

International approaches to the moderation of non-examination assessments in secondary education



March 2018

Ofqual/18/6364

Author

This report was written by Benjamin M. P. Cuff, from Ofqual's Strategy, Risk and Research directorate.

Contents

Executive summary	3
1 Introduction	5
2 Method	6
3 Findings	9
3.1 Overview of qualifications, terms used for NEA, and types of moderation	9
3.2 Purpose of non-examination assessment.....	10
3.3 Purpose of moderation	11
3.4 Sampling of pupils and centres	12
3.5 Process of moderation.....	13
3.6 Appealing moderation outcomes	14
4 Discussion.....	15
4.1 Is there a need for NEA, and if so, a need for moderation?.....	16
4.2 What are the alternatives to the current system of moderation?	17
4.3 What are the alternatives to the current system of sampling?	19
4.4 What are the alternatives to the current system of appeals?	20
5 Conclusions	21
6 Tables	23
References	42

Tables

Table 1. Overview of qualifications, terms used for NEA, and types of moderation..	23
Table 2. Purpose of non-examination assessment.....	26
Table 3. Purpose of moderation	29
Table 4. Sampling of pupils and centres.....	32
Table 5. Process of moderation.....	35
Table 6. Appealing moderation outcomes	40

Executive summary

Non-examination assessment (NEA), such as coursework, portfolios, the demonstration of practical skills, or performances, can offer a valid assessment of skills not easily measured by external examination. Despite their importance, NEA outcomes may be less reliable than those for exams, and so systems of moderation are often put in place to help ensure consistency of marking. The purpose of this report is to review how moderation is currently being delivered in a number of jurisdictions across the world at upper secondary, and in particular to consider what might be learnt in terms of the approach to moderation that is currently taken in upper secondary assessments in England.

To meet this purpose, jurisdictions were identified that have a population of greater than 1 million, and either have English as a primary language, or use English for official documentation. The final list comprised of 23 jurisdictions. For each, literature was sourced in relation to their main senior secondary school leaving qualification (ie equivalent to an English A level) to cover a number of pre-determined topics of interest: the stated purpose of NEA and moderation, how pupils and schools are sampled for moderation, how moderation is carried out, and what processes exist for how to appeal outcomes of moderation. All qualifications were high-stakes in the sense that they are the main secondary school qualification, which will be used for entry into higher education.

The findings of this review were that a number of different approaches to moderation are currently taken across the globe. These include consensus moderation (where teachers meet to discuss their marking, and agree upon the assessment standards), verification (where moderators evaluate centres' assessments, but centres have the responsibility for making corrective action), moderation by inspection (where moderators evaluate centres' marking, and corrective action is made based on their decisions), and statistical moderation (where centres' NEA marks are scaled to match the mean and spread of their exam marks).

Most jurisdictions include NEA as a useful method of assessing skills not easily measured via exams. Some jurisdictions also promote the formative benefits of NEA on teaching and learning (eg in improving assessment practices, and in preparing students for their final exams). Most jurisdictions use some form of moderation in order to monitor the consistency (reliability) of marking, and several also note the importance of checking that marking is being done in line with the intended assessment standards (ie valid, as well as reliable). Jurisdictions using social forms of moderation (ie consensus, verification, and moderation by inspection) all moderate samples of students' work, but several variations were noted in the exact sampling strategy used (eg whether samples should include the marking of all teachers within each centre). Some countries conduct additional audits to check marking not covered in the moderation samples. Information on appeals was limited for most jurisdictions,

but some differences were observed in terms of who can lodge an appeal (individuals or centres), whether the original centre marks or moderated marks are the focus of appeals, and whether a fee is levied on appealing outcomes.

Overall, while it was found that England does not appear to 'stand out' from the other jurisdictions (ie only a few aspects of our approach would be considered 'unusual' within the international landscape), various points for consideration can be raised in terms of what we might learn from these international approaches. The purpose of these points is not to provide recommendations as to whether the current system in England should or should not be changed. Rather, they are intended to provide a number of points of comparison for policy makers to bear in mind.

First, while many jurisdictions include NEA in upper secondary school qualifications, there are differences in the type of moderation model used (eg moderation by inspection vs. statistical moderation). Second it is also worth considering whether different methods of moderation might be used for different subjects. For example, in Hong Kong most moderation is statistical moderation, but expert judgements are used in cases where alignment is not necessarily expected between outcomes for NEA and exams (expected alignment is a requirement of statistical moderation). Third, the different sampling strategies adopted elsewhere are also worth reflecting upon, and whether there are any lessons to learn about an optimal approach. For example, other systems target the marking of different teachers. The possibilities of increasing the sample size, or sampling at county level might also be considered. Fourth, consideration might also be paid to the possibility of conducting random audits to check the marking of work not included in moderation samples, which is an approach taken by some jurisdictions. Fifth, it is worth reflecting on whether the more formative benefits of assessment and moderation might be given greater precedence in England, as this is an idea more strongly promoted by some other jurisdictions, but this would need to be considered in the context of the purpose and uses of these qualifications.

1 Introduction

Non-examination assessment (NEA) can be defined as “any type of assessment that is not externally set and taken by candidates at the same time under controlled conditions” (JCQ, 2017a, p. 3). They generally include assessment decisions that are made by teachers within centres (eg schools), which are then submitted to an awarding body. These might include tasks that are set internally by the centre or externally by the exam board. NEA offers an alternative to externally assessed exams, allowing one to assess a different set of skills via a range of methods such as coursework, portfolios, practical skills, or performance pieces. Many countries around the world use a combination of external exams and NEA for their main senior secondary school leaving qualifications (see for example McCurry, 2013).

