

# **Evaluating the summer 2015 results of A level French, German, and Spanish**



August 2016

Ofqual/16/6081

# Contents

1. Background.....	2
2. The 2015 summer assessments.....	3
Our research and exam board responses.....	3
Summer 2015 results .....	4
A level French .....	6
A level German .....	6
A level Spanish .....	7
Exam boards' analyses of their summer 2015 assessments .....	7
AQA .....	7
OCR.....	8
WJEC.....	8
Pearson .....	9
Stakeholder feedback .....	9
3. The 2016 summer assessments.....	9
Review of assessments for 2016 .....	9
4. Next steps .....	10
Awarding for 2016 .....	10
Consideration for reformed qualifications.....	10
Native speaker research .....	12
5. Conclusion .....	13
Table 1: Entry and percentages of students achieving A and above in A level MFL, 2013-2015 .....	5
Table 2: A level French outcomes at award, and differences from prediction for matched students at A and above .....	6
Table 3: A level German outcomes at award, and differences from prediction for matched students at A and above .....	7
Table 4: A level Spanish outcomes at award, and differences from prediction for matched students at A and above .....	7

## **1. Background**

In recent years, teachers and other stakeholders have expressed concerns about grading in A level modern foreign languages (MFL). These concerns are twofold: that too few students achieve the top grades compared to other A level subjects; and that the best students do not always achieve the best grades.

In June 2014, the Joint Council for Qualifications (JCQ), a membership organisation for exam boards, published a review<sup>1</sup> of MFL at A level which included an investigation into the proportion of A\* grades. In their report JCQ members committed to “review how mark schemes can best be applied”.

In September 2014, we published our own analysis<sup>2</sup> of the A levels in French, German and Spanish offered by AQA, OCR, Pearson and WJEC. This research was based on a detailed analysis of the design and working of all the exam questions and the associated mark schemes.

Our research found the assessments were effective in differentiating between students at lower grades, but less so at A and A\*. We recommended that exam boards increase the level of demand in some questions, so there were opportunities for the very best students to demonstrate their knowledge and skills in order to improve differentiation between the highest ability students.

We also found the mean marks in some elements of the assessments, such as speaking, were generally very high, and so these elements did not differentiate effectively between students of different ability levels.

One of the findings of the 2014 JCQ review was that students’ performance on the essay questions, in their exams, played a large part in determining whether they achieved an A\* or an A grade. This supported our findings that the majority of the other questions did not discriminate well at the top end of the ability range.

We recognise the design of the current specifications has a significant impact on the design of the assessments. Making fundamental changes to the question papers at such a late stage in the lifecycle of the specification is challenging. However, in response to these findings all the exam boards did make some changes to their assessments for the 2015 series and agreed to evaluate the effectiveness of those changes later that year.

This report mainly explains what we found when we looked at the 2015 assessments in these qualifications and how the exam boards responded to our findings. It also sets out some of the wider work we have been doing to address stakeholders

---

<sup>1</sup> <http://www.jcq.org.uk/media-centre/news-releases/mfl-review-press-notice>

<sup>2</sup> <https://www.gov.uk/government/news/improvements-to-be-made-to-a-level-foreign-languages>

concerns about the current MFL A level qualifications and explains how we have used the findings from the 2014 research to inform the reform of these qualifications.

## **2. The 2015 summer assessments**

### **Our research and exam board responses**

Our September 2014 research<sup>3</sup> set out a number of findings which we shared with each exam board. We identified the following concerns across the qualifications of the four exam boards:

- Some questions within the written papers did not differentiate effectively, particularly at the top end of the ability range.
- Not all optional questions were comparable.
- The spoken language assessments did not differentiate effectively, with most students scoring very high marks.
- There were also some adverse features of mark schemes, such as:
  - capping 'quality of language' marks;
  - application of scaling factors;
  - inconsistent design of levels of response mark schemes across languages and optional tasks.

Our research also set out some recommendations for all the exam boards. We recommended that they should:

- set out the principles for examiners to use when deciding what is an acceptable response, and how to mark responses that students appear to have pre-prepared ahead of the exam and/or those that simply lift responses from the stimulus material;
- undertake further analysis of mark data to understand why students' marks on different listening questions and on different writing questions were not well correlated;
- monitor the impact of any changes they made using appropriate metrics and report on this to Ofqual.

---

<sup>3</sup> <https://www.gov.uk/government/news/improvements-to-be-made-to-a-level-foreign-languages>

Exam boards reviewed the question papers and mark schemes they had written for use in 2015, in the light of the research and our recommendations, and made some changes. In doing so, they still had to comply with our published requirements such as the weighting and application of the existing assessment objectives and the current subject content. Some changed both question papers and mark schemes and others the mark schemes only.

