

# **Ofqual Investigation into OCR's Summer 2014 GCSE and A Level Marking Issues – Findings**



August 2015

Ofqual/15/5761

# Contents

Background .....	2
Methodology .....	3
OCR's review of the summer 2014 exam series.....	3
Findings from our investigation.....	5
Summary.....	5
Area of investigation 1: the adequacy of OCR's risk identification and contingency planning arrangements ahead of summer 2014 marking .....	7
The completeness of risk identification and contingency planning.....	7
Factors affecting risk identification and contingency planning.....	14
Implementation of 100 per cent e-marking using the scoris web assessor platform.....	20
The scoris web assessor functionality issues affecting marking .....	23
Area of investigation 2: the arrangements OCR had in place to track the progress of marking in the summer and to escalate issues when they arose .....	25
Tracking progress of marking and availability of management information.....	25
Internal risk escalation during marking.....	28
Area of investigation 3: the awareness and understanding of OCR management through the summer of risks to delivering marking on time .....	30
Decision making in response to issues affecting marking.....	30
Factors affecting management's awareness and understanding .....	32
Area of investigation 4: accuracy and completeness of OCR's reporting to Ofqual through the summer incident.....	37
Area of investigation 5: whether OCR provided appropriate assurance to Ofqual through the summer marking incident .....	40

## **Background**

1. Following marking issues being notified to us by OCR over the course of the summer 2014 series that put at risk the delivery of GCSE and A level results, we decided on 16th September<sup>1</sup> to carry out an investigation.
2. The terms of reference for our investigation identified the purpose as being to:
  - establish the facts and understand how the issues in relation to the marking of summer 2014 GCSE and A level qualifications occurred and what the consequences were for awarding and results;
  - satisfy ourselves that OCR's internal review was an open and honest review of summer 2014 marking to identify, recognise and accept the issues identified and to recommend action to mitigate the issues for any future series; and
  - to establish whether there was evidence to suggest we should be taking regulatory action.
3. The terms of reference identified the areas for investigation as:
  - the adequacy of OCR's risk identification and contingency planning arrangements ahead of summer 2014 marking, including an understanding of:
    - the completeness of risk identification and contingency planning;
    - factors affecting risk identification and contingency planning;
    - implementation of 100 per cent e-marking using OCR's web assessor platform (scoris web assessor); and
    - the web assessor functionality issues that affected marking;
  - the arrangements OCR had in place to track the progress of marking in the summer and to escalate issues when they arose, including an understanding of:
    - availability of management information;

---

<sup>1</sup> All dates in this report refer to the calendar year 2014 unless stated otherwise.

- operation of any internal risk escalation process;
- consequences for marking and awarding deadlines; and
- cross-business working;
- the awareness and understanding of OCR management throughout the summer of the risks to delivering marking on time, including an understanding of:
  - decision-making in response to issues affecting marking; and
  - factors affecting OCR management's awareness and understanding;
- the accuracy and completeness of OCR's reporting to Ofqual throughout the summer marking incident, including an understanding of:
  - the accuracy of data and information available to OCR; and
  - the effectiveness of reporting mechanisms; and
- whether OCR provided appropriate assurance to Ofqual throughout the summer marking incident.

## **Methodology**

4. Over approximately a five-week period we carried out a desk-based review of all relevant documentation, including reports from OCR's own review into the summer 2014 series and OCR's restructure transition. We also interviewed a number of OCR staff at all levels and from all relevant business areas.

## **OCR's review of the summer 2014 exam series**

5. OCR carried out its own review into what happened in summer 2014 and identified possible reasons why. On 20th August 2014 OCR produced a preliminary review report. This review was an instinctive view from OCR management as to what happened. OCR confirmed that the preliminary review report would act as a blueprint for the questions that its full review needed to address.
6. On 17th October 2014 we received an update from OCR on the themes that emerged from its full review:

- Issues with web assessor functionality led to a transient shortfall in marking, which recovered at least in the short to medium term. However it did lead to an increase in the number of assessors<sup>2</sup> dropping out.
- Fragmented governance at OCR was a critical weakness. OCR did not have a holistic understanding of the end-to-end process of marking through to the issue of results. This impacted on its understanding of the data exchange and UCAS deadlines.
- As a result of the fragmented governance, there was insufficient contingency planning and linking of risks.
- Web assessor functionality problems used up marking contingency planning time without OCR having a full understanding of the potential implications.
- OCR needs a better understanding of the marking and awarding processes and how to make the assessor body more responsive. OCR intends to apply this understanding to its planning for future summer series.
- OCR needs to implement better management information reporting.
- OCR needs better oversight for when things go wrong, including clearer accountabilities and a focus on risk escalation.
- OCR underestimated the impact of its restructure on the assessor community – specifically, the negative impact from the loss of the qualification manager role on:
  - the support provided to assessors;
  - the value of the qualification managers' working relationships with assessors; and
  - the mobilisation of assessors when marking rates deteriorated.
- With the qualification manager role removed, OCR's poor visibility of assessors' availability throughout the full period of marking became apparent, especially as key marking deadlines approached.

---

<sup>2</sup> OCR usually refers to examiners as assessors and we use this terminology throughout this report.

- There was inadequate planning for some assessor shortfalls.
  - OCR was not sighted on the commitments and availability of assessors throughout the whole marking period.
7. OCR provided us with a copy of its final review report on 3rd November 2014. Relevant evidence from this report has been included as supporting evidence for our investigation findings.

## **Findings from our investigation**

### **Summary**

8. Ahead of marking and awarding summer 2014 GCSE and A level papers, OCR had a number of significant and known risks not present in other recent summer series.
9. First, like other exam boards, in 2014 OCR offered no January series and re-sit opportunities for GCSE and A level. This linearisation meant that OCR had approximately 900,000 more scripts to mark in summer 2014 than in previous series.
10. Second, OCR had introduced an additional level of monitoring for summer 2014 and, as such, had anticipated that this would result in more assessors being stopped from marking than in previous summer series.<sup>3</sup>
11. Third, OCR introduced a more robust standardisation process intended to remove inconsistencies across the different stages of the standardisation process. OCR's final review report identified that its estimation that this new process would extend the timescales by two to three days was an underestimation and that in most cases it took up to a week to clear assessors through standardisation (p 42, AS09).
12. Fourth, in summer 2014 OCR moved to 100 per cent e-marking for the first time and implemented a new web-based version of their electronic script marking system (scoris web assessor). There were certain functionality issues with

---

<sup>3</sup> Across all exam boards, in every exam series, assessors are stopped from marking temporarily and/or permanently if the exam board's live monitoring of their marking suggests they are not following the marking standard.

scoris web assessor that contributed to OCR's subsequent significant marking issues later on in the marking window.<sup>4</sup>

13. Fifth, OCR had been involved in an intensive restructure programme over the preceding 12 months. As OCR's preliminary review report and transition delivery review reveal, the time and resources spent implementing the restructure, plus uncertainty among staff, had limited the effect of contingency planning undertaken for events such as the scoris web assessor performance issues.<sup>5</sup> OCR also introduced a new governance structure in January 2014 that emphasised the management of accountabilities, including business planning for each business area. Also, and crucially, the restructure had removed the qualification manager structure (approximately 120 personnel) who had previously been the primary operational interface with assessors during the marking period. The qualification managers were replaced by a new Examinations & Assessor team that had responsibility for monitoring and managing the quality and quantity of marking and determining what interventions and actions were needed. This was in conjunction with the 21 chairs of examiners who now had an enhanced role in recruiting and managing senior assessors. The change to the qualification manager structure took effect from January 2014.
14. The restructure transition was signed off as complete at the end of March/early April 2014. OCR commissioned a post-implementation review to look at the effectiveness of the Transition Delivery Group. We received a copy of the Transition Delivery Review report on 3rd November 2014. Relevant evidence from this review has been included in this report.
15. The executive summary of OCR's final review report stated: "OCR came very close to missing major external deadlines during the summer 2014 series. OCR demonstrated characteristic resilience in dealing with the summer's problems. Those problems should never have arisen."

---

<sup>4</sup> The problems with scoris assessor functionality from early June through to when performance stabilised in mid to late June did have an impact on marking rates as June progressed into July and towards marking completion deadlines. Yet while linked, they are also separate incidents and should not necessarily be conflated when considering the evidence.

<sup>5</sup> The preliminary review report also said: "A principal risk of undertaking the business restructure and re-organisation over the six months prior to the series was that it might result in insufficient planning and preparation for the summer series." This appears to contradict where the preliminary review report accepted that the restructure had limited planning.

16. The critical findings of our investigation, which run through this report are as follows:

- OCR's restructure impacted on the summer 2014 series. In particular, the qualification manager structure was removed without a full understanding and mapping of how qualification managers interfaced with assessors – including the support qualification managers provided to assessors during marking and managing the marking itself.
- OCR had limited understanding of assessors' availability.
- There was fragmented governance and a lack of clarity in key roles and responsibilities in managing marking.
- There was a lack of understanding among key senior managers of the end-to-end process of marking and awarding. This affected their ability to see the link between the scoris performance issues early in June and the potential for its impact on marking shortfalls and hitting marking deadlines.
- There was no cross-business contingency planning and risk identification.
- Senior managers were working with flawed understanding and assumptions.
- OCR's third-party control within a syndicate structure context was a contributing factor to some of the issues identified.

### **Area of investigation 1: the adequacy of OCR's risk identification and contingency planning arrangements ahead of summer 2014 marking**

#### **The completeness of risk identification and contingency planning**

17. Within the separate business areas of Assessment Standards, Examinations & Assessor and Operations there was evidence of some risk identification and mitigation, together with associated contingency planning, having taken place in preparation for the summer 2014 series. However, the standard of the documentation is variable. Specifically:

- The Assessment Management Group developed an action plan for managing assessor shortfalls together with pre- and post-assessment risk and control registers. The registers had two generic risks relating to the additional risks OCR faced that summer. One referred to assessor shortages relating to a number of factors, including new subjects moving to scoris. The other referred to unintended consequences from the 100 per cent e-marking project not being fully understood. Specific risks were not

identified within these generic risks. The risk registers summarised controls but additional details of the controls that related specifically to assessor shortages could be found in the assessor action plan (created by the Assessment Management Group on February 2014) and in further information submitted by OCR.

- The 21 chairs of examiners each completed risk assessments for their own qualifications. The Standardisation and Marker Monitoring risk assessment template identified risks for pre-standardisation checks, training requirements, panel design, at standardisation and monitoring of marking quality followed by an actions column. A sample of these risk assessments showed they were completed to a variable standard – ranging from detailed mitigating actions to hardly any detail at all. Also, the risk assessments appeared to be evolving documents not signed off by chairs until later in August. This raises the question of how and whether they were quality checked. Although these assessments were escalated to the Assessment Management Group, they did not feature as part of the product delivery risk register that covered many other aspects of assessment business processes. The Head of the Examinations & Assessor team was not fully sighted on these risk assessments, even though there are three risks identified within the panel design section that refer to assessors not completing marking on time and anticipating a significant impact from linearisation.
- OCR provided an update to Ofqual on 20th October, which also identified that “weekly contingency planning meetings were held to resolve any outstanding issues for standardisation meetings scheduled in the next two weeks”. These were also called standardisation monitoring meetings and ran from 14th March to 30th June. OCR’s final review report identified that “at the point of planning the meetings it was not clear exactly what issues they would need to address, but that with the transition and with the new processes there were likely to be issues that had not been directly planned for...” (p 38, Finding AS07). The meetings were not intended to replace the support previously given by qualification managers (for example, if principal examiners expressed difficulties in the use of scoris). In an interview between OCR and Ofqual it was confirmed that there were no formal agendas or minutes for these meetings. OCR also identified a number of other planning actions in preparation for marking, including: “a telephone support rota of experienced ex-QMs [qualification managers] established to deal with queries from senior examiners who might encounter difficulties during the standardisation/approval period”; “all senior examiners were provided with face to face or remote training to prepare them for roll-out of quality of marking initiatives”; “the modifying of

all relevant working instructions to reflect roll-out of quality of marking initiatives”; and “a rota of deputy chairs drawn up to be used as support for senior examiners”.