Despite their common usage, NEA outcomes are often considered to be more valid, but less reliable indicators of student ability compared to exams (Klenowski & Wyatt-Smith, 2014). This is because the reliability of different teachers’ judgements depends upon the consistency of those teachers’ interpretation and application of the assessment standards. While teachers might be reasonably expected to be able to determine the relative performance of each student within their class, determining the performance of their class relative to those in other centres is much more difficult (Elley and Livingstone, 1972, cited in Wilmut & Tuson, 2005). As such, teachers’ marks need to be quality assured via a process of moderation, so that those using the outcomes of these assessments (eg for accountability purposes, or by employers or university admissions officers) can have greater confidence in their validity (the importance of moderation for maintaining public confidence in assessments has been noted elsewhere – eg Stanley, MacCann, Gardner, Reynolds, & Wild, 2009). Moderation is particularly needed for high-stakes assessments, such as senior secondary school qualifications, where teachers have greater incentives to maximise their outcomes (outcomes on these qualifications often form the basis for accountability measures/performance scores for teachers/centres).

Various alternative approaches to moderation can be taken, several of which have been reviewed by Daly et al. (2011):

- **Statistical moderation** is a process of adjusting a centre’s NEA marks to (most commonly) match the level and spread of scores in that centre’s exam marks (eg an exam part of the same qualification). The rank order of NEA marks is not changed via this process. Various statistical moderation formulas are given by Williamson (2016), and the types of tests that can be used as calibration instruments have been discussed by Wilmut and Tuson (2005).
- **Moderation by inspection** is a process where subject experts are employed to review a sample of each centre’s marking. This is most commonly done remotely (ie centres post materials to their moderators), but moderators may

visit in person when this is not possible (eg to observe practical performances). Adjustments are made to the centre's NEA marks based upon the decisions made by the moderator(s).

- **Verification** is similar to moderation by inspection, although centres' marks are not adjusted according to the moderator's decisions; rather, feedback is generally given to the centre in order for them to make the appropriate corrective action. Verification often also focuses on how assessments are being delivered (ie the focus is not just on the outcomes of those assessments).
- **Consensus (or social) moderation** is a process where teachers meet in groups to discuss their marking, in order to standardise their interpretation and application of the assessment standards. No external moderator employed by the awarding body is involved in reviewing these decisions.

Each of these different approaches have been adopted within various education systems across the world. The current upper secondary system in England employs a moderation by inspection approach. While each exam board in England may handle certain details differently, the overall approach that they take is generally the same. More detail on this system shall be given later on in the report, and has also be given elsewhere (Cuff, 2017; Gill, 2015).

Some concerns have been raised about the effectiveness of the current approach in England. In particular, Cuff (2017) highlighted several risks to the reliability and validity of judgements that are being made by moderators. In order to better understand the moderation arrangements in England, it would be useful to know how the English system sits within the international landscape. By comparing our system with others, we can explore whether anything can be learnt from how moderation is conducted elsewhere. This is the purpose of the current report. The report begins by reviewing how moderation is conducted within a number of jurisdictions across the globe, before discussing what considerations might be made with regards to the current system of upper secondary moderation taken in England.

2 Method

The first task was to identify jurisdictions for review. There was no provision for the translation of foreign languages, so sovereign and non-sovereign states¹ were identified where English is used for official documents. This included jurisdictions

¹ Sovereign states (countries) have control over their own affairs and territories. Non-sovereign states have some autonomy, but are controlled by an external power (eg Hong Kong, which is a 'special administrative region' of China). For simplicity, all such territories are referred to as 'jurisdictions' in this report.

where English is the primary language, or where English is not the primary language, but is an official language, and therefore used for official purposes. For example, the primary language spoken in India is Hindi, but English is used for many official documents. Both England and Scotland were included in the review for Great Britain, as they operate very different moderation practices. To make the review more manageable, the list was further reduced via the exclusion of any jurisdictions that had a population of less than 1 million (in 2016 – see The World Bank, 2017).

The final list comprised of 23 jurisdictions. In Australia, Canada, South Africa, and the United States of America, education and assessment practices differ across states/provinces. For these, the largest state/province by population was reviewed; this was to not make the review unmanageable, and to avoid repetition (different provinces within the same nation may share many common practices). Some jurisdictions subscribe to multi-national organisations for assessment purposes, such as the West African Examinations Council (WAEC) and the Caribbean Examinations Council (CXC). To avoid repetition, these multi-national organisations were reviewed instead of each member nation separately. Any use of the word 'jurisdictions' in this report should also be taken to include these organisations. The final list of jurisdictions was as follows:

- Australia (most populous state: New South Wales)
- Canada (most populous province: Ontario)
- The Caribbean Examinations Council (CXC – Jamaica and Trinidad and Tobago met the original inclusion criteria. Other member nations include Anguilla, Antigua and Barbuda, Barbados, Belize, British Virgin Islands, Cayman Islands, Dominica, Grenada, Guyana, Montserrat, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, and Turks and Caicos Islands)
- England
- Hong Kong
- India
- Ireland
- Kenya
- Namibia
- New Zealand
- Pakistan
- Philippines
- Rwanda
- Scotland
- Singapore
- South Africa (most populous province: Gauteng)
- Uganda
- United States of America (most populous state: California)

- The West African Examinations Council (WAEC – Ghana, Nigeria, and Sierra Leone met the original inclusion criteria. Other member nations include The Gambia and Liberia)
- Zimbabwe

For each jurisdiction, a popular online search engine was used to source publically available documentation on NEA and moderation for their main secondary school leaving qualification². Efforts were made in all cases to glean information from official sources (eg government or exam board/awarding body websites). However, this was not always possible, and so some media sources were also used. Literature was sought with a number of specific questions in mind. The first question was intended to glean a broad overview of each jurisdiction's education/assessment system, with the remainder exploring moderation of NEA in more depth:

1. What is the main senior secondary school leaving qualification (ie equivalent to an English A level), what term is used to refer to NEA, and if applicable, what is the main type of moderation used (from the list presented earlier)?

