to be a result of teachers estimating higher grades for pupils from the most deprived areas than they actually go on to achieve, and predicting lower levels of achievement for those from least deprived areas than they subsequently achieve.

See also Charts 4, 5 and 6, in **Annex E**, which show estimation accuracy in 2019 by deprivation group. They show overestimation is greatest for pupils from the most deprived backgrounds and lowest for pupils from the least deprived backgrounds.

[Redacted information not in scope]

Enclosure 2 – Draft Rapid Review of National Qualifications Experience 2020 report

[Redacted information not in scope]

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 needed.

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 committed to developing assessment, which captures life skills and contribution to the wider school community. In this vein, young people would have liked a more holistic approach to the ACM.

1.1.1 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 earlier (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 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 process. (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)

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.

1.1.2 Local Authority support

The role of the local authorities appears to be crucial in respect of local moderation of the estimation process. 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 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 submitted rationales for variance to SQA (we do not have evidence that this was looked at by 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.

1.1.3 Coursework

Cancellation of coursework 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).

1.2 National Moderation

Perceived strengths	Perceived weaknesses
1. SQA designed a moderation system to adjust the centre’s estimates on	1. The moderation based on a purely quantitative approach. There was

<p>centre/course/grade level, taking into account historical patterns of attainment for each centre</p>	<p>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.</p> <ol style="list-style-type: none"> 2. Equality and equity issues 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 statistics. 4. Although the appeal system was in place, SQA did not appear to fully appreciate the impact that the moderated results would have on individual learners, their families, teachers, public opinion, et cetera.
Overall assessment	
<p>After examining this evidence we believe that 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.</p>	

1.2.1 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 that the sane 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. Furthermore, use of ranking was justified based on rather out-of-date literature not based on the context of Scottish secondary education system (e.g. Baird, 1997).

[Redacted information not in scope]

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 was abandoned at this point. ‘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 problems with this approach would be resolved through the appeal process.

[Redacted information not in scope] 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 be 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 believe [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.

[Redacted information not in scope]

Enclosure 3 – SQA factual corrections for Priestley report

SQA Fact Check on ‘Rapid Review of National Qualifications Experience 2020’ Final Report, September 2020

Page Reference	Extract from Report	Comment	Suggested change
[Redacted – not in scope]			
Estimation and Local Moderation			
[Redacted – not in scope]			
P12	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).	It would be beneficial if the Review also included SQA’s data on accuracy of estimates.	Additional text: “SQA’s 2019 data comparing estimates versus actual grade achieved using matched candidate data, shows the following level of accuracy 48%of grades estimated at National 5 matched the resulted grades; at Higher 44% and at Advanced Higher 43%”
P12	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 needed.	Further context and rewording here would aid understanding of SQA’s position. Significant patterns of divergence came as no surprise to SQA – and was highlighted in early papers, including to Qualifications Committee. SQA provided additional guidance to centres and historical estimate and results data, the purpose of providing the data was to allow centres to review at dept level whether they got it about right or that	Append text: “This is because SQA provided additional guidance to centres and their historical estimate and results data. The purpose of providing the data was to allow centres to review at department level whether they got estimates about right or that they tended to over- or underestimate. SQA was trying to support centres to improve their estimation accuracy and, as such had aimed through this support to bear down on the amount of

		<p>they had a tendency to over or under estimate.</p> <p>SQA was trying to support centres to improve their estimation accuracy. So: it was an aim not a hope, backed up by taking specific actions to help centres.</p>	<p>moderation required.”</p>
<p>[Redacted – not in scope]</p>			

Enclosure 4 – Email: Results 2020 - Submission - EqiA and CRWIA

From: [Redacted s38(1)(b)]@gov.scot

Sent: 02 December 2020 16:10

To: Deputy First Minister and Cabinet Secretary for Education and Skills
<DFMCSE@gov.scot>

Cc: DG Education, Communities & Justice <DGECJ@gov.scot>; Director of Learning <DirectorofLearning@gov.scot>; Pentland MK (Malcolm) <Malcolm.Pentland@gov.scot>; [Redacted s38(1)(b)]@gov.scot>; [Redacted s38(1)(b)]@gov.scot>; [Redacted s38(1)(b)]@gov.scot>; Wilson M (Mick) <Mick.Wilson@gov.scot>; [Redacted s38(1)(b)]@gov.scot>; [Redacted s38(1)(b)]@gov.scot>; McAllister C (Colin) <Colin.McAllister@gov.scot>; Communications DFM & Education <CommunicationsDFM&Education@gov.scot>

Subject: Results 2020 - Submission - EqiA and CRWIA

PS/Deputy First Minister and Cabinet Secretary for Education and Skills

Please find attached a 2 page submission seeking the Deputy First Minister's approval to publish the Equality Impact Assessment and Child Rights and Wellbeing Impact Assessment covering the policy decision to use teacher and lecturer estimates to award results for national qualifications in 2020.

Thanks [Redacted s38(1)(b)]

[Redacted s38(1)(b)]

[Relevant information from attachments redacted under s25(1):

<https://www.gov.scot/publications/coronavirus-covid-19-equality-impact-assessment-2020-results-using-teacher-lecturer-estimates/>]