- OCR's Operations area commissioned a number of short business readiness reports on the 100 per cent e-marking programme. These looked at OCR's approach to managing risks during the June 2014 series:
    - The first report *Scoris Web Assessor Performance Readiness* looked at the stability and performance of scoris web assessor. It outlined that performance issues with scoris web assessor had been previously reported by Cambridge Assessment markers in both the November 2013 and January 2014 series, but these had been addressed in the March 2014 release of the platform. It also contained half a page on what to do “if the worst happens” – which provided limited detail. For example, it identified the contingency of reverting back to the scoris classic platform but with no indication of how this would be done and against what criteria/set of principles. We found no evidence of a reversion plan in place ahead of OCR assessors starting to use the new web assessor platform.
    - The second report *June 2014 Series e-Marking Readiness* outlined how, following transition, capacity, process and functionality changes, and internal process ownership (specifically the removal of the qualification manager structure) were being managed.
18. Ahead of the summer 2014 series, there was no single, cross-business contingency plan and/or risk register that brought together and, where necessary, linked all cross-business risks and mitigating actions in relation to marking, awarding and post-assessment. OCR was unable to present any evidence to show that the planning for the summer series was reviewed and/or signed off at the Senior Management team/Responsible Officer level at any point. OCR's preliminary review report also acknowledged that a principal risk from restructuring during the six months prior to the summer series was insufficient planning and preparation for the summer series. A critical weakness affecting contingency planning and linking of risks ahead of the summer series was fragmented governance and that OCR did not have a holistic understanding of the end-to-end process of marking through to issue of results. Web assessor functionality problems ultimately used up any marking contingency planning time without OCR having a full understanding, at the time, of the potential implications.
19. OCR's final review report identified that: “There is less evidence of a common understanding of, and central responsibility for, the end-to-end awarding and

marking processes” (p 4)<sup>6</sup> and “OCR experienced significant changes that had an impact on the planning for the series” (p 5). It also said: “The shortened timescale available for pre-assessment planning hampered effective scheduling and the identification of critical dependencies.” And: “Reflecting the fragmented nature of governance, contingency planning was done in isolation, rather than in a joined way across the whole marking and awarding process. (p 5) ” Further, still referring to planning and management of risks, the final review report identified that: “Without a clear understanding of, or responsibility for, the critical milestones throughout the end-to-end process it wasn’t possible to understand the consequences of missing them. As a consequence there was insufficient contingency planning against OCR’s principal risks... contingency planning was done in isolation rather than in a joined up way across the whole marking and awarding process” (p 11). The final review report identified that the extent of business change required additional planning and concludes that “some of this planning activity was disjointed and a little too reactive – there was no real opportunity to say ‘this is the plan for the June series, we are ready to go’ from a steady position” (p 18, Finding Pre-PS02).

20. OCR’s transition delivery review identified “gaps in clarity and understanding limiting the extent to which Transition was able to control the risks around some of the activities in which it was involved. Not all planning activity was being overseen by Transition, and Transition’s role in all of the activity it engaged with wasn’t fully defined. The fragmented oversight of planning for the June 2014 series contributed to gaps in the end-to-end planning of process delivery which affected the smooth running of the June 2014 Series” (p 9). Specifically in relation to contingencies and testing, the transition delivery review identified that: “The focus of Transition was on the re-alignment of responsibilities for tasks within processes. Whether for reasons of time or failure to plan, insufficient emphasis was given to the testing of new arrangements of the development of contingencies in the event of failure (p 9).” And: “The lack of contingency & scenario planning, and subsequent testing of processes, limited the control that OCR could potentially exercise over the management of risks to the June 2014 series. When issues arise contingencies have to be developed ‘on the fly’, taking more time and resource than if contingencies are planned in advance (p 10).”
21. A more robust standardisation process, designed to address inconsistencies between various stages of the process and different scripts, was introduced alongside the move to 100 per cent e-marking. There is evidence to show that

---

<sup>6</sup> There is more detail about this in the section “Area of investigation 3” in this report.

this was another factor that impacted on marking as a whole, extending standardisation time to about a week in most cases. There is no clear evidence that on its own it was a significant factor in contributing to subsequent marking delays and the exceptional number of assessors who dropped out of marking. However, OCR's final review report identified that the new standardisation process "contributed to examiner drop out" and that "the extent of the drop outs was not factored in to any additional recruitment requirements because it was not known to be a risk" (p 29, Finding Post-E&A 03). Towards the latter stages of marking, when the standardisation of additional assessors became increasingly time critical, OCR developed the capability to hold additional face-to-face or remote standardisation meetings at any point up to 60 hours prior to the last marking deadline.

22. OCR, in common with the other exam boards, experienced a significant increase in the volume of scripts to mark (approximately 900,000 extra scripts according to OCR's preliminary review report) due to the withdrawal of the January series. Recognising this, and based on entry figures, OCR determined that it needed around 15,800 assessors for the summer series. Allowing for those already in OCR's total pool of assessors, this meant OCR aimed to recruit around 3,256 additional assessors (compared to the 787 additional assessors OCR determined it needed for summer 2013). This also anticipated a certain number of assessors who, as in any summer series, would drop out from marking and allowed for an increased number of assessors being stopped from marking due to the additional marking quality measures being introduced for the first time.<sup>7</sup>
23. OCR entered marking with assessor shortfalls on some panels. The *Deployment Shortage Report* dated 28th May shows 17 panel shortages – two graded red (GCSE Additional Science B and GCSE Physics A) with 20–22 per cent shortfall, and the remainder graded amber with shortfalls varying between 11 and 19 per cent (for GCSE and GCE History, GCSE and GCE separate sciences, GCSE English Literature, GCSE Religious Studies and GCE General Studies). OCR's own review identified that insufficient planning on some panel shortages was a factor. A number of key panels went into the series with a shortage of assessors and there was no contingency pool of OCR-trained and standardised assessors. In interviews, OCR was clear that the spread and number of panel shortages was not exceptional and was less than in previous series – so giving confidence that OCR's fallback contingencies (successful in

---

<sup>7</sup> In 2014, 348 assessors were stopped from marking, compared to 208 in 2013.

previous years) would again deliver marking to schedule. OCR also highlighted the national shortage of assessors in some subjects.

24. Our analysis of assessor shortages by unit for the last three summer series shows that in mid-April the number of units with assessor shortages was not significantly different from the same point in time in 2013 and 2012. In 2014, there were 57 units with at least 20 per cent assessor shortfall at this point in time and 47 units with 10–20 per cent assessor shortages. This compares with 51 units with 20 per cent+ and 53 units with 10–20 per cent in 2013, and 60 units with 20 per cent+ and 53 units with 10–20 per cent in 2012. It is acknowledged that in 2013 and 2012 OCR met the awarding deadlines.
25. OCR's final review report identified that: "During transition when the panels were handed over from Cambridge Deployment to Coventry Deployment it became evident at this point that the panels procedures had not been followed in the Cambridge Deployment team and panels had to be reviewed/amended/reconstructed..." (p 16, Finding Pre-E&A01). And: "Panel plans were rolled over from previous series. This was not the ideal scenario as many of the panel plans ultimately were found to contain anomalies such as team leaders managing too many assessors or allocations which were over-optimistic. These panel plans would have benefited from a review (currently underway for 2015 panels) and adjustment" (p 44, Finding Post-AS10).
26. Data from OCR shows that by the end of marking, the number of additional assessors required had risen from 3,256 to 5,048. The majority of these were required as a result of the very high number of assessors dropping out from marking in combination with the need to not just fill the gaps but catch up marking from time lost because of the scoris web assessor performance issues.
27. OCR recruited 3,744 additional assessors from the start of the recruitment campaign (September 2013) to the end of marking for the summer series. At any one point during marking there was a significant shortfall of assessors (1,495 by the end of marking). OCR does not fully record the reasons why assessors drop out, but it remains a possibility that the scoris performance issues experienced by OCR assessors impacted on the high drop-out rate. On 10th June the drop-out figure was 850 (compared with 517 at the same point in 2013), by 30th June it was 1,345 (compared with 958 in 2013), by 14th July it

was 1,608 (compared with 1,196 in 2013) and by 31st July it was 1,684 (compared with 1,285 in 2013), a 24 per cent increase.<sup>8</sup>

28. The contingency (although not formally documented as such in a plan), as articulated by senior managers, was to use tried and tested approaches. These included: allocating more marking to assessors who had successfully completed 50 per cent of their initial allocation; using financial incentives to increase existing assessors' marking rates (on 13th June OCR offered a retainer to secure the services of assessors affected by the scoris performance issues and, in the very latter stages of marking, used significant, additional financial inducements); and recruiting and standardising more assessors to make up for those dropping out.
29. There was no documented contingency plan exploring wider measures to deal with mass or exceptional loss of assessors over and above the tried and tested approaches described above – for example, testing whether there was full visibility of assessors' availability throughout the entire marking period up to marking completion deadlines, or exploring how assessors from other exam boards could be used to support shortages. An emerging theme from OCR's own review is that it was unsighted on the commitments and availability of assessors throughout the whole marking period. Towards the end of July, Ofqual suggested using assessors from other exam boards. However, it appears that this was not a contingency option considered within the industry as a whole.
30. Senior managers acknowledged that the risk of implementing web assessor and moving to 100 per cent e-marking leading to significantly higher numbers of assessors dropping out was not as clear as it should have been. OCR submitted information to Ofqual in October 2014 citing the Examinations & Assessor team's working instructions as the "contingency for panels where there are examiner shortfalls". This document gives instructions on, for example, adding new assessors when available or increasing assessor allocations. However, it lacks the detail necessary to be considered a contingency plan for dealing with mass and/or exceptional loss of assessors as experienced in 2014 – as evidenced by the significant shortfall OCR had during marking that summer. OCR's final review report stated: "OCR needs to review

---

<sup>8</sup> Figures provided by OCR. OCR's final review report references 1,688 assessors dropping out in summer 2014, compared with 1,122 in summer 2013. It notes that it is not known how many drop-outs were caused by scoris performance issues, but concludes that these drop-outs "undermined the contingency provision" (p 17, Finding Pre-E&A02). Note the 24 per cent increase should be seen in the context of OCR having recruited more assessors for summer 2014 marking to take account of linearisation.

the terms of engagement with examiners and assessors in terms of the end date of the task for Principal Examiners, the number of scripts examiners and assessors are allocated and some examiners and assessors behaviours” (p 17, Finding Pre-E&A03).

31. OCR’s final review report has identified that a joined-up and strategic approach is required for the management of the totality of Cambridge Assessment’s assessors – particularly for those shared by OCR and CIE. It recommends that the syndicate and OCR board should receive regular reports on how this approach is working (p 15, Matters for the Syndicate, OCR Board and Cambridge Assessment Corporate Board). In interview, not all senior managers recognised the opportunities that this could present in terms of resilience of assessors. OCR’s final review report stated: “OCR must allocate appropriate resource and priority to the rapid development and implementation of ‘OCRs Examiner and Assessor Strategy’ (p 15, Recommendation 11). It also identified the need for “A formal schedule of real-time governance meetings to support future series” (p 11, Recommendation 4). In addition, OCR stated: “There is a role for Ofqual and other stakeholders within the UK assessment industry to encourage further participation of teachers – with the support of their employers – in the marking process” (p 14).
32. There was also no documented specific contingency plan and/or associated risk register in place or scenario planning carried out should OCR potentially or actually miss any of the key marking deadlines. This included the last deadline for marks to be uploaded to UCAS. Only on 31st July did OCR begin to prepare a strategic response to UCAS, based on a number of scenarios, should OCR not have delivered all marks in time.
33. OCR’s final review report said: “There is no stated process for dealing with late marking that impacts on post-sign off processes, as it has never been an issue before” (p 33, Finding Post-RTS01). This applies to final marking activities immediately ahead of the issue of results, such as special consideration, null response checks, uniform mark scale conversion checks and assessor scaling. In addition, the final review report identified that: “OCR’s awarding schedules based on old business rules did not have enough in-built contingency to account for the impact of changed processes and unforeseen marking issues” (p 36, Finding Post-AS03). Also, there is no documented process for actions required by the Chair, screeners and the Processing team when an award meeting takes place with less than 85 per cent of marks on the system.