For those jurisdictions that have documented moderation systems in place,

2. What is the stated purpose of NEA?
3. What is the stated purpose of moderation?
4. How are pupils and schools sampled for moderation?
5. How is moderation carried out?
6. What is the process for appealing the outcomes of moderation?

Ultimately, these questions can be condensed down into the 4 main research questions for this project, regarding what can be drawn from international approaches in relation to the current system in England. This report does not intend to provide recommendations on how the current system should/should not be changed, but to present the alternative approaches that are taken elsewhere, alongside considerations about the current system in England.

1. Is there a need for NEA, and if so, a need for moderation?

² For England, while the focus here is on A levels, it is worth noting that the same process of moderation exists for GCSEs (taken at around age 16). A different method of moderation is adopted for Key Stage 2 assessments (taken at around age 11 – see Cuff, Howard, Mead, & Newton, 2018).

2. What are the alternatives to the current system of moderation?
3. What are the alternatives to the current system of sampling?
4. What are the alternatives to the current system of appeals?

It should be noted that the focus here was only to review national/state/provincial systems of moderation. Most jurisdictions promote internal (ie centre-based) moderation prior to external moderation being carried out. However, these practices are often dependent upon policies of individual centres, and so are not well documented. The exception to this rule is Ontario (Canada), where moderation is done internally within centres, however it does have a provincial strategy for consensus (internal) moderation.

It should also be noted at this stage that no guarantees can be made that the information presented within this report is an accurate reflection of practice; the review simply reflects how processes are described within the documentation that was found. Nevertheless, this is sufficient for the current purpose, which is not to document international practices per se, but rather to use examples of international documentation (ie intentions) as a means to consider our own system.

After sufficient information for each jurisdiction had been found, or at least an exhaustive search had been made, information was organised into a number of tables: 1 for each of the 6 search topics described previously. Full tables can be found towards the end of this report (Section 6), but findings are summarised below.

3 Findings

3.1 Overview of qualifications, terms used for NEA, and types of moderation

Table 1 presents information on the main senior secondary school qualification that each jurisdiction offers, and the typical terms that are used to describe NEA in each jurisdiction (eg 'continuous assessment', 'classroom assessment'). The particular use of these terms has little bearing on this review, but have simply been included to provide a starting point for those wishing to seek more information on each jurisdiction. Each jurisdiction's qualifications are 'high-stakes' in the sense that they are the main secondary school qualification, which will be used for entry into higher education. However, it is reasonable to assume that the uses to which each qualification is put will vary between jurisdictions.

All 4 forms of moderation presented previously (Section 1) are used by at least one of the jurisdictions included in this review. Two use statistical moderation, 1 uses consensus moderation, 3 use moderation by inspection (with 2 explicitly containing a statistical scaling element), and 2 use verification. Three use some form of expert judgements, but it is unclear whether these would be more appropriately described as moderation by inspection or verification (ie it is unclear from the documentation what actions are taken following moderation). Hong Kong uses statistical moderation or moderation by inspection (with statistical scaling), depending on the subject. For example, where the skills assessed by the NEA component of a course differ substantially to those assessed during the exam component (such as in design and technology), moderation by inspection is used in favour of statistical moderation (HKEAA, 2010, sec. 7.1). This is because it is deemed inappropriate to calibrate NEA marks with the exam marks in such cases.

India, Pakistan, and the United States of America (California) do not seem to employ NEAs for their qualifications, and so no moderation is undertaken. In addition, while Kenya, Philippines, Rwanda, Uganda do employ NEAs, they do not seem to have national moderation systems in place (or at least the lack of available documentation would seem to suggest so). The leaving certificate in Ireland does contain NEA, but this is externally set and marked in most cases. Because these 8 jurisdictions do not appear to carry out moderation, they are excluded from the remainder of Section 3, and Tables 2-6.

3.2 Purpose of non-examination assessment

Table 2, which contains statements on the purpose of NEA for each jurisdiction, has been included as moderation is only necessary to the extent that NEAs are necessary. A consideration of the purpose of NEA might also help to inform our thinking on moderation in later sections, such as how moderation might be used to help meet this purpose.

Various different rationales have been given by different jurisdictions, but the most common reason is (at least in the current sample) that NEA allows educators to assess skills that cannot easily be assessed via an external exam. For example, performance skills in the creative arts (music etc.) would be difficult to validly assess in a written test. Outcomes of NEAs are therefore believed to offer a more valid reflection of the abilities of students than exams in relation to certain types of skills.

Along the same lines of validity, Zimbabwe's rationale suggests that NEAs can offer an indication of students' abilities when working in different conditions (ie assisted versus unassisted work; group versus individual work). The WAEC further proposes that NEA outcomes are a truer reflection of students' abilities, because students are more comfortable taking assessments in their own classrooms, and do not have to

face the same pressures as they would in external exams (which might affect their normal performance).

Other rationales for the inclusion of NEA have also been put forward regarding the formative benefits that they can have. For example, Canada and Zimbabwe promote NEA as a way of improving student learning via assessment for learning, rather than assessment of learning (the latter would be more the case for summative exams). Singapore and South Africa also note that NEA provides a better indication of the journey taken to reach the end result, including levels of effort and progress made. Similarly, South Africa notes that NEA aids students in preparing for their final exams.

