

Date:

28 September 2016

Title:

Strategy, Risk and Research Update

Report by:

Michelle Meadows, Executive Director, Strategy, Risk and Research

Responsible Director:

Michelle Meadows, Executive Director, Strategy, Risk and Research

Paper for discussion/information

Open paper, annexes are closed.



Standards

1. Following the organisation re-structure, the new appointments in SRR have significantly increased the capacity and capability of the team to carry out additional analysis prior to the issue of results. This has had a considerable impact on the quality of the analysis and commentary that we can produce in a very short timescale. This will be further improved for 2017 as a result of recent appointments.

Summer 2016 results

2. Results for A level were very stable compared to previous years. This was not unexpected, given that there were no changes to the qualifications this summer. Results for the new AS subjects were also stable for 17 year-olds. Exam boards did not report any particular problems in meeting predictions, which suggested that students were not less motivated taking AS qualifications that no longer contribute to the final A level grade. It also suggests that the sample of students taking the reformed AS was representative of the students who took the legacy AS in previous years.

OFFICIAL – SENSITIVE

3. We published analysis of centre-level variability for both A level and AS. For AS, the main challenge in interpreting the results data was the difficulty in making comparisons between this year and last. We published additional analysis of the AS results for 17-year old students, which provided a more meaningful comparison. The stability analyses suggest that, overall, the schools/colleges who had decided not to enter students for new AS were no different from those who continued to enter students for AS.
<https://www.gov.uk/government/news/a-guide-to-as-and-a-level-results-summer-2016>
4. For new AS we agreed that statistical predictions are the most effective way to carry forward grade standards to new qualifications, so that students in the first cohort of a new qualification are not unfairly disadvantaged. But we also knew that a reduction in students taking these new AS qualifications might make the predictions less reliable. If that turned out to be the case, and exam boards needed to rely more heavily on looking at student work, we wanted to make sure that exam boards would use senior examiner judgement in a consistent way. For this reason, we worked with the exam boards to agree a set of principles for awarding the new AS qualifications, which they all agreed.¹
5. During awarding it emerged that AQA had misinterpreted these principles. We did not consider this misinterpretation to have affected the outcome of any of the awards, but we did make our interpretation of the principles clear to the exam board
<https://www.gov.uk/government/publications/letter-to-aqa-23rd-august-2016>
6. The proportion of students achieving A* in French, German and Spanish increased this summer, while the proportion achieving A or above remained stable. This was welcomed by some stakeholders who felt that there were too few A*s awarded relative to other subjects with similar proportions of students achieving A and above. Our position has always been that we have no evidence that A* was severe but that improving the way in which the papers discriminate at the top end might naturally lead to more A*s being awarded.
7. In GCSE there were some relatively large changes to the proportion of students achieving A*-C. Several factors contributed to this: the increase in post-16 entries for English and maths, cohort changes in GCSE English, and more entries for Ebacc subjects. Results for 16 year-old students were far more stable but still down. We were able to provide additional analysis to show why this was, for example, that

¹ See appendix 3 of the data exchange document
<https://www.gov.uk/government/publications/data-exchange-procedures-for-a-level-gcse-level-1-and-2-certificates/reviewing-gcse-outcome-data-received-from-awarding-organisations-as-part-of-the-data-exchange-procedures-summer-2016>

OFFICIAL – SENSITIVE

the overall cohort for a number of the Ebacc subjects was weaker this year than in 2015, and that the results for 16 year-olds in GCSE and IGCSE English/English language combined were very stable.

<https://www.gov.uk/government/news/a-guide-to-gcse-results-summer-2016>

8. Where we believe that exam boards' grade standards are not aligned within a subject, we have the power to require them to bring their award into line with other boards. We did this for Pearson GCSE mathematics this summer. Although the award was within tolerance, we judged that the grade standards were out of line with those of the other boards, and we were not persuaded by the additional evidence that the exam board presented to us. We therefore required the exam board to move one higher tier grade C boundary up by one mark <https://www.gov.uk/government/publications/ofqual-letter-to-pearson-12-august-2016>. The award remained within tolerance and was closer to prediction. We have since received a Freedom of Information request for more of the background information on which we based our decision.
9. As in recent years, we monitored the grading of CIE's IGCSE English closely. We were satisfied that CIE set their grade boundaries appropriately. Indeed, when the outcomes from GCSE and IGCSE English were aggregated and compared to those of 2016, there was very little change in the outcomes of 16 year olds.