### **Factors affecting risk identification and contingency planning**

34. There were a number of factors that affected OCR’s risk identification and planning for that summer series.

### **OCR's restructure over the preceding 12 months**

35. OCR senior managers stated that the decision to restructure was driven by organisational need. OCR used the opportunity, provided by having no January series in 2014, to carry out the restructure. As OCR's preliminary review report reveals (and interviewees, in general, agreed), the time and resources used to implement the restructure, plus uncertainty among staff, limited the amount of contingency planning that could be done for events such as the scoris web assessor performance issue.
36. OCR also introduced a new governance structure in January 2014 that emphasised the management of accountabilities, including business planning, at the level of each business area, but had not provided for that planning to be effectively brought together. OCR's final review report stated: "The restructuring process and externally driven timescales put pressure on OCR's ability to embed its change programme and meant that planning time for the delivery of the 2014 series was severely limited" (p 11). While this may have been a factor for OCR – although OCR knew well ahead what the external timescales were and in 2014 they were reduced as there was no January series – arguably another more important factor was OCR not fully considering the restructure in terms of the impact on managing marking and assessors.
37. OCR's transition delivery review identified the following points:
  - "The timescales involved limited the effectiveness of the Transition given that not only was OCR trying to accommodate internal change but also external change in terms of linearization and the loss of the January series. The timing of the restructure and the resulting legal consultation left very little time to undertake fully effective planning of new processes and or even the testing of existing processes with new responsibilities. Processes will now need to be reviewed, and where appropriate, resource re-assessed to ensure that requirements can be efficiently and effectively delivered" (p 4).
  - "The clarity of the role of the specific Transition work stream was also a limiting factor in its effectiveness. In some respects its role was clear, and agreement was reached on some specific processes. However, the requirements of the Transition work stream in other areas of activity, notably filling the gaps left by the removal of the QM role, and the handover between Transition and AMG [Assessment Management Group] were less clearly defined / adhered to. Assumptions were made about who was overseeing planning, and these assumptions went unchallenged, which led to gaps in the end-to-end planning for the June 2014 series. In all cases decisions were made based on the best understanding at the

time, some of which subsequently proved to be flawed.” And: “Due to the timing of the re-organisation and subsequent consultation there was insufficient time to re-engineer processes, and insufficient emphasis given to testing the existing processes or to pre-planning contingencies in the event that a process did not perform as expected. This did not mean a complete lack of contingencies as many have been tried and tested in parts over the years, but as they are often not documented they rely on individual knowledge and many of those individuals had either left OCR or now had a different focus” (p 4).

- “In effect OCR had to plan to deliver mainly legacy processes with far less resource than it had had in June 2013. As a result some staff were incredibly stretched and asked to fulfil roles for which they were not adequately prepared, while some processes were fragmented due to the need to find resource rather than develop appropriately connected end-to-end processes” (p 6).
  - “In terms of process detail, verbal assurance was provided on 29 May that working instructions had been reviewed as fit for delivery of the June 2014 Series. No more definite criteria were given. A brief audit of working instructions conducted on 3 June found some inconsistencies in the assurances given to the Transition Delivery Group, including the Missing & Incomplete Results Working Procedures labelled ‘*Draft*’ with no assurance they had been approved, and no assurance had been received or sought from Assessment Standards outside of that given by Examiners & Assessors. This was not followed up by a more structured approach to re-assurance, with clear criteria, but a second request for verbal assurance” (p 8).
38. Crucially, the restructure had removed the qualification manager structure that had been the primary operational interface with assessors during the marking period. All senior managers agreed that the insight that qualification managers had in the micro-managing of marking on a unit-by-unit basis and the understanding of individual assessor relationships and communication was not formalised for the assessor population as a whole within any documented working instructions. When the qualification manager structure was dismantled the probability is that much of this knowledge was lost. OCR’s own review has identified this as a key factor: that the impact of restructure on the assessor community was underestimated – specifically, the close support of assessors the qualification managers took on, the value of their working relationships with assessors, and how the loss of the qualification manager role made mobilisation of assessors more difficult when marking rates deteriorated. During the investigation OCR expressed the view that the role of the qualification manager was fundamentally untenable going forward. Furthermore, it said that the

marking issues raised early in 2014 around history and English supported this view.

39. OCR's final review report stated: "Weaknesses in legacy arrangements have hitherto been compensated for by the ad hoc personal relationships that existed between assessors and QMs" (p 14, Working with the assessor community). And: "It is apparent that the full scope of the QM's role was not captured by the change programme (p 14)." In addition, the final review report concluded: "Fundamentally, the QMs seem to have fulfilled a 'security blanket role' for some senior examiners, and the full impact on confidence of removing that was not anticipated in transition planning (largely because it shouldn't have to have been)..." (p 39, Finding Post-AS07).
40. The final review report also identified that OCR failed to understand the role qualification managers played in managing assessors' "frustrations" regarding system issues: "The lack of scoris knowledge within the business led to a drop in the level of in-series support for senior assessors, particularly at SSU [Standardisation Setup] meetings" (p 48, Finding Post-AS13). It also said: "The accountability of Senior Assessors was perhaps also not made clear enough to them in terms of their responsibility to manage the teams to achieve targets, or the timeframes this would span. The responsibility of managing teams to achieve marking deadlines had in most cases been taken up by the QMs in the past" (p 50, Finding Post-AS14).
41. Another impact of the restructure was in relation to the "QM/QL Task List" (a strand of work to assign responsibility for tasks formerly undertaken by qualification managers). OCR's transition delivery review identified that: "This was achieved, and owners were agreed for all tasks as part of Transition. However, the list originated before Transition began, and many of these tasks were not fully understood by OCR until the new owners began to implement them, long after they had agreed their staff resources in 2013, by which time it was difficult to do anything other than use temporary staff. This had resource implications for some parts of the business" (p 8).
42. OCR's preliminary review report identified that, following the restructure, the new Examinations & Assessor team had been under-resourced from the outset. A data analysis role for the Examinations & Assessor team was only recognised as being necessary as the series progressed. An additional quality support and monitoring manager also had to be added. There is no significant body of evidence to suggest that these resource issues had a major impact on OCR being able to respond in the latter stage of marking by the end of July/early

August, although it did affect the service to assessors.<sup>9</sup> Resources from elsewhere in the organisation, including some ex-qualification managers, were drafted in to support the deployment of assessors and the telephoning of assessors who had either finished marking or had dropped out to see if they were available to take on additional marking.

43. The restructure had also not addressed some single-person dependency vulnerability. A number of critical roles required during the course of the marking and awarding process were staffed by the same individuals. OCR's final review report identified that before the restructure, the qualification managers provided a trained source of knowledge about scoris and/or ModMan:<sup>10</sup> "Post the restructure there was knowledge gap and an over reliance on one individual in the business who was the single point of contact regarding scoris/ModMan issues/knowledge" (p 28, Finding Post-E&A01). It also said: "The issue of poor management information was not adequately recognised in planning and its consequences foreseen" (p 6, Informed decision making). And: "There was insufficient 'expert' business resource to support OCR's data/management information requirements during the live delivery cycle (p 13)."
44. The final review report briefly outlined concerns about how OCR's "historical business model" managed its assessor community: "OCR's change programme recognised the shortcomings of its historic business model and dispensed with QMs" (p 14).
45. The question of whether OCR considered reforming the qualification manager model instead of dispensing with it did not form part of this investigation.

### **OCR's placement within wider syndicates**

OCR is one business unit within the Cambridge Assessment Syndicate. The syndicate also contains CIE and Cambridge English awarding organisations. The syndicate is provided with technical services (including scoris) by a third-party supplier, RM Education. The interface between OCR and RM is through another syndicate function, the Group Infrastructure Services Directorate. OCR's final review report identified that OCR's relationship with the Infrastructure Services Directorate and how the cross-business stream ESM group works has been brought into question. OCR's final review report identified that: "Two cross-stream governance groups (with representation from RM, IS and the 3 Business units – OCR, CIE and Cambridge English) are defined within the Managed Service Agreement" (p 23,

---

<sup>9</sup> There is more detail about this in the section "Area of investigation 3" in this report.

<sup>10</sup> OCR's system for managing on-screen moderation of internal assessment.

Finding PS10). The first of these groups – “The ESM Strategic Operations Board (largely, in place to oversee the strategic aspects of the managed service and the highest point of escalation (p 23)” –had not met for at least 12 months before the summer 2014 series as it was previously agreed it would meet by exception only. Following the summer marking incident, the strategic group had re-started regular meetings. The second of these groups – “The ESM Operating Board, charged with management of the managed service, risk/escalated issue management, commercial oversight” – had met “over recent months (p 24)”. In late May/early June there is evidence that it had responded to scoris performance issues.

46. OCR was of the view that the relationship with the Infrastructure Services Directorate had not worked well. It had been difficult to get assurances that the Infrastructure Services Directorate and therefore RM fully understood the scale of what assessors were reporting to OCR about the scoris web assessor performance issues in the early days, and therefore the implications for marking. OCR’s final review report stated: “The customer-supplier relationship between OCR and Group ISD [Infrastructure Services Directorate] has not worked well, given that these two entities reside within the same Group. OCR has worked on the presumption that ISD has the requisite technical and contract management skills. OCR has therefore taken a ‘light touch’ approach to the governance of ISD. This suggests that OCR did not challenge ISD sufficiently as to the robustness of the reassurances being received from RM... This requires an overhaul of the OCR-ISD relationship to ensure Cambridge Assessment retains the benefits of a centralised IT function, whilst recognising the need for improvement of RM” (p 5). In addition, the final review report identified that: “There is little evidence to suggest OCR’s staff demonstrated sufficient abilities to act as an ‘intelligent customer’. OCR has not been very good at articulating its requirements and the changes it requires, for example in respect of the new standardisation approach” (p 12).
47. OCR does not have a formal service level agreement with the Infrastructure Services Directorate for how it supports OCR in working with RM. RM does not have a service level agreement directly with OCR in relation to ESM – rather it relies on service levels as defined in the contract between the Infrastructure Services Directorate and RM. In interview, the Infrastructure Services Directorate was not clear where ownership/responsibility lay for the fact that scoris web assessor did not function properly for OCR assessors. It could not recall who from its side of the business reviewed the testing evidence from RM. The Infrastructure Services Directorate believed it was OCR’s responsibility to be satisfied that sufficient testing of, and contingency planning for, the first volume use of the scoris web assessor platform was carried out.

### **OCR's relationship with UCAS**

48. OCR's preliminary review report stated that its relationship with UCAS has been "arm's length and limited. Other than a formal letter, there were few formalised protocols and processes and no explicit agreement on specifics such as the size of acceptable amendment files (p 7)". In interviews, OCR acknowledged this. The nature of this relationship impacted on OCR's rationale when it came to providing Ofqual with assurances, as well as its planning for marking contingencies as measured against key marking deadlines. OCR's final review report identified that: "Importantly the significance of failure to deliver the required data to the 7 August UCAS deadline was not fully appreciated or factored into the series planning" (p 5).<sup>11</sup>

### **OCR's understanding of the end-to-end marking and awarding process**

49. OCR did not have a common and effective understanding of, or central responsibility for, the end-to-end marking and awarding process. OCR's final review report (p 4) confirms this, describing OCR thinking "of itself in organisational/structural terms as 'What Operations does' and 'What Assessment Standards does' rather than focusing on the end-to-end process that drives outputs". This was apparent during the interviews with OCR and is also demonstrated by a lack of centralised contingency planning and risk identification documentation.<sup>12</sup>

### **Processes were not always fully understood or mapped**

50. Not all of the sub-business processes that were at the end of the main marking period were fully understood and properly mapped. In interviews, OCR acknowledged not being fully sighted on these processes and the time needed to complete them. OCR's final review report said: "There is no stated process for dealing with late marking that impacts on post-sign off processes, as it has never been an issue before" (p 33, Finding Post-RTS01). This applies to final marking activities immediately ahead of the issue of results such as special consideration, null response checks, uniform mark scale conversion checks and assessor scaling.