Part of Singapore's rationale includes the suggestion that NEAs are more manageable than nationally delivered exams, both in terms of costs and efficiency. Greater national organisation is needed for exams, and with greater cost incurred (ie to pay markers), whereas NEAs are generally delivered and marked within centres.

It is unfortunate that documentation was not found on why India, Pakistan, and the United States of America do not seem to include NEAs in their secondary school leaving qualifications. It is therefore unclear whether or not those jurisdictions disagree with other jurisdictions' rationales, or whether they agree but have instead prioritised some competing consideration(s). Nevertheless, findings suggest that the majority agree with the notion that NEA is an important method of assessment. Ontario (Canada) in particular seems to promote this idea, given that the Ontario Secondary School Diploma appears to be solely based upon NEA outcomes.

3.3 Purpose of moderation

Where NEA is implemented, most jurisdictions recognise the need for some form of national/state/provincial moderation strategy. Whilst it is unclear why Kenya, Philippines, Rwanda, Uganda do not seem to have done so, it is possible that this is due to a lack of resource – Kenya, Rwanda, and Uganda in particular have a relatively low GDP per capita (The World Bank, 2016).

Each jurisdiction's rationale for moderating NEA marks are presented in Table 3. As with the previous sub-section, various reasons are given for this. However, all seem to fall under 3 main themes here: the need for reliability, the need for validity, and the need to evaluate and improve the assessment practices of centres. As a side note, these broadly align with the purposes of moderation that have been described by Klenowski and Wyatt-Smith (2014, Chapter 5).

Most jurisdictions highlight the need to achieve consistency between different centres' assessment decisions, so that reliability of outcomes can be achieved across the jurisdiction. More specifically, there is the recognition that teachers within different centres often assign different tasks, and mark those tasks in different ways. Moderation is therefore needed to ensure that all students are graded according to

the same assessment standards, regardless of who marked their work. Some jurisdictions only emphasise the need to ensure consistency between different centres, whilst some also note that consistency should be ensured between each teacher marking within each centre. It is possible that the former group of jurisdictions assume that intra-centre consistency is achieved via centres' internal moderation processes, prior to them submitting their decisions for external moderation.

Some jurisdictions also emphasise the need to ensure that all teachers are consistently marking in line with the intended standards of the NEA. In other words, assessment decisions not only need to be reliable (ie consistent between different centres), but also need to be valid in relation to the assessment standards (ie the standards against which teachers judge their pupils should be the same as the standards outlined in the course/assessment specification). The WAEC notes that assuring the validity of assessments through moderation can help to maintain public confidence in qualifications and awarding organisations.

A few jurisdictions also promote moderation as a means to evaluate and enhance assessment practices within schools. This is particularly true of Canada, which notes benefits of moderation for teachers' professional development, for example by helping teachers to focus their teaching on specific learning objectives. Through their verification process, Scotland aim to confirm that assessments are being appropriately delivered in centres, and that assessments align with national guidelines. In other words, moderation is not only used to monitor outcomes, but also to monitor the design and delivery of assessments.

3.4 Sampling of pupils and centres

The various approaches to sampling are given in Table 4. Sampling largely depends upon the method of moderation that is employed by each jurisdiction. For example, those adopting a statistical moderation approach sample all marks submitted by all centres (although some outlier marks may be excluded from the calculation). Sampling all marks is easy to achieve in statistical moderation, because the process is largely automated. Hong Kong uses statistical moderation, but checks the outcomes of that process using subject experts – all pupils and centres are sampled for calibration, but samples are selected for the expert judgement element.

Those that employ subject experts to review centres' decisions all select a sample of scripts from individual centres, as moderating every script in this manner would prove unmanageable. Typically, these samples constitute a relatively small number of scripts from each centre (eg 5 or 6 scripts, or 10% of the entry). In some jurisdictions (eg England and Namibia), samples are reviewed in stages (ie a moderator will review a set of scripts, and then if they are not content with the marking, that moderator will review a further number of scripts).

Samples usually aim to cover the full range of marks. For example, the CXC samples scripts with the highest, median, and lowest mark for each centre, and those midway between the highest and median and between the lowest and median. In some jurisdictions (eg England and Hong Kong), samples are chosen by the exam board(s), whereas in others (eg Namibia and New Zealand), the centre is responsible for selecting their own sample. New Zealand and Scotland only moderate a selection of courses for each centre in each year. While most jurisdictions seem to focus efforts only at a national or state/provincial level, South Africa and Zimbabwe select samples at centre, district, and provincial level. The purpose of this is to evaluate consistency at each level of the system. Only Namibia explicitly requires examples of each teachers' marking to be sampled (the rest only seem to sample at a centre level).

In addition to 'usual' moderation, the CXC and New Zealand also samples a number of centres for random audit. For example, CXC moderators visit centres to evaluate marking that had not been included in the main moderation sample. Presumably, this is to ensure that all of the centre's marking is appropriate, not just for those students included in the sample.