MFL native speaker research

10. Data from the 4,748 questionnaires has now been entered and we will shortly be able to start the analysis. We expect to be able to report later this year, probably in November.

Inter-subject comparability

11. At the May Board meeting, members considered paper 7/16 which provided an update on the evidence gathered about inter-subject comparability in GCSEs and A levels. At that time, we were still working on the implications of an analysis that had been prompted by a letter we had received in April from prominent science organisations. We have continued to work on the implications of this new analysis – which we are currently calling Comparative Progression Analysis (CPA) – and have convened a technical group to explore its significance in greater detail on 23 September. This group comprises experts from Ofqual, exam boards and academia.
12. To support deliberations by this technical group, we have produced an 'overview' report. This explains CPA, relates it to other techniques, illustrates the kind of outcomes which it produces, and raises questions concerning its significance. We have also produced a 'reflections' report,

OFFICIAL – SENSITIVE

which identifies a variety of different ways in which the approach might be rationalised, or theorised. The group will also discuss a ‘critique’ of the approach, which has been produced by colleagues from Cambridge Assessment. Ofqual has undertaken a number of extensive replications of the original analysis, which it will also feed into deliberations.

13. A brief update on the group’s discussions will be provided to the Board when it meets. Outcomes will also be discussed with the Standards Advisory Group, with wider groups of stakeholders, and then again with the Board in November.

Regulatory Strategy

14. Following discussion of the overarching regulatory strategy at the July Board, on the 9th September Board members met for a workshop to continue to discuss strategy in the context of two case studies, one in GQ and one related to technical education qualifications given the government’s skills plan.

15. Board members’ contributions to the development of strategic thinking in both areas was greatly appreciated. The Executive Directors for General and Vocational Qualifications will inform the Board of the outcome of the discussions when they give their updates.

16. At the end of October members will be invited to contribute to the next stage of our strategy development at the Board’s strategy day.

Strategic Risk

17. The Strategic Risk Register is shown in Annex A. After revising the format of the register in July, there has been a refresh of VQ related risks. This has resulted in a number of changes that the Board are invited to note:

- a. A new strategic risk (17) related to delivery of the Functional Skills reform programme has been added;
- b. The risk (03) related to VQ awarding organisation capability and capacity has been significantly updated;
- c. The risk (15) related to Technical Education has been updated and revised down, based on activity since the Board’s last meeting.

18. Further changes to the register include closing the summer 2016 maintenance of standards risk, and updating the risk related to setting maintaining standards in reformed and unreformed qualifications in the future.

OFFICIAL – SENSITIVE

19. The Board will also note that our people risk has increased slightly in light of ongoing challenges in recruiting to specialist areas, but is rated medium-low.
20. Board members have asked for clarity on the use of terminology in our strategic risk register related to awarding organisations. We will use three terms:
- a. Awarding Organisations (AOs) to refer to all regulated entities;
 - b. VQ Awarding Organisations (VQ AOs) to refer to regulated entities delivering VQs, or for those delivering both GQs and VQs – to refer specifically to the VQ delivery aspect of their business;
 - c. Exam Boards to refer to entities providing GCSEs, AS and A levels.

Entity Risk

21. We are beginning to use a new approach to monitoring and analysing entity level risk. Our intention is to create a 'single view of risk' for each regulated entity – bringing together in one place the range of data, intelligence and evidence we hold on them. This is intended to provide easier access to the complete risk picture for each AO, to underpin our compliance activity. This system will interface with our core IT systems and will require development. Hence, the timeline is currently being agreed. In the meantime, we will continue with our current approach to the production of AO profiles.

Risk Analysis

22. At the VOCAG meeting on 9th September we presented a paper setting out our view of risks related to technical education. This report drew on the available evidence, intelligence from stakeholders and expert opinion of colleagues in the organisation to identify enduring, transitional and future risks related to implementation of the Skills Plan. Members of VOCAG provided valuable feedback on the paper, which will now be refined to capture our view of risks.
23. The paper is intended to inform further activity to gather evidence related to the scale of impact each risk presents and steps to assess our assurance of compliance against relevant conditions.

Systemic Risk

24. Board members will have received an invite to participate in the analysis of systemic risks, to inform our six monthly update. As last year, members are invited to participate in a comparative judgement

OFFICIAL – SENSITIVE

exercise to establish the impact and likelihood of the different risks occurring. The risks included have been updated to reflect the current state of the qualifications system.

25. The results of the exercise will be presented to the Board at the October strategy day. In that session, we will invite the Board's consideration specifically of areas where we may wish to undertake further analysis, to inform future action.