### **Implementation of 100 per cent e-marking using the scoris web assessor platform**

51. OCR has been moving towards 100 per cent e-marking for a number of years. Prior to that summer OCR used both e-marking and traditional paper marking of

---

<sup>11</sup> There is more detail about this in the section "Area of investigation 3" in this report.

<sup>12</sup> There is more detail about this in the section "Area of investigation 3" in this report.

scripts. OCR decided to move to 100 per cent e-marking for summer 2014 using scoris web assessor, which had enhanced functionality compared to the previous system. In 2013, a Programme Board for transition to 100 per cent e-marking was set up.

52. A number of short business readiness reports about the 100 per cent e-marking programme summarised OCR's approach to managing risks around the June 2014 series. These were produced for OCR's Senior Management team ahead of summer 2014 marking. One relates to scoris web assessor performance readiness and outlines, in principle, that if all else fails there is the ability to switch the majority of components back to the previous platform (scoris classic). OCR's final review report stated: "Planning and work undertaken in seeking assurances from RM (documented at both the Programme Board and Operating Board) together with OCR's planning outlined in the 'WA Performance' and 'e-Marking' readiness reports demonstrated the due diligence/scenario planning undertaken for this migration. These were, ultimately, however unsuccessful in preventing the issues encountered" (p 21, Finding Post-PS01).
53. However, other evidence appears to contradict this finding. First, interview evidence suggested that the principle of switching back to scoris classic as a contingency was based on an assumption or expectation that any reversion would be more limited than the reversion that occurred (in the end, 175 out of 500 components were moved back to scoris classic). Second, OCR's final review report identified that scoris web assessor was "Not planned for, there was an outline contingency for scoris failure but not for scoris issues on scale they occurred... The impact on standardisation in terms of time lost at meetings and time taken to clear examiners was not fully considered..." (p 40, Finding Post-AS08). Furthermore, OCR's final review report stated that OCR "did not properly consider and identify risks around moving all panels to scoris Web Assessor in terms of load testing and required RM support. We possibly relied on information and assurances from RM and lacked thorough load testing" (p 48, Finding Post-AS13).
54. OCR was aware of the difficulty of moving long-form question marking back to scoris classic (although this only affected a small number of components). OCR's preliminary review report identified that scoris classic did not have the functionality for long-form questions and therefore the contingencies for such papers were limited, unlike the shorter questions, which could be moved to scoris classic.
55. The 100 per cent e-marking Programme Board papers identify that a risk-reduction non-live pilot to prove the use of complex optionality (a feature of web assessor designed to deal with OCR question paper optionality) did not go ahead before the first live use of web assessor in summer 2014. When

interviewed, OCR was unable to provide any further detail/documentation on why this was the case.

56. It appears that the testing carried out by RM as part of its planned release of web assessor for summer 2014 was based on a set of assumptions that did not sufficiently reflect the actual number of concurrent markers experienced in the June 2014 series. Additionally, and more crucially, it did not reflect how OCR assessors would use the system in live marking and as part of standardisation set-up. It is unclear from the interviews and documentation provided by OCR what sign-off there was by OCR or the Infrastructure Services Directorate regarding the effectiveness of the testing. OCR was of the view that it relied heavily on the Infrastructure Services Directorate for this, and the Infrastructure Services Directorate was evasive on this point. Much of the assurance assumed by OCR (and probably the Infrastructure Services Directorate) came from the use of scoris web assessor by CIE markers in volume in 2013 and OCR's more limited use of it in 2013 with no significant functional issues being reported.
57. OCR's final review report identified that the flaws in the programming code in scoris web assessor, which caused the poor performance, were not identified either in live running for other RM customers despite relatively high volumes of scripts, or by the volume testing undertaken by RM. This was based on "RM's assessment of the numbers of concurrent markers suggested by the capacity planning exercise but may not have taken full account of the higher than expected numbers of markers using the mark by question option or the scale of onsite concurrent standardisation setup activity" (p 25, Finding Post PS10). Furthermore, the final review report identified that: "CIE's November series was seen as a test for OCR to determine a wholesale move to scoris web, this test has proved to be insufficient in terms of numbers" (p 30, Post-E&A05).
58. RM did recommend (as indicated in the business readiness documentation) that assessors should use the familiarisation mode in advance of marking. This was to help identify any user/functionality issues. On 14th April 2014, assessors were e-mailed to inform them of the new web assessor release and of the familiarisation mode. The information was also included in the scoris web assessor marker guide on OCR's website. No evidence has been provided to show whether, and if so how, either OCR or the Infrastructure Services Directorate collated and used information on which assessors used the familiarisation mode or any feedback from them about functionality.
59. In December 2013, a number of email communications were sent to assessors to advise and explain about the use of scoris web assessor and the type of user equipment required by the assessors. However, there was no systematic approach to recording, checking or chasing up assessors to ensure they had read and acted on this information and no evidence was provided by OCR of

checking, even on a random sample basis. OCR's final review report identified that OCR "should probably have anticipated the level of assessor unhappiness with changes and the loss of the QM who many principal examiners had established relationships with... the loss of examiner confidence in OCR probably contributed to un-willingness to mark more scripts... further down the line. We could also have communicated changes to senior examiners more effectively" (p 39, Finding Post AS07).

60. There was no detailed plan for reverting to scoris classic in the event of a significant functionality issue with web assessor. A reversion plan was developed in the second week of June by OCR, the Infrastructure Services Directorate and RM. On 9th June, OCR circulated a paper outlining its "Plan B" if issues with scoris web assessor continued and performance was unacceptable. This paper identified that OCR's visibility of RM's contingency plan was poor and that "Given the significance of this system OCR's contingency planning should have considered the possibility of system failure and have had the skeleton of contingency plan situation where the system is not available".
61. In preparation for the introduction of 100 per cent e-marking using scoris web assessor and potential technical issues being experienced by assessors, RM increased its help desk resource (by 60 per cent). However, going into marking the help desk only worked office hours (10.00 to 18.00 weekdays and 09.00 to 17.00 weekends). Later on in June, in response to the scoris functionality issues, the help desk did operate an out-of-office-hours service. When interviewed, both OCR and the Infrastructure Services Directorate expressed surprise that RM's help desk was not accessible outside office hours – although it appears this was in line with the contract in place with RM. It appears that neither OCR nor the Infrastructure Services Directorate had checked this, or if they had, did not challenge the arrangements. At the height of the functionality issues, RM was unable to cope with the volume of queries from assessors. This resulted in assessors further being delayed in starting or continuing with marking. OCR's final review report confirms this ( Finding Post E&A13 p 33, Finding Post PS04 p 22).

### **The scoris web assessor functionality issues affecting marking**

62. OCR's evidence indicates that it first informed us of the scoris performance issues on 2nd June (preliminary review report, Appendix 1). Our records show that the first recorded notification (an email) was on 6th June, followed by an official event notification on 13th June. It was probable that in the first week in June, OCR, the Infrastructure Services Directorate and RM did not fully understand the full nature or scale, and therefore implications, of the emerging functionality issues.

63. A number of interviewees expressed the view that they found it difficult to get satisfactory responses from the Infrastructure Services Directorate and RM. OCR's final review report stated: "It was not clear during the early part of the marking window as to the precise nature of the problem unfolding. There was a considerable amount of 'noise' coming into OCR... This intelligence was not collated in a central point and hence early warning of the impending problem was missed" (p 49, Finding Post AS13). However, by 10th/11th June, OCR, RM and Infrastructure Services Directorate crisis teams had been formed. Over the next few days, options and more detailed planning were developed for a reversion of units from scoris web assessor back to scoris classic, with the first units being reverted on or around 12th June. By, or slightly before, 24th June, 175 (out of 500) components had been reverted to scoris classic.
64. The principal issue with functionality resulted from a number of flaws in the programming code for scoris web assessor's stored procedures (marking transactions). When used too often, as a result of how OCR assessors operated the system (for example, higher than expected numbers of assessors using the mark by question option), the functionality was adversely affected. The system 'froze'. Even when scoris web assessor was operating, the ability of assessors to log in and move around the system without 'hanging' was affected. The ability of the system to refresh quickly was also impacted.
65. OCR also introduced a new, more robust standardisation process for summer 2014. Where there was a large number of assessors being standardised, and all logging in at the same time, this added to scoris web assessor's functionality problems and further slowed down the system.
66. On or around 13th June, OCR offered a [REDACTED] retainer to assessors impacted by scoris performance issues. However, as identified earlier in this report, a significant number of assessors still dropped out from marking. OCR's final review report identified that "it took 2 weeks to stabilise scoris and some of the technical solutions proposed and adopted (at RM's instigation and endorsed/not challenged by IS) led to further outages and performance issues. The impact on the marking window and on OCR assessors was protracted<sup>13</sup> and recovery

---

<sup>13</sup> It is unclear what OCR means by "protracted" in the context of the impact on marking shortfalls, particularly for GCE qualifications, later on in July. In OCR's update to Ofqual on emerging themes from its review (at a meeting on 17th October) OCR stated that the scoris performance issues were a "red herring" when it came to the marking crisis as it had developed in July. In addition, the final review report provided data about batch marking and stated that "scoris issues may have had less impact on standardisation meetings and clearance and subsequent delays in marking than might be suggested by how far these panels then missed batch 2 dates by" (p 41, Post-AS08). Arguably, this implies that the cause of marking shortfalls may lie elsewhere.

(resolution of backlogged support calls and the marking rate) took longer than desirable” (p 21, Finding Post PS02).

## **Area of investigation 2: the arrangements OCR had in place to track the progress of marking in the summer and to escalate issues when they arose**

### **Tracking progress of marking and availability of management information**

67. As evidenced earlier in this report, OCR's restructure removed the qualification manager structure. Qualification managers essentially had been responsible for all aspects of their qualifications including being the main operational focus for monitoring marking progress together with the Assessor Deployment team. The qualification manager structure was replaced with an enhanced role for OCR's 21 chairs of examiners, specifically in recruiting senior assessors, who reported to the Head of Assessment Standards working with a new, more centralised Examinations & Assessor team that reported to the Head of the Examinations & Assessor team. This structure, led by the Examinations & Assessor team, was responsible for monitoring the quality and quantity of marking.
68. For summer 2014, the mechanism to track and monitor marking consisted of a scoris marking status dashboard together with a watch list (and later on a separate section within it called a worry list) showing GCSE and GCE unit marking progress. OCR's preliminary review report identified that the dashboard was only introduced after the scoris performance issue had been identified. From mid-June, the scoris dashboard and watch list was circulated daily internally to all key senior managers, together with the actions by chairs of examiners and the Deployment team resulting from the watch list information. From, on or around, 24th June, a small group of senior managers met daily to monitor the progress of marking data and decide on actions (these were recorded and titled as Assessment Management Group actions).
69. In interview, Assessment Management Group members appeared to be unclear about whose lead responsibility it was to monitor marking, ensure action was taken and, if necessary, escalate concerns. OCR's final review report stated: "What was not clear was the reporting/accountability relationship between the Head of E&A [Examinations & Assessor team] and the Director of Operations whose responsibility it was to ensure that results were published to agreed timelines. In particular, the accountability for monitoring and chasing marking progress was insufficiently clear – was it E&A or Ops? There was perhaps an assumption that it would take care of itself. And even if the accountability was known at senior level, then responsibility for monitoring progress of marking at subject/panel level was not clear... nobody had ownership for specific panels as

in the past” (pp 43–44, Finding Post AS10).<sup>14</sup> Furthermore, the final review report identified that: “The newly re-configured AMG may not have reacted as swiftly and strongly to the emerging late marking issues as the situation later transpired to have required” (p 22, Finding Post PS03).<sup>15</sup>

70. The scoris dashboard provided an overview of marking progress in 2014 compared with the previous year. It also showed the total number of GCSE and GCE marks entered on file in the last day and week. The watch list provided, at unit level, a comprehensive range of information on marking progress including entries, number of scripts marked to date, percentage marking complete, remaining number of scripts to mark, number marked in last day/week, and projected completion date at the current daily/weekly rate. Each unit was Red/Amber/Green (RAG) rated and had commentary from the Chair of Examiners about what actions were being taken. The worry list identified those units where the projected marking completion date was further away (and crucially, beyond the key marking deadlines) or where the marking rate had deteriorated substantially. The judgement as to which unit went on the watch or worry list was determined on the basis of a number of factors, including qualification type, date of award, whether marking had achieved the 50 per cent point, and previous daily and weekly marking rates.
71. Overall, Assessment Management Group members who used this mechanism thought it generally effective as a monitoring and flagging tool and an improvement on the previous system (a report form of marking progress maintained by each qualification manager but with less detail). OCR had used the watch list tracking mechanism before in the previous summer series – used by exception for a small number of panels where there were known assessor shortages and then to monitor any other panels where marking progress looked to be slowing.
72. However, OCR’s final review report identified that “the level of detail in the Watch list, allied to the lack of reference to the significant dates for data submission, would not have highlighted the seriousness of the marking delays without additional detailed analysis” (pp 55–56, Findings XB04). This position is difficult to understand given that many of the projected marking completion dates, based on the previous 24-hour and week marking rates, were clearly

---

<sup>14</sup> This links to a finding in the transition delivery review about sign-off and planning at the high level but no assurance of detailed provision.