3.5 Process of moderation

As one can see from Table 5, there are a range of approaches to moderation that are taken internationally. However, these can all be collapsed under 5 general approaches (ie 'no moderation' plus the 4 described in Section 1):

- Kenya, Philippines, Rwanda, and Uganda use NEA as a form of assessment in their qualifications, but do not seem to have a national strategy for moderating outcomes (or at least the lack of available documentation would seem to suggest so).
- Canada promotes a system of consensus moderation, where teachers within schools are encouraged to meet and discuss pupils' work to arrive at a consensus about the assessment standards. Outcomes do not appear to be moderated at a provincial level, but there is an expectation that all teachers mark according to the provincial standard.
- New Zealand and Scotland operate a verification model. These jurisdictions dispatch verifiers (moderators) to centres to check that teachers are appropriately marking in relation to the standards. Where irregularities are found, centres are responsible for making corrective action.
- The CXC, England, and South Africa use moderation by inspection. Moderators remark a sample of scripts. For the CXC and England, all marks for each centre (ie not just those included in the moderation sample) are then scaled based

upon the relationship between the centre's and moderator's marks. This is done statistically in England, using a regression line; the method is unclear for the CXC. For South Africa, moderators' recommendations are passed on to 'subject advisors', who then adjust the centre's marks (how exactly these adjustments are made is unclear).

- Australia, Hong Kong and the WAEC use statistical moderation. Broadly, NEA marks are scaled so that the NEA mean equals the exam mean, and the spread of NEA marks equates to the spread of exam marks. Outliers are sometimes excluded from calculations of the mean/spread for a centre, if their inclusion would be deemed to have an unfair impact on overall outcomes. The rank order of a centre's marks are not changed in statistical moderation.

Hong Kong employs subject experts to review the outcomes of statistical moderation, to check for appropriateness (eg for some centres, the distribution of NEA marks may be legitimately different from their exam marks). For subjects where exam marks are not expected to align with NEA marks (eg design and technology – HKEAA, 2010, sec. 7.1), Hong Kong instead employs expert judges to remark a sample of scripts, and then NEA marks are scaled with those marks, rather than against exam marks. This would more appropriately fall under the category of moderation by inspection.

Namibia, Singapore, and Zimbabwe make use of expert judgements. However, it is unclear whether these jurisdictions operate a model of verification or inspection because details on the outcomes of moderation could not be found (eg which party has responsibility for making corrective action).

3.6 Appealing moderation outcomes

Table 6 outlines the various systems that are put in place for students/centres to appeal the outcomes of moderation. Documentation was notably scant here, and no information could be found for several jurisdictions. It is difficult to know in these cases whether no systems of appealing moderation outcomes are available, or whether they are just not documented publically. Other jurisdictions clearly had systems in place to appeal the outcomes of exams, but it was often unclear whether these systems could also be used to lodge appeals against moderation.

For those that had documented systems in place, there appears to be some variation in terms of who can actually lodge an appeal to the relevant body. Some jurisdictions (Australia, Hong Kong, Namibia, and South Africa) seem to allow individual students to appeal directly, whereas others (England, New Zealand, Singapore, Scotland, and the WAEC) seem to only permit centres to lodge an appeal (ie candidates would have to appeal via their centre).

There also seems to be variation in terms of which decisions are the object of these appeals. For example, candidates in Australia seem to be appealing against the original decisions made by their teachers (centres are required to make corrective action after a review of moderation). In contrast, appeals to exam boards in England relate to the outcomes of moderation, meaning that appeals are made for the centre as a whole, but individual marks cannot be challenged. It is somewhat unclear what actually occurs during some jurisdictions' appeals. For example, Hong Kong states that appeals are to be used for 're-checking' marks, but not 're-assessment' of performance. This perhaps suggests that appeals are to be used to check for administrative errors only.

Further variations also exist as to whether fees must be paid as part of an appeal. While most countries do not seem to require fees to be paid, England exam boards charge a fee if the centre's original marks are not reinstated (ie the moderator's decisions were upheld), and Singapore always charges a fee to make an appeal.

England appears to be unique in the fact that exam boards operates an 'automatic protection' rule for reviews of moderation (at least this did not appear to be mentioned within any of documents reviewed for other jurisdictions). This rule means that individual students' grades cannot be reduced as a result of a review of moderation. Because all of a centre's marks would be changed following an appeal (individual marks cannot be challenged), this rule is in place to protect individuals who did not wish to lodge the appeal.

4 Discussion

This section discusses what might be learnt from these findings in relation to the system that is currently delivered in England. As noted earlier, it does not intend to provide recommendations. Rather, the purpose of this section is to highlight a number of areas that policy makers might consider when reflecting on our current position on moderation.

The first overall conclusion that can be drawn is that in many respects, England does not appear to 'stand out' from the other jurisdictions reviewed here. While a number of different approaches to moderation are taken across the world, England is not alone in operating a moderation by inspection model. The stated purpose of NEA in England is "[to assess] skills and knowledge [that] cannot be assessed through exams" (Ofqual, 2017b), which aligns with the purpose proposed by many other jurisdictions. The same can be said for the stated purpose of the moderation of NEAs in England, which aligns with the purpose given by most other jurisdictions. The fact that a number of scripts from across the range of abilities are sampled for each centre is also typical practice for jurisdictions practising moderation by inspection.

While most of the more specific details are shared between England and at least 1 other jurisdiction at upper secondary, 3 details do not seem to be shared elsewhere (at least not in the publically available documentation). Firstly, England appears to be the only system which applies a tolerance to moderators' judgements. Here, a centre's marks are only changed where moderators disagree beyond a certain number of marks (small differences in opinion are often considered to be reasonable/justifiable, given that standards for NEAs are not precisely defined – see Maxwell, 2002). This is not mentioned in documentation of the CXC and South Africa, which operate similar models of moderation by inspection. This does not necessarily mean they do not do so, however, and it is also possible that a 'natural' tolerance is of course being applied when moderators are making their decisions (ie the decision to change marks might only be made when disagreements are considered noteworthy). Secondly, England appears to be the only jurisdiction that offers 'automatic protection' of marks during appeals. This means that marks cannot be reduced as a result of an appeal (known as a 'review of moderation'). Thirdly, England is the only jurisdiction in our sample that appears to charge a fee based upon the outcome of an appeal. Most others either do not charge a fee or apply a fee regardless of the outcome.

