

# Moderation of internal assessments

---

An explanation for centres of how adjustments to marks are made

---

## Introduction

When results are sent to centres, details of candidates' final marks for internally assessed units/components are included. By comparing these marks with the original centre marks, centres can see whether their marking in a particular unit/component was accepted without any change or whether adjustments were made. Further details about the reasons for any adjustments are provided on the feedback form.

When an adjustment has been made, centres may wish to know how the final marks were determined. This document explains the procedure which is used.

## Why is moderation necessary?

Ofqual's *General Conditions of Recognition* require moderation of assessments marked by centres to be carried out.

Awarding marks entails human judgement so, when a moderator looks at a centre's marking, it is unlikely that there will be precise agreement. Two people assessing the same piece of work may disagree on whether enough of a skill has been demonstrated or whether an answer is clear. Even one person may mark a piece of work, then re-mark it later and make slightly different judgements without being 'wrong' either time.

It is possible to reduce differences with practice, experience and discussion, but they will always exist. A small **tolerance** is therefore allowed on internal assessment. If the differences between a moderator's marking and a centre's marking are within this tolerance, the differences can be taken as a legitimate variation in judgement and the centre's marks can be accepted. If the differences are outside the tolerance, an adjustment will be needed to align the centre's standard of marking with the agreed AQA standard.

## How is moderation carried out?

In most AQA internally assessed units/components, a sample of work from each centre is provided for a moderator (either by post or when the moderator visits). The moderator starts by considering part of this sample (called the **sub-sample**).

If the moderator agrees with the marks in this sub-sample (to within the specified tolerance), then no change is made to the centre's marking.

---

If the centre's marks for any of the work in the sub-sample are outside tolerance, the moderator considers further work from the centre's sample, and any necessary adjustment to the centre's marks is made using the regression technique described below.

Precisely how much work is inspected by the moderator at this stage depends on the pattern of differences between the centre's marks and the moderator's marks.

If there is a consistent pattern (even where the differences are large), a fair adjustment can be made from inspecting the work of a relatively small number of candidates.

If the pattern is less consistent, the whole sample needs to be inspected. In exceptional circumstances, an adjustment which is fair to all candidates cannot be found from the sample, and at that point the moderator will request further work (often the work of all candidates) from the centre.

The marking of the sample (or sub-sample) is **not** a re-assessment of the individual candidates involved. It would be unfair to amend the marks of these candidates on a different basis from the marks of the other candidates at the centre. Instead, the sample is intended to be **representative** of the marking standard at the centre, in order to provide an indication of whether an adjustment is needed and to determine the nature and scale of that adjustment.

Therefore, there is a requirement that the centre's marking is internally standardised. For further information on internal standardisation see the AQA website ([aqa.org.uk/coursework](http://aqa.org.uk/coursework)) and the JCQ booklets *Instructions for conducting coursework*, *Instructions for conducting controlled assessments* and *Instructions for conducting non-examination assessments* on the JCQ website.

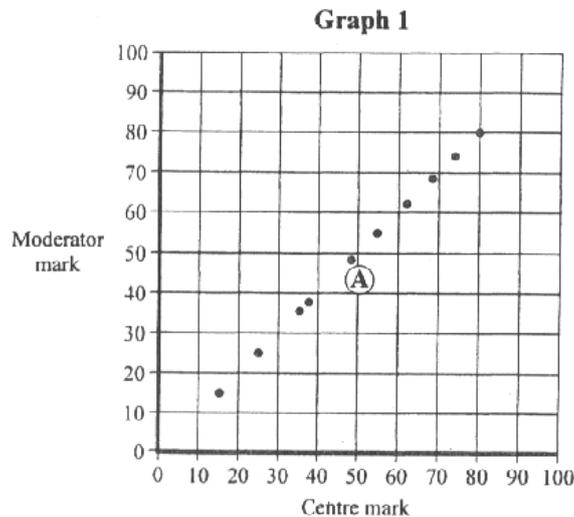
Where centres are in a consortium (ie several centres are working together in a subject with joint assessment arrangements), AQA must be informed. AQA will then take account of the consortium arrangement when carrying out moderation, for example by applying the same set of adjustments to all the centres if the marking is out of tolerance.

For further information see the AQA website and the application form for centre consortium arrangements, available on the JCQ or AQA website.

### **How is regression used to make adjustments?**

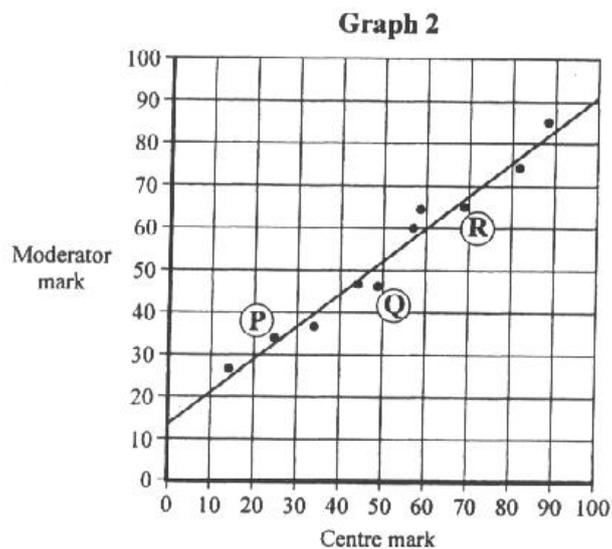
Regression compares two sets of data (in this case centre marks and moderator marks) in order to find the relationship between them. To understand how it works, think about what happens if a graph is plotted to show centre marks and moderator marks for the candidates in a sample.