<sup>15</sup> Further information about the Assessment Management Group is provided in the next section of this report; Internal risk escalation during marking

shown in the July watch list to be well beyond the issue of results days in August.

73. Assessment Management Group members were not clear about whether the watch list percentage of marks data referred to 100 per cent of marks or the 85 per cent of whole candidate marks needed for awarding (when unit marks are aggregated the percentage of whole candidates on file will be lower). OCR's final review report stated: "Daily AMG meetings identified at risk qualifications in terms of having less than 85% of the marks on the system in time for the award, however as this data was at unit level only it did not account for the number of aggregated candidates at qualification level. In many cases where individual units were above 85%... the overall % of graded candidates (aggregated) was below 85% and the awards could not be screened" (p 36, Finding Post AS05).
74. Prior to summer 2014, there was "poor or non-existent management information" on historic marking rates of individual papers, so no marking rate benchmarks were available. Marking rate information for each component was gathered and monitored in 2013 by individual qualification managers. There was no centralised data set for 2013 that could be compared to 2014. A number of interviewees agreed that this information was not known. However, they felt that this was not as critical as the lack of visibility of assessors' availability – specifically, not having a clear view of when they might stop marking, when they could be available to start marking again, and what marking they did for other boards. There was no electronic system in place with a complete picture of all OCR assessors' marking availability. This presented significant problems for OCR in the latter stages of marking.
75. OCR's final review report identified that: "OCR had a limited understanding of assessor's availability during the series" (p 14). And: "OCR was not aware of the availability of examiners outside the normal marking window as there are no systems or processes to collect and store this data." It also stated that "when the marking shortfall was identified, OCR was not easily able to marshal the assessor community effectively". In addition, the final review report identified that decision-making was affected in the early stages of marking by not having sufficient real-time management information (p 12, this possibly refers to availability of assessors) and that "Visibility of the progress of marking volumes during the series was poor at the start of the series" (p 44, Finding Post AS10).
76. OCR's final review report stated that as a consequence of the summer series issues, "damage to examiners morale" may impact on their willingness to accept subsequent invitations to mark for OCR (p 14, p 32, Finding Post E&A10).

### Internal risk escalation during marking

77. OCR had an incident escalation process in place (*Managing and Reporting Issues* document) where incidents/issues were escalated to OCR's Risk & Compliance team. Once a risk moved from a potential to a real event, it followed OCR's *Managing and Reporting Issues* process. While a risk remained a potential event, escalation followed governance outlined in the *OCR Risk Management Strategy* (January 2014).
78. The speed with which the scoris functionality performance issues came to light, in early June, meant that this escalation process was not used as the main mechanism for escalating to senior managers in the early days. OCR's final review report identified that: "Different people within OCR, especially Process Managers supporting Standardisation Setup (SSU) meetings and Chairs of Examiners, became aware of these issues through anecdotal reports on performance" (p 7).
79. The Assessment Management Group, as the principal management operational group for the Assessment business area, would normally escalate to OCR's Senior Management team. In practice, the speed with which the scoris functionality issues developed and the fact that an OCR crisis team was put together by 10th/11th June consisting of the majority of OCR's relevant senior managers meant that escalation had in reality taken place, albeit not through formal Assessment Management Group channels.
80. There is evidence to show that, in the lead-up to and during the summer 2014 marking period, there was confusion about the role of the Assessment Management Group in coordinating the managing of marking, and, where necessary, in escalating issues.
81. OCR's final review report identified "a lack of clarity" in the role of the Assessment Management Group (p 10). In addition: "There was a lack of clarity on who was accountable for monitoring and chasing marking progress – and who had oversight of the end-to-end marking and awarding process." The final review report concluded: "There were different perspectives on the role of AMG. [Some] felt it was to monitor and manage the series and [others] felt it should provide assurance of sign-off (even though historically the AMG has never had an active role in pre-assessment or post-assessment signoff). AMG saw the prime responsibility for managing processes as sitting with Business Process Owners, with AMG as a forum to monitor, challenge and reach agreement on points at issue. This is what was reflected in the AMG Terms of Reference revised at the start of 2014. The Directors were not part of the process of revising the ToR when, as stakeholders, they should have been" (p 54, XB03). The final review report goes on to identify that: "The newly re-configured AMG may not have reacted as swiftly and strongly to the emerging late marking

issues as the situation later transpired to have required” (p 22, Finding post-PS03). It also stated that: “AMG failed in its management role having become deflected from its main and most important task – that of having oversight of a live series” (p 44, Finding Post-AS10). Nonetheless, it is difficult to understand how this confusion was not identified and addressed well ahead of the series or at least during marking, especially when late marking issues emerged. The reconfigured Assessment Management Group started on or about 24th June and then met daily through the whole of the marking period. It consisted of senior operational managers. They received the watch lists that showed marking shortfalls during July, as did a group of key OCR directors who also contended that they were monitoring marking progress.

82. OCR's transition delivery review identified that in February 2014 the chairs of the Assessment Management Group and the Transition Delivery Group agreed that responsibility for monitoring standardisation would revert to the Assessment Management Group after high-level sign-off of the new arrangements (p 8). However, the relationship between the two groups may not have been clearly understood by all concerned.
83. In relation to the progress of marking in July, there is evidence of some formal risk escalation from the daily meeting of Assessment Management Group members to OCR's Senior Management team as a whole and on an individual basis as marking progressed. The Senior Management team decided on greater use of financial incentives for assessors.
84. OCR's Senior Management team members did not attend daily marking monitoring Assessment Management Group meetings, although they were copied into the daily circulated dashboard and watch lists. There is some evidence to support the Senior Management team's assurances that they were sighted on marking through the dashboard and watch lists and did informally question marking progress. OCR's final review report stated: “Marking progress continued to be monitored daily by both the AMG... and separately by a group of three directors covering Operations, Assessment Standards and Risk and Compliance and the CEO” (p 7).
85. On 24th July, two members of the Screening team in OCR's Research and Technical Standards identified concerns that imminent deadlines for providing results data to Ofqual and UCAS would be missed due to marking shortfalls. This issue was escalated to the CEO, the Director of Risk & Compliance (who was on leave), the Director of Operations and the Director of Assessment Standards. OCR's final review report identified that following this escalation “a new crisis management team was formed on 31 July to manage the situation” (p 8). Our understanding from the emails provided to us by the Director of Risk & Compliance (who had returned from holiday on 30th July) in his role as Chair

of the Crisis Management team was that this team may not have been fully functioning until 2nd August. However, when it was formed, the Crisis Management team included all relevant OCR directors as well as the Head of the Examinations & Assessor team and the Head of Assessment Standards. It also included the Chief Executive/Responsible Officer who had been contacted by OCR colleagues and had returned from leave early on 2nd August.

### **Area of investigation 3: the awareness and understanding of OCR management through the summer of risks to delivering marking on time**

#### **Decision-making in response to issues affecting marking**

86. A key decision in relation to the scoris web assessor functionality incident in early June was whether to revert to the classic platform. The evidence suggests that OCR drove the decision to revert to scoris classic and was appropriately involved in developing the options and more detailed planning for reversion.
87. Some decisions during the main marking period, the timing of management interventions and lack of understanding of roles and responsibilities when making those decisions, suggests that more work needs to be done to reinforce and embed OCR's new governance framework.
88. First, the lack of clarity about the role of the Assessment Management Group in effectively managing the marking and understanding, where necessary, where to escalate issues. This has been explored in "Area of investigation 2" above.
89. Second, and linked to the lack of clarity about the Assessment Management Group, OCR's final review report identified that: "The accountability of Senior Assessors was... not made clear enough in terms of their responsibility to manage the teams to achieve targets..." (p 50, Finding Post-AS14). The finding goes on to identify that OCR was "still devising internal processes and working instructions at the time we were communicating with assessors". And: "This responsibility of managing teams to achieve marking deadlines had in most cases been taken up by the QMs in the past."
90. Third, as marking rates deteriorated towards the end of June, it was misjudged, in hindsight, not to resort to financial incentives earlier and in a more incremental way as marking went further into July. The time taken by OCR management to make the decision to offer financial incentives, in face of the emerging situation, came late (at least one senior manager acknowledged that in hindsight this should have come seven to ten days earlier) and was then rapidly extended. OCR's marking position might have been stronger if the use of financial incentives had been more progressively implemented from an earlier time. In interview, it was clear that there was a conflict between offering greater

financial incentives when, as OCR felt, some assessors were exploiting the marking situation. Alongside concerns about the precedent this may set for future series. Consideration of these issues contributed to the delay in the decision. The proportion of assessors who “took commercial advantage” of OCR was not information specifically requested by the investigation.<sup>16</sup>

91. Fourth, OCR’s preliminary review report identified as an example of “poor quality decision making” the decision at the outset (before marking started) not to bring in temporary staff to support the Examinations & Assessor team, preferring instead to use overtime to maintain cover. Management stepped in late to remedy this. It would appear that the resourcing of the Examinations & Assessor team factored in the use, at peak times, of temporary staff. An example of this was when Coventry Deployment inherited the panel preparation from Cambridge Deployment and had to carry out remedial work. While the Director of Assessment Standards agrees it was the wrong decision not to bring in temporary staff earlier, the Head of the Examinations & Assessor team does not, taking the view that it was better for efficiency and morale of the Examinations & Assessor team to use overtime for existing trained staff rather than train up temps.
92. OCR’s final review report stated: “Once this issue was recognised and due to the number of new staff transferring to the Coventry Deployment Team a decision was made to transfer the temp budget to overtime budget as the existing Coventry team felt to work overtime in order to rectify the panels was the most efficient option over training new and temporary staff” (p 16, Finding Pre-E&A01). Ultimately, due to the scoris functionality issues, a lack of resources did affect the level of support that the Examinations & Assessor team could provide to assessors calling and emailing about scoris performance issues (many were not answered in a timely way). On the flip side, there was some support for the view put forward by OCR’s preliminary review report that the decision to centralise the Examinations & Assessor team on one site made crisis management easier.
93. Fifth, there may not have been sufficient use made of former qualification managers who had experience of closely managing the marking in latter stages of the series. On the face of it, the Assessment Management Group action matrices show that it was only from 29th July that former qualification managers were involved in the daily watch list actions and used to contact assessors

---

<sup>16</sup> As part of factual accuracy checking of this report, OCR was invited to provide data on the number of assessors as a whole or by subject who took commercial advantage of OCR. OCR is of the view that the time required to draw out detailed analysis would outweigh the value it would add to the report.

regarding their availability to mark. It is not possible to know whether earlier involvement of a number of former qualification managers in the latter stages of marking would have significantly altered the marking situation towards the end of July.