4.1 Is there a need for NEA, and if so, a need for moderation?

Returning to the research questions presented earlier, one might first consider whether NEAs and moderation are needed within national qualifications.

While there may be some concerns with the reliability of NEA outcomes (for England, see Cuff, 2017), it would seem that the majority of the international community does regard NEA as a worthwhile endeavour. They are seen by most jurisdictions to offer a more valid form of assessment for those skills not easily assessable via external exams, and are also seen by some jurisdictions to offer various formative benefits for teachers and pupils that are perhaps not available through exams. It would seem sensible, therefore, to retain NEA where the validity of qualifications might otherwise suffer, assuming that moderation can secure sufficiently reliable and valid outcomes.

Where NEA is implemented, most jurisdictions have adopted some form of national/state/provincial moderation strategy. Most jurisdictions highlight the need for moderation to ensure that marking is being done consistently between centres. Some jurisdictions, including England, also note in their rationale that consistency between teachers should be achieved. In its rationale, England also emphasises the need for moderation to ensure the 'accuracy' of outcomes (ie validity in relation to the assessment standards). While not all jurisdictions explicitly include this idea in their rationales, it is perhaps implicit. As the WAEC notes in their rationale, ensuring that outcomes are valid helps to maintain public confidence in qualifications. Targeting both reliability and validity through moderation certainly seems like a wise approach.

One idea that England does not seem to hold in great importance (at least explicitly) is the use of NEA and moderation for formative purposes. Ontario (Canada) in particular seems to be a big supporter of this, emphasising assessment for learning and as learning, not just assessment of learning, in its assessment strategy (Ontario Ministry of Education, 2010). Several jurisdictions suggest that NEA can improve student learning, including helping them to prepare for their exams. Ontario also promotes moderation as a means to help teachers focus their teaching. Those operating a verification model (New Zealand, and Scotland) also use moderation to improve assessment practices within centres. Klenowski and Wyatt-Smith (2014, p. 74) promoted these formative benefits, describing moderation as being “integral to the entire process of effective teaching and learning”. A consideration for England may be whether these more formative aspects of assessment and moderation might be given greater emphasis in assessment and moderation moving forward.

4.2 What are the alternatives to the current system of moderation?

Perhaps the most obvious distinction to make across the different jurisdictions is of the different types of moderation that can be delivered (ie consensus moderation, verification, moderation by inspection, and statistical moderation). Linn (1993) described these different approaches as lying on a continuum of ‘rigour’, with consensus moderation being the least rigorous, followed by verification and then moderation by inspection. Statistical moderation can be considered to be the most rigorous because one can easily sample all marks, and because it does not rely upon the judgements made by individuals, there are no concerns of human error or bias. Concerns of inconsistency and bias have been raised in relation to the current moderation by inspection approach taken in England (Cuff, 2017), and other research has reported “measurable impacts” of moderator inconsistencies on students’ grades (Taylor, 1992, cited in Johnson, 2011, p. 43). It is therefore worth reflecting on the other methods that are available.

An important consideration when reflecting upon these methods is how NEA outcomes might be used within a jurisdiction. Each qualification reviewed here is high-stakes in the sense that they are all used for entry into higher education and employment. Nevertheless, jurisdictions may still vary in what they would deem to be an acceptable level of rigour for moderation, in order to be considered sufficient for their particular usage of outcomes. While reflecting on the appropriateness of these alternative approaches for the context in England, it is therefore worth bearing in mind how sufficient each approach may or may not be in relation to how outcomes are being used. As A level outcomes are currently used for many high-stakes decisions in England (e.g., for university and employment decisions, and for school-accountability purposes), one of the more ‘rigorous’ forms of moderation would perhaps seem most appropriate.

Given the above, therefore, one might assume that statistical moderation would be the approach for jurisdictions to aspire to, given that it offers the most 'rigour'. However, achieving such rigour is highly dependent upon there being suitable calibration instrument available (eg an exam). The outcomes of statistical moderation will only give a valid indication of student ability in relation to the NEA standards if the calibrating instrument measures essentially the same content or construct, and at the same level of cognitive demands (Linn, 1993). Too little overlap between the NEA and the exam renders the exam unsuitable as a calibration instrument, but too great an overlap creates redundancy in having 2 assessments (Smith, 1978, cited in Wilmut & Tuson, 2005). As one might imagine, it is often difficult to find an instrument that can meet these conditions. The challenge is greater when considering the fact that the main purpose of NEA is to target skills not easily assessable via exam, meaning that it is unlikely that an external exam would have sufficient overlap with the NEA. Statistical moderation would therefore seem to be at odds with one of the main purposes of moderation in England: to ensure that outcomes are an accurate reflection of the NEA standards. In fact, Maxwell (2002) argued that statistical moderation isn't moderation at all, suggesting 'scaling' to be a more appropriate term, as this approach does not ensure that students' marks are aligned with the intended standards of the NEA, but rather the standards of the exam.

Further issues with statistical moderation have been discussed elsewhere (eg Williamson, 2016; Wilmut & Tuson, 2005). For example, there is a lack of opportunity for providing meaningful feedback to centres on their assessments. This would again seem to be at odds with the formative purposes of moderation promoted by several jurisdictions. The relationship between NEA scores and exam scores would also need to be the same for each centre for scaling to work. However, the strength of correlations between teachers' judgements and external tests can vary from teacher to teacher (Martínez, Stecher, & Borko, 2009). The exact method of statistical adjustments would also have to be carefully chosen, as different formulas can result in quite different outcomes for individuals (Williamson, 2016).