## **Summer 2015 results**

For all A level subjects, exam boards predict the likely proportion of the national cohort that will achieve each grade, based on their prior attainment at GCSE. This prediction is the starting point for exam board awarding committees. For MFL, as with other subjects, we would normally expect outcomes to be reasonably close to predictions (either just above or below) unless there is evidence why this would not be appropriate, having taken into account students' work, and reports from senior examiners on how the exam questions worked. We monitor how closely each exam board meets those predictions and, before the results are issued, exam boards send their expected results for each qualification to Ofqual.

We made some changes to the way in which we expected A levels in French, German and Spanish to be awarded in 2015. We stressed to the exam boards that, for summer 2015, we did not expect the percentage of students achieving A\* in A level French, German and Spanish to be lower than those predictions, unless there was compelling evidence to justify this. In our September 2014 report, we modelled the potential impact on the A\* grade of improving the discrimination of the assessments.<sup>4</sup> Our analysis suggested that if the assessments discriminated more effectively between the very able students, then this might naturally lead to more students achieving A\*. We amended our rules so that, if this did happen, we would not constrain any such increase.

We wrote this into our regulations for summer awarding<sup>5</sup> so exam boards had to provide additional evidence for any French, German or Spanish grade boundary decisions which would have resulted in the percentage of students being awarded A\* being below that predicted (as well as any awards where the percentage was well above the prediction).

---

<sup>4</sup> See section 7 of the report

<sup>5</sup> Link to summer 2015 data exchange procedure:

<http://webarchive.nationalarchives.gov.uk/20151010034910/https://www.gov.uk/government/publications/data-exchange-procedures-for-a-level-gcse-level-1-and-2-certificates>

The table below shows the number of students taking French, German and Spanish, together with percentages of students who achieved grades A\* and A, for 2013, 2014 and 2015.

		<b>2013</b>	<b>2014</b>	<b>2015</b>
<b>French</b>	Entry	11272	10433	10328
	Grade A*	6.5%	6.6%	8.1%
	Grades A*/A	38.5%	37.6%	37.3%
<b>German</b>	Entry	4242	4187	4009
	Grade A*	8.2%	8.9%	8.3%
	Grades A*/A	41.4%	40.6%	39.1%
<b>Spanish</b>	Entry	7651	7601	8694
	Grade A*	6.7%	7.7%	8.1%
	Grades A*/A	36.0%	35.6%	35.0%

*Table 1: Entry and percentages of students achieving A and above in A level MFL, 2013-2015*

At A\* there was an increase in the proportion of students in French and, to a slightly lesser extent, in Spanish. In German, the proportion decreased by 0.6% from 2014.

We considered the data the boards provided and found that the boards' outcomes for 18 year-old students were in line with predictions at grades A and A\* and within the reporting tolerance.

We published our analysis of the comparability of grade standards in December 2015<sup>6</sup>. The extract below (tables 2, 3 and 4) shows the percentage of matched students (those 18 year-old students who had GCSE results from 2013) and the differences between the predictions for those students and the actual outcomes at the award. These percentages (of matched students) do not correspond with the percentages published for all students in August by JCQ. This is because the matched students are a sub-set of the overall entry. The overall entry will include students who might be re-taking the qualification, mature students who would not have taken their GCSEs in 2013, and other students who do not have GCSE results from 2013. Exam boards use predictions based on matched students to guide their

---

<sup>6</sup> See the third report and data here <https://www.gov.uk/government/news/summer-series-access-arrangements-and-inter-board-comparability>

decisions on grade boundaries. They then apply those grade boundary decisions to all students and they report the outcomes of those decisions to us in relation to matched students and in relation to the overall entry.

### **A level French**

Table 2 below shows that in A level French all four boards came within tolerance of the prediction, and all were above prediction at grade A and above.

Exam board	Outcome at award A*	Difference from prediction	Tolerance	Outcome at award A* and A	Difference from prediction	Tolerance <sup>7</sup>
AQA	8.0%	+0.7	+2%	38.9%	+0.5	+/- 1%
OCR	8.4%	+1.5	+2%	39.2%	+1.7	+/- 3%
Pearson	6.9%	+0.2	+2%	36.1%	+0	+/-2%
WJEC	7.6%	+1.9	+2%	33.0%	+0.1	+/-2%

*Table 2: A level French outcomes at award, and differences from prediction for matched students at A and above*

### **A level German**

Table 3 below shows that in A level German, AQA, OCR and Pearson were within tolerance and above prediction at both A\* and A. WJEC was within tolerance at A\* but slightly out of tolerance (0.1% above prediction) at A. We accepted WJEC's rationale for an award that was +3.1% above the prediction for grade A and above. It was not possible for WJEC to meet predictions for both A and A\*. The alternative, which WJEC considered, would have meant A\* being below predictions. WJEC judged that an outcome that was just outside tolerance at grade A was preferable.