94. Finally, it was only recognised on 2nd August that a number of units on the worry list were AS units and so were not entirely made up of candidates who were seeking university places. OCR's preliminary review report stated: "Had this been realised sooner resources could have been more effectively targeted earlier on."

### **Factors affecting management's awareness and understanding**

95. A general point in relation to the scoris functionality issues during June is the context of OCR operating in a syndicate structure. This arguably means OCR relies heavily on a syndicate function (the Group Infrastructure Services Directorate) to hold the necessary technical expertise to challenge the syndicate's third-party electronic marking provider about scoris performance. This has been explored in more detail in "Area of investigation 1" above.
96. In relation to assessing the increasing deterioration in marking rates and the implications for achieving key marking deadlines during July, there is evidence that OCR senior managers worked on the basis of a number of flawed understandings and assumptions.
97. First, that there was insufficient holistic understanding of the end-to-end process of marking through to issue of results. When the scoris performance functionality issues arose in June there was insufficient awareness at the time, or more critically as marking progressed into July, of the link between this and the potential impact on, and implications for, achieving marking deadlines later on in early August. It is possible that the previous qualification managers would, generally, have had this picture.
98. OCR's final review report identified that: "There is less evidence of a common understanding of, and central responsibility for, the end-to-end awarding and marking processes" (p 4). It was evident from interviews and the lack of cross-business joined-up planning documentation that OCR did not have a strong culture of strategic or operational end-to-end approach towards marking for summer 2014 – rather, more a silo business approach. The final review report went on to identify that: "OCR thinks of itself in organisational/structural terms as 'What Operations does' and 'What Assessment Standards does' rather than focusing on the end-to-end process that drives outputs" (p 4). Furthermore, still referring to planning and management of risks, the final review report stated: "Without a clear understanding of, or responsibility for, the critical milestones

throughout the end-to-end process it wasn't possible to understand the consequences of missing them" (p 11).

99. Second, there was insufficient clarity between the Director of Assessment Standards/Head of the Examinations & Assessor team and the Director of Operations about whose responsibility it was to ensure that results were published to agreed timelines and hence where the primacy of ownership and responsibility laid for this.
100. Third, there was insufficient awareness about the availability of assessors. OCR's own review, supported by interview evidence, identified that not knowing that a number of assessors were imminently going on leave or had other commitments later on in July, had a particular impact on marking rates at this critical time - when completion of marking to key deadlines was already at risk in a number of subjects.<sup>17</sup>
101. Fourth, OCR did not sufficiently understand the exact definition and expectation of UCAS about what the "main results file" to be transferred to UCAS on the initial deadline of 7th August meant – particularly, in terms of the percentage of GCE marks that should be on this file and what was then acceptable to be sent through in subsequent amendment files up to 12th August (the last upload date before issue of A level results on 14th August). It would appear that the anticipated or acceptable volume of late marks beyond 7th August was not clarified between OCR and UCAS until the meeting between OCR, Ofqual and UCAS on 31st July. A meeting did take place between OCR Operations and UCAS, with follow-up emails, from 25th to 28th July, but this appears to have been more about timings of files. Ofqual's note of the 30th July telephone call between Ofqual and the then UCAS Director of Operations shows that when asked to give an indication of the percentage of marking missing from the main file upload on 7th August that would be "significant" he said 5 per cent, but that he would need to talk to colleagues as UCAS did not have an agreed position on this.
102. Key OCR senior managers viewed 12th August as the UCAS critical deadline and did not have a clear understanding of UCAS's expectations about the main and amendment mark files. When interviewed, the Head of the Examinations & Assessor team appeared unclear about the UCAS dates, but finally referred to 12th August. The Director of Assessment Standards could not recall the last marks upload dates at interview, but they were known to her during the series. While not documented by OCR, it was reported to the investigation that an OCR

---

<sup>17</sup> Refer to sections "Area of investigation 1" and "Area of investigation 2" in this report.

manager attended meetings with UCAS in late 2013. The manager reported that UCAS pushed for earlier data from the exam boards but was told that earlier data would be less complete. UCAS seemed happy with this if it meant it could start processing university applications sooner. The UCAS transfer date was included in OCR's General Qualifications Results Processing Schedule that was circulated early in 2014, although, as confirmed by the final review report, this date "did not subsequently factor into the marking targets or discussions once OCR was looking at contingencies to counter delays in marking" (p 53, Finding XB01). The final review report also points to an agreement between exam boards and UCAS dated June 2013 that implies that UCAS accepted that the initial data file would not be complete and that no statement as to the acceptable completeness of the initial (main) file was provided (p 53, Finding XB01).

103. However, on 11th July, the Chief Executive of UCAS wrote to OCR's Chief Executive stating: "We have reached the point in the annual cycle of events where I write to you to seek your assurance that there are no anticipated issues with the supply of results to UCAS in early August." The letter went on to state: "The agreement that we have with you, through [JCQ], stipulates that UCAS must receive your results information in the agreed format by 12 o'clock noon on Thursday 7 August." On 31st July, OCR's Director of Operations emailed UCAS confirming that OCR "will issue the main file with results information by mid-day 7th August. Any amendment files will start Saturday 9th August." While there is nothing in the letter that defines the percentage of marks in that main file (UCAS's expectations of the completeness of marks on 7th August), it implies that the expectation is for complete matched data for all subjects so it can be made available to universities on 9th August, rather than incomplete data.<sup>18</sup>
104. OCR demonstrated a lack of understanding of the critical role that the data from exam boards plays in the very early days of allocating university places. It appears that only after the joint meeting between OCR, Ofqual and UCAS on 31st July did OCR senior managers understand fully the significant effect that an anticipated shortfall of marked A level papers would have on a significant percentage of students going into clearing.

---

<sup>18</sup> As part of factual accuracy checking, OCR has made two points about this statement. First, custom and practice of providing data to UCAS over a number of years does not bear out this statement. Second, OCR cannot reasonably be expected to infer from the tone of a letter what it is that UCAS requires. Rather, we would expect the requirement to be explicit.

105. OCR's final review report stated: "OCR had no way of knowing that certain number of missing results could impact upon clearing in the manner referenced by UCAS at a meeting between OCR, DfE [Department for Education], Ofqual and UCAS on 31 July. This meeting was the first time UCAS mentioned such a threshold and impact" (p 53, Finding XB01). However, it should be noted that the 11th July letter from UCAS to OCR (referred to above) clearly identified that: "The admissions process has a critical dependency on your supply of results and any delay or problems with the format or contents of results files could have a significant impact on the admissions process..." OCR did not have a level of understanding about the impact that missing results could have on A level students in the university admissions process. The investigation recognises that it is not known whether senior managers in other exam boards have this understanding. This, in turn, raises the perspective of the industry-wide nature of this situation and the need to explore whether greater cross-industry clarity on this issue is needed.
106. Fifth, there was also a lack of understanding of some end-of-marking business processes that are necessary for the completion of marking and therefore the issuing of results. Key senior management at OCR acknowledged not fully understanding some end-of-marking processes that affected delivery of results to centres, such as the completion of marking necessary for the calculation of Z scores and realisation that some AS units could contribute to aggregation for A level candidates. With reference to the reconciliation of marking processes, OCR's final review report identified that: "This is a complex process which requires input from E&A and Ops. There are no clear lines of responsibility with regards to the tasks necessary at the end of the marking period..." (p 31, Finding Post-A&A07). And: "E&A did not understand its role in this process..." The report also stated: "Due to the complex process the ownership of this will never be clear."
107. In addition, OCR's final review report identified that the "current arrangements for monitoring examiners who have been stopped post-standardisation do not support effective pre-results checks" (p 27, Finding Post-AP02). This hampered the timely and accurate provision of information to Operations so that incomplete re-marks could be suppressed through the use of the Q score (pending status).
108. Sixth, on 25th July, Ofqual received very late notification from OCR's Director of Assessment Standards that there was a risk OCR could not provide all the A level outcome data (deadline 28th July) in compliance with Ofqual's data exchange document, which is published as a regulatory document. Ultimately, OCR did manage to comply. The data exchange outcome data deadlines were not identified on OCR's June 2014 GQ Results Processing Schedule, which identified all key summer series dates, which may have contributed to what

appears to be a lack of visibility about the deadlines from some senior managers.<sup>19</sup>

109. Seventh, senior assessors, as a body, did not fully understand their role in managing progress of marking within their teams. In addition, it is probable that it was not fully understood that some senior assessors needed more support than others to use scoris and prepare for standardisation meetings. OCR's final review report identified that: "The accountability of Senior Assessors was... not made clear enough in terms of their responsibility to manage the teams to achieve targets..." (p 50, Finding Post-AS14). The finding goes on to identify that OCR was "still devising internal processes and working instructions at the time we were communicating with assessors". And: "This responsibility of managing teams to achieve marking deadlines had in most cases been taken up by the QMs in the past."<sup>20</sup> OCR's final review report stated: "What became evident was... some senior examiners had become over reliant on the QM/L [Qualification Managers/Leaders]...." (p 30, Finding Post-E&A04).
110. Eighth, the degree of support required by assessors in terms of queries during a summer series was not understood. In previous series, a majority of these would have been dealt with by the quality managers. This resulted in the Examinations & Assessor team having unexpected demand, without more resources, placed on parts of the team, particularly deployment and monitoring and support functions. The impact was confirmed by OCR's final review report: "Es&As did not receive the support we aimed to provide because calls and emails coming in to E&A were not always answered in a timely fashion... resulting in us not meeting our SLA [service level agreement] to Es&As" (p 32, Finding Post-E7A09).
111. Another factor that may have affected some senior managers' awareness and understanding is the lack of experience in their current roles of managing marking and awarding of an OCR GQ summer series. We understand that a number of the top team, specifically the Director of Assessment Standards and the Head of the Examinations & Assessor team, had not been in their current roles managing an OCR GCSE and A level summer series of marking and awarding before summer 2014. The Director of Assessment Standards was OCR's previous Head of 14–19 Qualifications a number of years ago, and previous to her current role was the Responsible Officer equivalent in Bahrain

---

<sup>19</sup> There is more detail about this in the section "Area of investigation 5" in this report.

<sup>20</sup> Ofqual has not seen any documentation that outlines the scope of the Senior Assessor's responsibility to manage teams to achieve targets, so it is difficult to assess the weight of this finding.

overseeing the whole process and standards. The Head of the Examinations & Assessor team was appointed Interim Head of Assessor Management in April 2013. However, for the 2013 summer series these responsibilities did not include deployment and monitoring of assessors.

112. Finally, the holiday absence of some key senior managers during June, July and early August could also have impacted on their awareness and understanding of both scoris performance issues and when marking rates deteriorated. Our understanding is that the Head of Assessment Standards was on leave from 2nd to 12th June, the Director of Operations was on leave from 23rd to 27th June, and the Director of Risk & Compliance returned from ten days' leave on 30th July.
113. We also understand that the Chief Executive Officer/Responsible Officer started leave on 27th July, intending to return around 9th August. He returned from leave for the meeting of the Management Board for GCSE and A Level Reform in London on 30th July, then returned to his holiday. In interview, the Chief Executive Officer stated that he was in communication with his senior managers and returned from leave on 2nd August to join the Crisis Management team.

#### **Area of investigation 4: accuracy and completeness of OCR's reporting to Ofqual through the summer incident**

114. OCR's chronology of events indicates that OCR first informed Ofqual of the scoris performance issues on 2nd June (preliminary review report, Appendix 1). Our records show that 6th June was the first recorded notification (an email to the Acting Director of Regulatory Operations), followed by an event notification on 13th June. It was probable that in the first week in June, OCR, the Infrastructure Services Directorate and RM did not fully understand the nature or scale, and therefore the implication, of the emerging functionality issues. A number of interviewees expressed the view that they found it difficult to get satisfactory responses from the Infrastructure Services Directorate and RM. OCR's final review report stated: "As a result of system problems with scoris web assessor approximately 10 marking days were lost. It was a number of days before the full extent of the problems was identified" (p 30, Finding Post-E&A05). And: "It was not clear during the early part of the marking window as to the precise nature of the problem unfolding. There was a considerable amount of 'noise' coming into OCR... This intelligence was not collated in a central point and hence early warning of the impending problem was missed" (p 49, Finding Post AS13). However, by 10th/11th June, OCR, RM and Infrastructure Services Directorate crisis teams had been formed. Towards the end of the second week, OCR provided more detailed daily updates on scoris performance as it developed a better understanding of what the performance issues were and the

effectiveness of solutions being adopted by RM and the Infrastructure Services Directorate.