Markets

26. We are currently monitoring any potential systemic implications that could arise from a joint project between Crown Commercial Service (CCS) and the Association of Colleges (AoC). The project aims to create a procurement vehicle for colleges to acquire qualifications more efficiently and with greater transparency to improve purchasing decisions. As proposals are confirmed, we will consider these in light of other market changes that will result from implementation of the Skills Plan.

Research

A level Mathematics accreditation

27. We have supported the accreditation of A level mathematics in a similar way to GCSE Science, using a comparative judgement method in order to support judgements about exam demand. We conducted the work to tight schedule to be able to feed the outcomes of the study into the accreditation panel meetings which started at the end of July.
28. The panels found the results of the study very helpful in supporting the accreditation decisions. Evidence from the study is cited in decision reports which have now been sent to exam boards. There will be subsequent rounds of the comparative judgement in order to support the decision making for resubmissions.

Quality of marking

Research into Online Standardisation

29. In July we reported on progress on the online standardisation research, a study aiming to take a close look at the online standardisation experience using a detailed qualitative observational methodology as well as capturing some large scale data through a survey.
30. Since July, online examiner surveys have been constructed based on the observations of live standardisations. To date, we have around

OFFICIAL – SENSITIVE

13,000 responses. The surveys will close shortly. Data analysis is underway. We plan to report in early 2017.

Enquiries about Results and Progress 8 in 2015

31. One implication of the new Progress 8 and Attainment 8 accountability measures is that there should be focus on attainment across all grades compared to the current focus on grade C. In recent years, we have seen that a disproportionate number of Enquiries about Results are from candidates receiving a grade D and just below the grade C boundary. *This sentence has been redacted as its publication would be prejudicial to the effective conduct of public affairs.*
32. In 2015, 300 or so schools opted into the new accountability measures early. We collected the Enquiries about Results data from all boards so that we were able to compare the volume and pattern of the opt-in centres with those of a matched control group (matched on the basis of school type, attainment profile and Enquiries about Results activity in 2014). Unexpectedly, we found that there was no overall difference between the opt-in centres and matched centres in 2015 Enquiries about Results. There was no difference in terms of overall volume, distribution of original grades, or Progress 8 subjects versus non-Progress 8 subjects.
33. Of course, this may reflect the fact that this was a pilot and the pattern of requests for reviews may change this year. Nonetheless, the findings, and the caveats, were shared with boards in the summer to help in their planning and resourcing.
34. On 14 November we will hold an invited symposium to launch the publication of our marking consistency work. The event will be aimed at teachers and teacher associations but invitees will also include academics, exam board representatives and selected media. To provide context to our findings, speakers include those who have researched marking consistency in Higher Education and internationally. There will also be a panel discussion to give a voice to teachers, examiners and teacher educators. If Board members would like to attend, please let us know.

Proficiency specification

35. In the VQ arena, we are developing a programme of research to investigate what might count as good or bad practice in specifying the proficiencies (constructs) that qualifications are designed to assess. Proficiency specification embodies two related activities: constructing the proficiency (identifying what needs to be assessed); and representing the proficiency (communicating what needs to be assessed). During the first phase of this programme, the focus is upon proficiency representation.

OFFICIAL – SENSITIVE

36. The work combines desk-based research and research literature mapping. The desk-based research is exploring similarities and differences between the ways in which different types of regulated qualifications represent their proficiencies. The aim of this part of the project is to provide a foundation and direction for future research, through the development of an analytical framework, plus useful research questions.

Research Conferences and Publications

37. In early September we published the work on the Sawtooth project on the Ofqual website. Further, we will shortly be publishing the work that supported the GCSE Science Accreditation and our work on a new method of inter-board screening for comparability.

38. The Standards Chair recently presented at the International Association for Educational Assessment. Researchers will also be presenting work at Research Ed, the British Educational Research Association this month and the Association for Educational Assessment Europe Conference in early November. These activities are important for dissemination, gaining insight into research activities and methodologies generally and for increasing the credibility and standing of Ofqual.

Paper to be published	YES
Publication date (if relevant)	After the meeting
If it is proposed not to publish the paper or to not publish in full please outline the reasons why with reference to the exemptions available under the Freedom of Information Act (FOIA), please include references to specific paragraphs	<i>See guidance on exemptions below</i>

ANNEXES LIST:-

ANNEX A closed

OFFICIAL – SENSITIVE

ANNEX B closed