Table 1 above shows that the percentage of the overall entry achieving A\* or A was lower in 2015, compared with 2014 and 2013, while table 3 shows that outcomes for the matched students were higher than the predictions. This suggests that the students who could not be matched to GCSEs in 2013 were, in general, weaker than in previous years.

---

<sup>7</sup> Tolerances are based on the size of the entry, with larger entry specifications having a narrower reporting tolerance

Exam board	Outcome at award A*	Difference from prediction	Tolerance	Outcome at award A* and A	Difference from prediction	Tolerance
AQA	5.9%	+0.1	+2%	34.1%	+0.2	+/-2%
OCR	8.6%	+1.9	N/A <sup>8</sup>	40.7%	+2.8	N/A
Pearson	6.5%	+0.9	+2%	34.7%	+2.7	+/- 3%
WJEC	5.4%	+0.8	+2%	32.1%	+3.1	+/- 3%

Table 3: A level German outcomes at award, and differences from prediction for matched students at A and above

### A level Spanish

Table 4 below shows that in A level Spanish all four boards were within tolerance (where tolerance was set) and in all but one case (AQA grade A) they were above prediction.

Exam board	Outcome at award A*	Difference from prediction	Tolerance	Outcome at award A* and A	Difference from prediction	Tolerance
AQA	7.8%	+0.5	+2%	33.9%	-0.3	+/- 1%
OCR	7.3%	+0.7	N/A	36.7%	+4.9	N/A
Pearson	6.0%	+0.2	+2%	28.9%	+0.1	+/-2%
WJEC	6.5%	+0.9	+2%	31.7%	+2.7	+/- 3%

Table 4: A level Spanish outcomes at award, and differences from prediction for matched students at A and above

### Exam boards' analyses of their summer 2015 assessments

This section provides a summary of the exam boards' analyses of the impact of the changes they had implemented in response to our recommendations.

#### AQA

AQA reported that, in response to our September 2014 research, they asked their lead assessment writers to re-consider the 2015 papers and identify specific questions where the level of demand could be increased. This resulted in small

<sup>8</sup> Where the number of matched students is below 500, no reporting tolerance is applied



changes to some questions and the introduction of further guidance for markers on the application of the marking criteria for the speaking assessment. They also removed the cap on marks for quality of language in the essay questions so that the marks were no longer limited by the mark given for the content of the essay (recommendation 4 in our September 2014 report).

AQA subsequently provided analyses reports to show the extent to which these changes had an impact on the difficulty of the papers and on the way in which they discriminated between students of different abilities. It appears from these reports that changes made by AQA had less of an impact than intended. However, the reports do show there are some examples of the mean mark being lower than in previous years and examples of the discrimination being better in some questions than in previous years.

## **OCR**

OCR reported that they carried out a comprehensive review of their assessments and mark schemes in response to our research and recommendations. OCR provided us with their report and several supporting documents. OCR reported that they decided not to make many changes to their questions or marking schemes as they were concerned this could have an adverse impact on candidates so close to the exam series.

In 2015, OCR focussed their improvements on their approach to marking certain items – the translation task at A2 and writing exercises at AS and A2.

The item-level statistical information available for the translation exercise showed that the item performed well across the three languages, with even mark distributions and relatively low facility indices<sup>9</sup>. OCR reported that they evaluated how the marking criteria for the essay questions were being used and they provided annotated practice scripts to the marking teams for the June 2015 series.

## **WJEC**

In response to our September 2014 research, WJEC reported that they made changes to their question papers and mark schemes for 2015 to ensure there was sufficient variety in item difficulty to discriminate more effectively between students at the top level. WJEC also submitted a technical report which included statistics about how well the assessments in 2015 functioned. These statistics appeared to show the changes made had a positive impact on the extent to which the A2 written paper discriminated effectively, at least when compared with the 2014 papers. The mean mark for the paper decreased for all three languages, when compared with the 2014

---

<sup>9</sup> The facility index for each question shows the percentage of students who scored full marks. A high facility index suggests that the question did not discriminate well.

papers, by half a mark in French, two marks in German, and by three and a half marks in Spanish.