115. During the week commencing 16th June, OCR's daily email updates to Ofqual included a range of information required by Ofqual: total number of standardisation meetings completed relative to plan and 2013; percentage cumulative marking completed; churn rate of markers in 2014 relative to plan and 2013; total panel vacancies relative to 2013; and components being switched to scoris classic. The 17th June data identified that: the cumulative percentage of marking completed was 18.18 per cent, compared to 19.30 per cent at the same point in 2013; the dropout of assessors had risen from 1,024 the day before to 1,068 on 17th June, compared with 664 in the same week in 2013; and there were 1,092 panel vacancies (7.05 per cent) compared with 743 (5.6 per cent) at the same point in 2013. On 19th June, the Director of Risk & Compliance in an email update confirmed that OCR was "close to returning to a business as usual operation". In addition, he advised that, given that the focus was now on marking rates, information on panel vacancies and marker churn rates had been excluded. He also confirmed that total percentage of marks received was 22.63 per cent, compared with 22.62 per cent at the same point in 2013.
116. On 20th June, it was agreed at a teleconference between Ofqual and OCR that OCR's daily reporting would be replaced with weekly reporting (every Friday) – reporting the total number of standardisation meetings completed, cumulative percentage marking completed, and system stability. On 25th June, Ofqual wrote to OCR summarising the teleconference that took place on the 20th June and requesting a further teleconference on 27th June and also that OCR provide on a weekly basis, starting 3rd July, the percentage of marks on file at the time of each award. OCR provided this data by updating an Ofqual template each Friday (with the exception of the first Friday where OCR explained that no awards had taken place as several had been rescheduled).
117. On 27th June, OCR reported to Ofqual that it was back on track. As a result, Ofqual stood its incident down, not requiring daily or regular reporting of scoris performance-related data other than the percentage of marks on file. Over the next two weeks or so, two further marking issues (double marking and special consideration issues) were reported to Ofqual. Following this, on 15th July, Ofqual requested further marking data from OCR (specifically the percentage of marking completed by component compared to the same point last year) and information about marking tolerances. On 17th July, OCR responded asking why we wanted this information (our email did not explain why), given that it had agreed with Ofqual that the only reporting would be weekly percentage of marks on file at time of award. A further email exchange took place on 17th July providing some clarification to OCR, but OCR continued to challenge the

request. Ofqual sent a final email advising OCR that Ofqual would meet with OCR the following day to take this request forward.

118. OCR expressed concern that from its perspective, the request for data distracted from what it was trying to achieve and pressed Ofqual for a rationale. When Ofqual met with OCR, OCR provided Ofqual with the scoris dashboard and watch lists, among other information. This marking data could have been offered earlier. OCR expressed to the investigation that it took some time for Ofqual to clarify the reasons for its request so that OCR understood the reasons. When clarification was given, the data was provided promptly.
119. Following Ofqual's meeting on 18th July, OCR agreed to provide the scoris dashboard and watch lists on a weekly basis. From 26th July, OCR complied with Ofqual's request to provide the scoris dashboard and watch lists on a daily basis.
120. The data and other information contained within the scoris dashboard and watch lists were based on best available marking rate data (OCR's cautioned that all live time data can be affected by small time delays). The data was based on an automated system using previous days/weeks marking rates and then RAG-rated dependent on/prioritised against the award date. When Ofqual requested the dashboard and watch lists from 18th July, the same dashboard and watch lists were provided as those provided to OCR senior managers.
121. In early July, OCR notified Ofqual of a double-marking issue. OCR's timeline of key events in its preliminary review report indicates that it notified the regulators of a double-marking issue on 8th July. It is unclear how this was done as we have formal notification of this incident recorded as 9th July.<sup>21</sup>
122. In relation to the special consideration issue, OCR's own timeline does not record the date it first notified us of this, but we have formal notification recorded as 11th July.
123. On 24th July, two members of the Screening team in OCR's Research and Technical Standards identified concerns to OCR's Risk & Compliance team that imminent deadlines for providing results data to Ofqual and UCAS would be missed due to marking shortfalls. The Risk & Compliance team escalated the issue (by email) to the Chief Executive Officer, the Director of Risk &

---

<sup>21</sup> OCR, as part of the factual accuracy checking, stated that it believes the preliminary review report cited an incorrect date. It gave the date that OCR circulated its internal draft notification rather than the date that the formal notification went to Ofqual.

Compliance (who was on leave), the Director of Operations and the Director of Assessment Standards.

124. On 25th July, OCR notified Ofqual by email that outcome data for 13 A level specifications were at risk of not being submitted to Ofqual in line with the summer 2014 GCSE and A level data exchange document (published as a regulatory document). The deadline for submission was 10.00am on 28th July. Ultimately, all specification outcome data was provided by the submission date, although for several subjects it was not fully complete. (It excluded partial absences – but OCR had made Ofqual aware this might be the case some months earlier. At that point, it was accepted in principle, as it was not considered to materially affect the review of outcomes.) However, the data exchange submission deadlines were not identified within OCR's summer 2014 GQ Results Processing Schedule, which identified key summer series deadlines. OCR's final review report, referring to awareness of Ofqual data requirements, identified that "there is perhaps not sufficient awareness outside of Assessment Standards [of] the importance of all of the dates and particularly as happened in the summer, the implications for not having sufficient data available at the given date" (p 54, Finding XB02).

### **Area of investigation 5: whether OCR provided appropriate assurance to Ofqual through the summer marking incident**

125. There were a number of points during June and July where OCR provided information (either by email or recorded when summarising a meeting with OCR) that Ofqual could reasonably take as assurances.
126. First, an assurance in relation to the scoris web assessor functionality issues. On 9th June, OCR's Director of Risk & Compliance provided an email update to Ofqual (Compliance & Monitoring team) about the scoris performance issues, stating: "Given the early stage of marking within OCR, overall marking progress remains on track." Although it was probable that marking had been affected by the functionality issues, and indeed continued to be during June (OCR's final review report stated that ten marking days were lost), given the relatively early point of marking then it would be difficult to prove that "overall" marking was not still on track at this point.
127. The email went on to say: "The solutions identified on Friday were successful with performance back to the levels expected." A review of communications between the Infrastructure Services Directorate and OCR does bring into question whether this was appropriate in terms of providing the full picture. On 9th June, the Director of Risk & Compliance circulated an options paper about what the Plan B should be: "Given the performance of issues noted during this last week, OCR needs to be clear on what its Plan B is if the issues continue and performance is unacceptable." Also, later that day, an email from the

Infrastructure Services Directorate to OCR (the Director of Operations but not including the Director of Risk & Compliance) identified that, while the platform continued to operate comfortably within capabilities, “there remains a small number of significant functional issues that occur consistently”. On 10th June, RM’s help desk extended its opening hours from 09.00 to 22.00 until further notice. An email on 11th June from the Infrastructure Services Directorate to crisis team members (but not including the Director of Risk & Compliance) referred to having to agree a clear implementation plan for switching to scoris classic. Another email (but again not including the Director of Risk & Compliance) stated: “RM believes that so far 25% of the untrustworthy constraints in the database have been updated and are now trusted.”

128. Second, an assurance about meeting awarding deadlines in an email of 18th June from OCR’s Director of Risk & Compliance to Ofqual as part of OCR’s daily updates<sup>22</sup> stated: “We are watching this very closely and clearly will be taking actions where we believe that there is a heightened risk that we might not hit the Awarding and Last Marks Upload deadlines. At the moment we are confident that we can still meet these deadlines but of course this picture varies day to day as we progress further into the marking phase and more units have their marking completed.” Although this assurance is caveated with “the picture varies day to day”, at this point it did not appear to fully consider the future potential implications of assessor dropout and panel vacancy data identified in OCR’s daily update email of 17th June to Ofqual. In addition, at this point in June, as confirmed by OCR’s final review report, a significant number of marking days had been lost (p 30, Finding Post-E&A05). In interview, the Director of Risk & Compliance acknowledged that with hindsight assurances provided to Ofqual were overly optimistic, but they were genuinely based on OCR’s understanding of the management information it had.
129. Third, there were statements from OCR, contained within daily scoris performance updates on 19th and 20th June, that OCR was close to returning to business as usual. Both confirmed that standardisation meetings continued to complete on time and to plan during the day, that cumulative percentage marking completed for all subjects was very close to the same point last year, and that the system was stable. However, OCR also identified that it was concerned about 22 per cent of the total number of units where standardisation was complete where an interim or final marking deadline had not been met or was in danger of not being met. It is not clear from the information how this 22 per cent compares with the same point last year, but the 19th June reference

---

<sup>22</sup> This assurance was also referred to by the Chief Regulator in her letter to OCR’s Chief Executive Officer on 6th August.

does identify that the number of units had risen by nine (16 per cent) since the previous day. However, the email is also caveated that all figures fluctuate on a daily basis and so not to read too much into individual figures.

130. Fourth, Ofqual's *Summary of OCR's Issues Summer 2014* stated: "OCR gave assurance on 27th June it was back to business as usual levels and it had not experienced any further performance issues with scoris during the week." It is not clear whether this assurance is referring to the performance of scoris or marking rates (compared to same point last year), or both. Ofqual's specific note of the afternoon meeting on 27th June between Ofqual (Entry, Exit & Enforcement Manager and Monitoring Manager) and OCR (Chief Executive Officer and Director of Assessment Standards) does not record use of the words "assurance" or "back to business as usual levels". However, the information provided by OCR strongly suggested that scoris was performing properly and that marking was not an issue for concern: "...that 33% of marks were now on the system.<sup>23</sup> This is slightly behind the position in 2013 but nothing that OCR is concerned by."
131. Fifth, on 8th July, OCR (in an email from the Director of Assessment Standards) provided a response to Ofqual's request for assurance following OCR's notification that a number of award dates had been rescheduled. OCR stated: "The more we move into the middle and towards the end of the session the less likely it is that we will move award dates. I would think that it would be highly unlikely that we will move any awards after the end of this month." It concluded by saying "and to reassure Ofqual that we are on track to meet all of the key dates for awarding this summer".<sup>24</sup> However, on 14th July, OCR sent an updated awarding schedule (in response to Ofqual checking, as part of the data exchange procedures, why OCR had not submitted some outcome data following award meetings) showing a further 12 GCSEs and three GCEs with changed awarding dates. In ten of these (almost all GCSE) the award date was pushed back, several by a significant number of days. The Director of Assessment Standards stated that she was not aware of any rescheduled award dates other than the ones notified more formally to Ofqual on 3rd and 4th

---

<sup>23</sup> The problems with scoris assessor functionality from early June through to when performance stabilised in mid to late June did have an impact on marking rates as June progressed into July and towards marking completion deadlines. Yet while linked they are also separate incidents and should not necessarily be conflated when considering the evidence.

<sup>24</sup> This assurance was also referred to by the Chief Regulator in her letter to OCR's Chief Executive Officer on 6th August and copied to the Minister Nick Gibb.

July. We have not been able to determine whether these awards were rescheduled before or after the assurance given on 8th July.