If the moderator re-marked a sample of ten candidates and agreed with every one of the centre marks, the result would be as shown in Graph 1. For example, candidate A was given 49 marks by both the centre and the moderator.



In reality moderation produces something more like Graph 2: candidate P was given 24 marks by the centre and 34 by the moderator, candidate Q was given 49 marks by the centre and 46 by the moderator, and candidate R was given 69 marks by the centre and 65 by the moderator.

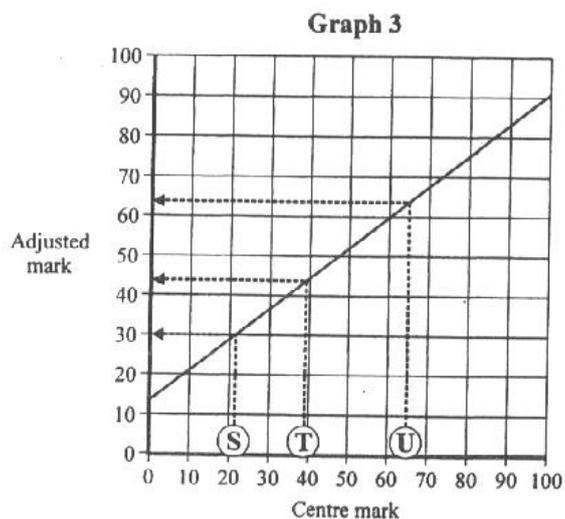
The points do not lie precisely on a straight line but there is obviously a trend. A line has been drawn to show this trend. This line is known as the regression line.



AQA uses the regression line in order to adjust the marks of candidates in a centre. Graph 3 shows the line from the example given above. This can now be used to read off the adjusted mark for every candidate at the centre.

For each candidate, the line gives an estimate of the mark that the moderator would have given. The estimate is based on the sample of work that was actually inspected by the moderator.

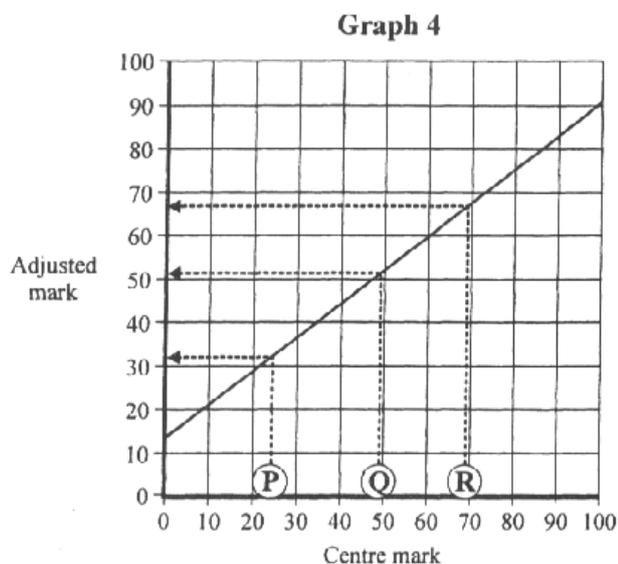
Three examples are shown in Graph 3. The centre gave Candidate S a mark of 21, Candidate T a mark of 39 and Candidate U a mark of 65. These marks are adjusted, using the regression line, to 30, 44 and 64 respectively, as shown by the broken lines.



To preserve the centre's rank order of marks and to ensure that candidates whose work was sampled are treated no differently from other candidates at the centre, the marks of **all** candidates are adjusted using the regression line, whether or not their work was seen by the moderator.

The procedure – which is used to find the best match of the centre's marks to the general standard – does not guarantee that candidates in the sample will get the mark which the moderator gave to their work.

Graph 4 shows what happens to the marks of Candidates P, Q and R from Graph 2. For example, Candidate Q has a centre mark of 49 and a moderator mark of 46 (as shown in Graph 2), but this candidate's final (adjusted) mark is 52, as shown in Graph 4.



Candidate	Centre mark	Moderator mark	Adjusted mark
P	24	34	32
Q	49	46	52
R	69	65	67

---

## Feedback to centres

AQA sends a report to each centre which covers matters such as the appropriateness of the tasks (where they are set by the centre), the accuracy of the assessment and the centre's administration.

## Enquiries about results

In line with other awarding bodies AQA offers a post-results review of moderation. This is intended to check that the assessment criteria were fairly, reliably and consistently applied in the original moderation process. For details please see the JCQ *Post-Results Services* booklet, available on the JCQ website, and the Post-results Services page of the AQA website ([aqa.org.uk/prs](http://aqa.org.uk/prs)).

## Application of the system to all centres

For each centre where the sub-sample is out of tolerance, provisional adjustments are made as illustrated in Graph 4. If the adjusted marks are only marginally different from the centre marks (in other words, if they fall within tolerance), the centre marks are accepted unchanged. Otherwise, the adjustments suggested by the computer are normally applied but in some cases they are reviewed manually and overruled.

## Summary

In summary moderation has three possible outcomes.

- (i) Where there are only small differences between the centre's marks and the moderator's marks for the candidates in the (sub-)sample, **the centre's marks are accepted for all candidates.**
- (ii) Where the differences are greater but the moderator generally agrees with the centre's rank ordering of the candidates, **the regression technique is used to adjust the marks of all candidates.**
- (iii) In exceptional cases, where the moderator disagrees significantly with the centre's rank ordering, either
  - the moderator re-marks further work (possibly the work of all candidates) before the final marks are determined
  - or the centre is asked to review its marking and then a new sample is selected for moderation.

Marks are not altered unless necessary, and then only in a manner which treats all candidates and centres equitably.