## **Pearson**

Pearson reported the actions they had taken in response to our research. Their actions focussed on the mark schemes for assessments as our research did not find that there was an issue with the level of demand of their papers. For example, they introduced a new approach to marking the A2 translation question to avoid the need to scale the marks awarded. They convened a meeting of senior examiners for both the writing papers (units 2 and 4) and the speaking units to review the marking criteria. They amended the relevant marking guidance on their website.

Pearson also provided detailed statistical analyses, including facility and discrimination indices. This showed that, in general, despite the work undertaken, the discrimination indices for the reading and writing questions were lower than in previous papers and lower than the 2013 papers which were the subject of our analysis.

## **Stakeholder feedback**

Stakeholders have told us, and publicly reported, that they were disappointed with the summer 2015 awarding outcomes. We received a number of anecdotal reports that the rank-ordering of the most able students had not improved significantly. There was still concern that the most able students did not get the highest grade.

This suggests that the changes exam boards made in 2015 in response to our recommendations did not fully address concerns with the qualifications.

# **3. The 2016 summer assessments**

## **Review of assessments for 2016**

As a result of the exam boards' analyses, and continuing stakeholder concerns, we reviewed the 2016 papers and mark schemes for AQA, OCR and WJEC against our requirements (the General Conditions of Recognition) before the exams were taken. We excluded Pearson from this exercise as our 2014 research did not find any issues with the level of demand of their papers.

Our review took place following the production of live papers. This meant there was limited time available for the exam boards to make any changes to their papers before the examinations started. Instead, we identified some scope for the exam boards to adjust their mark schemes to provide for better discrimination and we asked the exam boards to consider how they could incorporate these. We did not suggest they should change their exam papers as this would introduce additional risks to the delivery and marking of the assessments.

The exam boards committed to using the findings of our review of the 2016 papers as they trained their examiners and as they considered the need for additional guidance for examiners that would support effective differentiation.

## **4. Next steps**

### **Awarding for 2016**

In June we published the regulatory requirements<sup>10</sup> which support our monitoring of the summer 2016 awards. For A level French, German and Spanish, exam boards must base their predictions on data showing the relationship between GCSE performance in 2013 and A level outcomes in summer 2015, rather than continuing to use the 2010/2011 baseline. This will avoid unintentionally undoing the effect of any changes to the 2015 and 2016 assessments.

We will continue to apply a tolerance of 0 to +2 percentage points at the A\* grade. We expect that outcomes at A\* will not be lower than the predictions, unless there is compelling evidence to support that.

In setting these requirements our regulations should complement any work the exam boards may do to improve the way in which their qualifications discriminate across the ability range.

### **Consideration for reformed qualifications**

Our 2014 research did not solely focus on the changes exam boards could make to improve the assessment of modern foreign languages. It also made a specific recommendation to inform the wider reform of these qualifications.

The current assessment objectives<sup>11</sup> require that:

- 30-35% of marks test students can *Understand and respond, in speech and writing, to spoken language*
- 40-45% of marks test students can *Understand and respond, in speech and writing, to written language*
- 25% of marks test students can *Show knowledge of and apply accurately the grammar and syntax prescribed in the specification*

However, the 2014 research found that knowledge and understanding of wider, cultural aspects of the country where the language is spoken was also being rewarded in the current assessments. This is not a stated assessment objective of

---

<sup>10</sup> See for example: <https://www.gov.uk/government/publications/data-exchange-procedures-for-a-level-gcse-level-1-and-2-certificates>

<sup>11</sup> The intended balance of knowledge, skills and understanding to be assessed

the current qualifications and the approach taken to assessing cultural aspects was not being done so consistently.

Our research recommended the development of new assessment objectives in this subject to improve the consistency and structure of the new qualifications:

*The absence of cultural aspects of knowledge and understanding from the assessment objectives should be considered in the criteria for the reformed specifications as part of the on-going consultation process.*

We developed assessment objectives<sup>12</sup> for reformed modern foreign language A levels to reflect the requirements of the revised subject content.<sup>13</sup> The content was developed by an independent body, the A level Content Advisory Board (ALCAB) on behalf of the Department for Education (DfE).

The subject content for reformed modern foreign language A levels explicitly requires students to consider cultural and social themes (past and present) in the countries where the language of study is spoken. As a result, the fourth assessment objective (AO4) states:

*Show **knowledge and understanding of**, and respond critically and analytically to, **different aspects of the culture and society of countries/communities where the language is spoken.***

AO4 is worth 20% of the marks for the new qualifications. To ensure the awarding objectives are applied consistently and accurately by the exam boards, we have also published guidance<sup>14</sup> which explains how we expect exam boards to interpret assessment objectives. This guidance sets out the key areas of emphasis in each assessment objective and the particular meaning for the subject of any key terms and phrases used; defined terms are shown in bold text, followed by their definitions.