Another possible issue with statistical moderation is the potential washback into school behaviours. If the NEA scores are moderated to align with the examination scores, it makes more sense for schools to allocate more resource and teaching focus for the examination content and preparation to maximise scores for the examination so that the NEA scores might be 'pulled up'. This might undermine the very role and purpose of including NEA.

Given the issues with statistical moderation, Wilmut and Tuson (2005, p. 60) noted that they "find no compelling reason to opt for statistical moderation in place of the most immediate alternative of moderation by inspection". The current system in England of moderation by inspection seems to be reasonably well placed on Linn's (1993) continuum, offering greater control over outcomes than would consensus moderation or verification, but without compromising on validity in relation to the NEA standards, and should allow for more formative feedback to be given to centres than

just statistical moderation. Nevertheless, further reflection may be needed on the stated purpose of moderation in England: “[to determine whether] the criteria against which learners’ performance is differentiated are being applied accurately and consistently by assessors in different centres” (Ofqual, 2017a, para. H2.2). If accuracy in relation to the NEA standards is desired, then statistical moderation on its own is unlikely to satisfy this purpose, unless a suitably aligned calibration instrument can be found. Subject experts are better placed to make those kinds of judgments, and so the current moderation by inspection model seems more appropriate. However, subject experts are less able to evaluate the consistency of judgments being made between centres (eg each expert only moderates a small number of centres), and issues with reliability have been identified (Cuff, 2017).

Another alternative is offered by Hong Kong, where statistical moderation is used for most subjects, but subject experts are also employed to review samples of work for each centre. The purpose of this is to check whether the outcomes of statistical moderation are appropriate. This offers a contingency for those centres whose distribution of NEA marks legitimately differs from the distribution of their exam marks. The system in Hong Kong also addresses the other issue that suitable calibration instruments may be difficult to find in some cases. Where this proves to be the case in the Hong Kong system, inspection by moderation is instead used. Adopting different approaches depending on the particular qualities of each qualification might be something else for those in England to consider.

Although the purpose of this literature review was to seek to better understand other arrangements in place for moderation, an alternative way of marking NEA is for the examination body to organise the marking of all candidates work in a similar way to that of examinations, thus negating the need for teacher marking and the need for a check on teacher marks. Such a system would retain NEA, but dispense with teacher marking and the need to provide checks.

4.3 What are the alternatives to the current system of sampling?

The sizes of samples moderated by exam boards in England are broadly comparable with those elsewhere, and may even be somewhat larger when one considers that samples are expanded wherever issues are identified with a centre’s marking. Nevertheless, it is still worth reflecting upon whether the size of samples is sufficient in light of the purpose of moderation being to ensure that marking is both reliable and valid. Further research would be needed to answer this question, and so no claims are made for or against sufficiency here. On the one hand, increasing the standard sample size might offer greater assurance that all marking is reliable and valid, and may increase the ability of moderators to provide more meaningful feedback to centres (increasing the formative capabilities of moderation). On the other hand, increasing samples might place unnecessary burden on the system in terms of manageability and cost, if the current strategy is indeed sufficient.

One concern that could be levied against the current approach to moderation is that the assumption that the sample generalises to all the other work in the centre may not always be met. There are no guarantees that teachers' marking is consistent between pupils included in the moderation sample, and those that are not included. More specifically, there may be a concern that teachers' marks might be inflated for non-moderated work, to improve outcomes for accountability/performance table purposes. A system similar to that operated by the CXC might be considered to combat this. In addition to usual moderation, the CXC samples a number of centres for random audits in order to evaluate marking not included in the main moderation sample. Further burden would again be placed on manageability and cost, but this might be considered for England, as it may help to ensure consistency in all marking, not just marking sampled for moderation.

It is interesting to note that samples in England do not have to cover each teacher; rather, sampling is only at a centre level. This seems to be a common approach taken by most jurisdictions, as only Namibia appears to explicitly sample marking from each teacher. However, by not targeting all teachers, the current sampling strategy in England may go against its stated purpose of moderation: to ensure consistency, "regardless of the identity of the Assessor, Learner, or Centre" (Ofqual, 2017a, para. H2.2 – emphasis added). While it is likely assumed that intra-centre consistency (ie between teachers within a centre) is achieved via centres' internal moderation, no guarantees can be made under the current sampling approach that this is always being done appropriately. It may therefore be worthwhile considering whether to include all teachers in the sampling strategy, in order to check inter-teacher, as well as inter-centre, consistency.

In South Africa and Zimbabwe, moderators review samples of work at a centre, district, and provincial level. In other words, the consistency of marking is evaluated at each level of the system. Again, this could be something to reflect upon for England. Currently, all moderators attend standardisation sessions to encourage consistency, and their decisions are checked by team leaders and principal (head) moderators, but there are currently no formal mechanisms in place to check that standards are being consistently applied across different counties. However, it is likely that moderators' allocations reflect a regional variety when moderation is administered postally, and is not 'visiting moderation'. Any such activities for visiting moderation would need to be balanced against the burden to manageability and cost.

4.4 What are the alternatives to the current system of appeals?

International systems of appealing the outcomes of moderation are not well documented (at least not publically). Nevertheless, having a system of review or appeal in place is obviously worthwhile, particularly for those operating systems of moderation that rely upon judgements made by individuals (including England). Given that one of the main purposes of moderation in England is to ensure the

validity of outcomes, a system of appeal (which is known in England as a 'review of moderation') is needed to correct any mistakes made by moderators.