132. We identified a total of 19 award meetings delayed (counting GCSE Geography B twice as it was delayed twice). In her email of 29th October, the Director of Assessment Standards said that she thought only 16 award meetings had been delayed, noting that this was a low proportion (8 per cent) of the total number of award meetings (192) and “would not judge that as significant”. She added that nine of these meetings were delayed because of scoris issues. These delayed meetings were “one day’s worth of screening”. As OCR had “more than one day screening contingency” in its awarding schedule, these delays were “not an indication of seriousness”.
133. Although this analysis is accurate, it does not acknowledge the relatively high proportion of award meetings scheduled for early July that were delayed. Our analysis of awarding meeting data shows the following:
- All of the first five award meetings (26th June to 8th July) were delayed – although OCR attributed only one delay to scoris-related issues.
  - Up to and including 16th July there were 18 award meetings originally scheduled, 11 of which were delayed – four due to scoris-related issues.
  - Ninety-nine of the 122 GCE award meetings were always scheduled to take place in the very last week of OCR’s awarding schedule (21st to 25th July), where the scope to delay these meetings is much less if marking deadlines are to be achieved.
134. These delays could have been an early warning indicator of the marking issues that materialised later in July. However, we do not know how many award meetings were rescheduled in previous summer series due to changes in OCR’s auditing of its awarding schedule. Without this data it is not possible to determine if this situation with delayed awarding meetings in the first half of July was unusual.
135. OCR’s final review report also made some relevant findings. First, it identified that: “Daily AMG meetings identified at risk qualifications in terms of having less than 85% of the marks on the system in time for the award, however as this data was at unit level only it did not account for the number of aggregated candidates at qualification level. In many cases where individual units were above 85%... the overall % of graded candidates (aggregated) was below 85% and the awards could not be screened” (p 36, Finding Post AS05). Second: “Whilst delays to awards on the whole were due to insufficient marks available at the time of the meetings, a number of other issues had a significant impact on the timely running of award meetings. The resource requirement for the

agreed support to awarding... was underestimated” (p 37, Finding Post-AS06). Third, it identified that: “In the June 2013 series Chairs reported ‘any other difficulties at award’ for 27 specifications, in June 2014 this increased to 81 specifications.” Fourth, it identified that: “OCR’s awarding schedules (based on old business rules) did not have enough in-built contingency to account for the impact of changed processes and unforeseen marking issues” (p 36, Finding Post-AS03).

136. Another indicator of marking progress is the proportion of whole candidate marks on file at the time of each award. The data provided to us as part of data exchange showed 203<sup>25</sup> GCSE and GCE awards made in summer 2014. Of these, 37 awards (18 per cent) were made with fewer than 85 per cent of candidate marks on file at the time of the award meeting. In comparison, in summer 2013 only around eight award meetings (4 per cent) were made with fewer than 85 per cent of candidate marks on file. However, of the 37 awards that took place with fewer than 85 per cent of marks on file, only one meeting took place in the first week of awarding, four meetings in the second week and 32 in the third and final week – therefore, arguably, this had limited use as an early warning indicator.
137. However, if we look below the 85 per cent benchmark to the awards where there was a significantly lower percentage of marks on file than for the same award in summer 2013, the data shows a marking ‘gap’ emerging by mid-July. Our analysis indicates that in summer 2014, 51 awards were made with significantly fewer whole candidate marks on file at the time of the award meeting than in summer 2013.<sup>26</sup> In the first week of awarding there was only one such award. But in week two (15th to 21st July) there were ten awards: five where the difference was between 5 and 10 per cent, and five where the difference was greater than 20 per cent. This was an indicator that in some subjects, significantly less marking progress had been made than at the same point in summer 2013. This picture became more marked by the final week of

---

<sup>25</sup> Figures differ from the total cited by OCR (192). This is due to differences in methodology – for example, we have counted the eight different GCSE Art & Design specifications as separate awards.

<sup>26</sup> For the purposes of our analysis, we have assumed that all of the 2013 award meeting dates were held on the same dates as the summer 2014 schedule. We have defined a “significant” difference as greater than 5 per cent – for example, 90 per cent of marks on file in summer 2013, 85 per cent in summer 2014. OCR has questioned the above assumption, stating that the factors that affect awarding dates are many and varied, and it is unlikely that awarding meetings were scheduled on the same day.

summer 2014 awarding, with 40 awards made where the proportion of marks on file was at least 5 per cent lower than in summer 2013.

138. Sixth, on 18th July, Ofqual met with OCR. The note of that meeting records the following:

- Eighty-three per cent of scripts had been marked compared to 89 per cent on same day in 2013. OCR had marked over 500,000 more scripts than at the same point last year, given that the total volume to be marked in 2014 was significantly higher than in 2013. There were approximately 200,000 GCE scripts and 300,000 GCSE scripts still to be marked.
- OCR stated that it did not have visibility of individual assessors' marking plans.
- OCR noted that it had a slower start to marking due to scoris performance issues.

139. OCR also stated that at that time there was no GCSE/GCE qualification where it was believed that results would not be ready in time for results day. In interview, the Director of Assessment Standards stated that what she meant was that all results would be with centres by results day. The watch (worry) list for 18th July showed 22 GCE units with projected completion of marking dates at current daily/weekly rate beyond even the very latest last mark upload to UCAS deadline (12th August) and about half of those had less than 70 per cent of marking completed. Also there were five GCSE units (not including another five where dates were entered as "unknown") beyond 18th August (the deadline on the GQ Processing Schedule), with only one being above 70 per cent marking complete.

140. In addition, OCR's final review report identified that: "Marking once again began to deteriorate in the week of 14 July and the traditional tools available to OCR such as the instigation of marking centres... were deployed. However, the impact on marking rates of these interventions was slow and failed to have a material impact on marking rates ahead of one of the regular marking updates meetings with Ofqual on 18 July" (p 8). There are two points about this statement. First, it gives a sense that Ofqual and OCR were holding regular update meetings at this point in July. This was not the case: the only reason this meeting occurred was at Ofqual's insistence, due to OCR failing to provide information requested in previous days.<sup>27</sup> Second, this suggests that OCR's

---

<sup>27</sup> See the section "Area of investigation 4" (paragraph 94) in this report, which also outlines OCR's different perspective to that of the investigation team.

“traditional tools” were not enough to deal with a larger than anticipated assessor shortfall, although OCR expressed the view that it had no grounds to believe that the “traditional tools” were not enough to deal with the shortfalls that arose.

141. On 24th July, it was identified that imminent deadlines for providing results data to Ofqual and UCAS would be missed due to marking shortfalls. OCR's final review report identified that following this escalation “a new crisis management team was formed on 31st July to manage the situation” (p 8). There are two points about this. First, following escalation, it is unclear why members of the Research and Technical team felt compelled to escalate this and why, as it also appears, senior managers had not already identified the risk and acted upon this (for example, by bringing the Crisis Management team together earlier). No explanation has been provided for why this was the case. Second, our understanding from the emails provided to us by the Director of Risk & Compliance (who returned from holiday on 30th July) in his role as Chair of the Crisis Management team was that this team may not have been fully functioning until 2nd August.
142. Seventh, on 25th July, there were morning and afternoon teleconferences between Ofqual and OCR. The note of the afternoon conference records OCR stating: “They remain confident that all GCE and GCSE marking will be complete in time for the results days.” In interview, OCR explained that it was meant that all results would be with centres by results day. At this teleconference, OCR agreed to contact other exam boards regarding the possibility of using their assessors for GCSE and GCE English Literature and GCE Biology. Ofqual also received the first weekly update (as agreed at the 18th July meeting), which showed that the gap between percentage of marking completed at the same time last year, which had been 6 per cent on 18th July, was just 4 per cent. Sixteen GCE units were identified as high risk in terms of completion of marking (Ofqual document *OCR Issues – Summer 2014*). The 25th July watch (worry) list showed ten GCE units with projected completion dates beyond even 12th August (with six history units with projected completion dates into September/October or beyond) and four GCSE units (English Literature and History) with projected completion dates into September through to November.
143. On 26th July, a daily update on marking progress from OCR (from the Director of Assessment Standards forwarding an email from the Manager Data Reporting) identified that with 14 days to go before GCE Last Marks Load (working to a UCAS marks upload date of 7th August), there were 74,439 GCE script marks outstanding. This meant that 5,319 marks per day were required. It identified that the previous day's rate of marks on the system was over 7,000. For GCSE, 131,938 marks were outstanding, requiring a daily rate of 6,283,

with over 15,000 marks having been put onto the system in the previous day. The daily update on 27th July showed the previous day's rate of marks being entered onto OCR's system as still above the daily required rate – 5,950 for GCE and 12,950 for GCSE. On 28th July, OCR offered to pay assessors [REDACTED] the script marking rate for units most at risk.

144. Eighth, on 30th July, further assurance was offered by OCR – first, at the meeting of the Management Board for GCSE and A Level Reform in London, attended by Ofqual's Chief Regulator. She reported (by email) that OCR's Chief Executive Officer had attended the meeting (although on holiday) and had stated that OCR "were expecting to finish marking ahead of results day but he was less sure of his facts on the UCAS deadline". Second, the note of the afternoon teleconference between Ofqual and OCR records that OCR was still achieving the necessary daily marks for both GCSE and GCE. Resources had been found to identify assessors who had previously dropped out of marking and to contact them by telephone. In addition, 100 AQA assessors were available to mark GCSE English scripts and some CIE assessors would shortly start to mark GCE Biology scripts. The Director of Assessment Standards stated that OCR remained "quite confident" that all GCE marking would be complete in time for results day. In interview, the Director of Assessment Standards stated that she meant getting all results to centres by results day. The 29th July watch (worry) list shows eight GCE units with projected marking completion after the 12th August UCAS last marks deadline.
145. On 1st August, a joint meeting between Ofqual, OCR and UCAS took place in London. UCAS joined the Ofqual/OCR daily teleconferences. OCR identified four GCE subjects (English Literature, Geography, History and Sociology) that were not expected to have marking completed in time for the main file upload to UCAS on 7th August. Part of OCR's plan to address this included offering assessors [REDACTED] times the normal script marking payment as incentives to assessors.
146. On 4th August, based on the latest data from OCR and that UCAS now believed the marking situation was manageable, Ofqual took the view that OCR was likely to meet UCAS's expectation of main file upload date of 6th to 7th August.
147. In interview, the Director of Assessment Standards was asked for the rationale as to why she was confident about achieving marking deadlines and why this justified the assurances given to Ofqual. She stated that it came from being sighted on all relevant daily marking progress data and believing, as did her managers, that tried and tested means of increasing marking rates where necessary, such as increasing allocations to existing markers and use of financial incentives, would work as they had done in previous series. She

believed in the assurances being given, but stated that providing assurance is a difficult judgement and she had to consider factors such as the usual last-minute rush of marking that can bring marking within completion, and as such it is important to “hold your nerve”. There was an acknowledgement that, in hindsight, the assurances were overly optimistic and naive.

148. The Director of Risk & Compliance in interview also acknowledged that OCR's confidence about ultimately making all the marking deadlines was, in hindsight, overly optimistic. This was based on OCR's understanding of the management information it had.
149. OCR's final review report identified that the Director of Operations “did not feel he was aware of the lateness of the marking at the end of July, despite receiving the Watch list every day. The level of detail in the Watch list, allied to the lack of reference to the significant dates for data submission, would not have highlighted the seriousness of the marking delays without additional detailed analyses” (p55, Finding XB04). The final review report stated: “An appropriately focussed assessment of the current status of key elements of the series was not available to senior management.” While the watch lists provided to us by OCR do not show significant dates for submission of data, this position is still difficult to understand given that the projected marking completion dates based on the previous 24-hour and weekly marking rates are clearly shown in the July watch (worry) list and many of these dates were shown as well beyond results days in August.

We wish to make our publications widely accessible. Please contact us at [publications@ofqual.gov.uk](mailto:publications@ofqual.gov.uk) if you have any specific accessibility requirements.



© Crown copyright 2015

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](http://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: [publications@ofqual.gov.uk](mailto:publications@ofqual.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.uk/ofqual](http://www.gov.uk/ofqual).

Any enquiries regarding this publication should be sent to us at:

Office of Qualifications and Examinations Regulation

Spring Place  
Coventry Business Park  
Herald Avenue  
Coventry CV5 6UB

2nd Floor  
Glendinning House  
6 Murray Street  
Belfast BT1 6DN

Telephone 0300 303 3344

Textphone 0300 303 3345

Helpline 0300 303 3346