The guidance for AO4 states:

**Respond critically and analytically** may include, but is not limited to:

- *selecting relevant material,*
- *presenting and justifying points of view,*
- *developing arguments,*
- *drawing conclusions based on understanding, and*

---

<sup>12</sup> <https://www.gov.uk/government/publications/gce-subject-level-conditions-and-requirements-for-modern-foreign-languages>

<sup>13</sup> <https://www.gov.uk/government/publications/gce-as-and-a-level-modern-foreign-languages>

<sup>14</sup> <https://www.gov.uk/government/publications/gce-subject-level-guidance-for-modern-foreign-languages>

- *evaluating issues, themes, and cultural and social contexts.*

**Countries/communities** means at least one country or community. However, specifications should include reference to a range of countries and communities where the language is spoken.

*In the context of this assessment objective, knowledge and understanding should be demonstrated through the Learners' ability to communicate information about, and to demonstrate their appreciation of, different aspects of culture and society.*

The new modern foreign language A levels, developed by the exam boards, have all been subject to accreditation, so we could check they met our new requirements.

We also gave copies of the 2014 research to members of the accreditation panel, as part of their initial training, to alert them to the issues identified in previous assessments.

The accreditation panel specifically considered the level of demand of the questions in the sample assessments and whether the mark schemes were likely to differentiate effectively. They also considered the assessment strategies<sup>15</sup> that accompanied each submission.

Reformed AS and A level French, German and Spanish qualifications will be taught from September 2016. The first AS assessments will take place in 2017, and the first reformed A level assessments will be taken in 2018.

## **Native speaker research**

Modern foreign language qualifications are intended to assess the foreign language skills of students whose first language is English. One of the concerns of stakeholders is that the performance of native speakers of the language being assessed might distort the grading and as a result disadvantage the non-native speakers. There is a view that the proportion of students who are native speakers is increasing. There is currently no routine collection of information about native speakers taking modern foreign language A levels and so we have no evidence to support or refute these concerns.

To help address this, we asked students entered for A level French, German, Spanish, Italian or Russian in 2016 to complete a questionnaire to give us more information about their language acquisition outside school/college. We will compare

---

<sup>15</sup> Assessment strategies are developed by each exam board to demonstrate how they will, for each qualification, design, set and deliver high-quality, effective assessments on an ongoing basis.

this information with the results for those students, to consider how students with different native speaker characteristics performed in the assessments.

The information collected from students will not be shared with the exam boards and will not affect students' results. However, it will provide us with a greater understanding of the prevalence of native speakers and the grades they are awarded. We will then consider whether any techniques might address any issues we may find. For example, if we found there was a significant proportion of candidates in one or more of the languages whose results were better than their overall prior attainment at GCSE suggested they would be, we could consider, in discussion with the exam boards and other stakeholders, whether there would be advantages to collecting additional data on A level MFL students so that they could be excluded from the predictions the exam boards use.

We will report on this work in autumn 2016.

## **5. Conclusion**

We continue to take the issue of grading in A level modern foreign languages seriously and recognise stakeholders' concerns. We have taken these into account in considering how we currently regulate A level modern foreign languages and to inform our reform of these qualifications. The activities we have set out in this report should not be taken in isolation of wider and more general work we do to regulate A level and other qualifications.

Exam boards will continue to use predictions, as is normal practice, when awarding A level modern foreign languages for this year. As in other subjects, exam boards can come to us with evidence to support an award that is significantly above (or below) the predictions, although we recognise there are challenges in producing such evidence.

We have reiterated our expectations that the questions and mark schemes should allow for appropriate differentiation of the highest performing students.

We will evaluate the impact of the exam boards' changes to their assessments for 2016 and consider whether additional changes might be required for 2017. The current qualifications will be awarded for the last time in 2017, other than for students re-sitting them in 2018.

We have also considered our research findings as part of the reform programme and during the accreditation of the reformed modern foreign language A levels that will be taught from September 2016. We will evaluate the effectiveness of those changes through our wider monitoring of exam boards' implementation of the reforms.

Our additional native speaker research should tell us later this year whether the take-up of the qualifications by native speakers of the languages distorts the overall outcomes, as some stakeholders believe.

We will continue to keep A level modern foreign languages under review to make sure that the most able students achieve the grades they deserve.

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 2016

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit <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

Telephone 0300 303 3344  
Textphone 0300 303 3345  
Helpline 0300 303 3346