It might be worth reflecting on who can lodge an appeal, and to what end. Exam boards in England, as with some exam boards in other jurisdictions, only allow centres to lodge an appeal, and individuals would have to appeal through their centres³. Other jurisdictions, however, allow students to appeal to the awarding body directly. The reason why England requires individuals to appeal through their centres is because any post-review mark adjustments are applied to the whole centre, due to the scaling algorithm that is used to adjust centres' marks. This means that individual marks cannot be changed. Arguably, this causes issues of fairness for individuals, as it may be more difficult for them to appeal if their centre is unwilling, or unable to pay the associated fee if their appeal was not successful (England operates an outcomes-based fee system for appeals). Nevertheless, the 'automatic protection' rule does mean that when centres lodge an appeal, individual marks cannot be lowered – the absence of this rule might present issues of fairness for individual candidates who were happy with their original marks.

5 Conclusions

Through a consideration of how other jurisdictions deliver moderation, a number of reflections can be made about the current system of moderation in England. Firstly, policy-makers may wish to reflect upon whether the formative aspects of NEA and moderation might be given greater emphasis. Secondly, while it may appear that a moderation by inspection approach might be the best way of ensuring that outcomes reflect the intended NEA standards, alternative approaches to moderation might still be considered. This might include reflecting upon the possibility of operating slightly different models of moderation for different qualifications. Thirdly, consideration might be paid to the current sampling strategy, and whether changes might be beneficial with regards to the size of samples, what work is sampled (ie which marks across the range of achievement), whether all teachers should be sampled, and whether moderation should monitor outcomes across counties. These considerations of sampling in particular need to be balanced against considerations of manageability and cost.

Careful thought is always needed when looking at the outcomes of an international review. In weighing up the various pros and cons of alternative approaches one needs to consider their applicability and suitability for translation into another system, with its own context of qualification purpose and use. Nevertheless, this review

³ In England, individuals can appeal within the centre, prior to marks submission to awarding bodies, if they believe that their mark is not appropriate.

provides a point for reflection on moderation as an important quality assurance mechanism.

6 Tables

Table 1. Overview of qualifications, terms used for NEA, and types of moderation

Country/territory	Main senior secondary school leaving qualification	Typical term used for non-examination assessment (NEA)	Type of moderation
Australia (New South Wales)	Higher School Certificate (HSC)	School assessment	Statistical moderation
Canada (Ontario)	Ontario Secondary School Diploma (OSSD)	Assessment and evaluation (all assessments are school based)	Consensus moderation
Caribbean Examinations Council (CXC) ^a	Caribbean Advanced Proficiency Examinations (CAPE)	School based assessment	Moderation by inspection with statistical scaling
England	GCE A level	Coursework / non-examination assessments	Moderation by inspection with statistical scaling
Hong Kong	Hong Kong Diploma of Secondary Education (HKDSE)	School based assessment	Statistical moderation with checks by subject experts, or moderation by inspection with statistical scaling (depending on the subject)

*International approaches to moderation of non-examination
assessments in secondary education*

India	All India Senior School Certificate Examination (AISSCE)	Not applicable (No NEA in Class XII)	-
Ireland	Leaving certificate	Coursework	No moderation – coursework is externally set and marked.
Kenya	Kenya Certificate of Secondary Education (KCSE)	Continuous assessment	Appears to be no national moderation (no information found)
Namibia	Namibia Senior Secondary Certificate (NSSC)	Coursework	Inspection/verification (information unclear)
New Zealand	National Certificate of Educational Achievement (NCEA)	Internal assessment	Verification
Pakistan	Higher Secondary School Certificate (HSSC)	Not applicable (No NEA)	-
Philippines	High school diploma	Classroom assessment	Appears to be no national moderation (no information found)
Rwanda	Advanced General Certificate of Secondary Education (AGCSE)	School based assessment	Appears to be no national moderation (no information found)
Scotland	Highers	Internal assessment	Verification

Singapore	GCE A level	Group project – Most courses are assessed by exam, but students also have to complete a group project alongside their other studies	Inspection/verification (information unclear)
South Africa (Gauteng)	National Senior Certificate (NSC)	School based assessment	Moderation by inspection
Uganda	Uganda Advanced Certificate of Education (UACE)	Classroom assessment	Appears to be no national moderation (no information found)
United States of America (California)	High school diploma	Not applicable (No NEA)	-
West African Examinations Council (WAEC) ^b	West Africa Secondary School Certificate of Education (WASSCE)	Continuous assessment / school based assessment	Statistical moderation
Zimbabwe	ZGCE A level	Continuous assessment	Inspection/verification (information unclear)

^a CXC membership countries include Anguilla, Antigua and Barbuda, Barbados, Belize, British Virgin Islands, Cayman Islands, Dominica, Grenada, Guyana, Jamaica, Montserrat, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Trinidad and Tobago and Turks and Caicos Islands

^b WAEC membership countries include The Gambia, Ghana, Liberia, Nigeria, and Sierra Leone

Table 2. Purpose of non-examination assessment

Country/territory	Statement	Source
Australia (New South Wales)	"Assessment tasks allow students to show what they know, understand and can do in ways that may not be possible in a written examination."	(NESA, 2013)
Canada (Ontario)	"The primary purpose of assessment and evaluation is to improve student learning"	(Ontario Ministry of Education, 2010, p. 6)
Caribbean Examinations Council (CXC)	"School-based assessment...is intended to assess certain knowledge, skills and attitudes... which are not easily assessed in external examinations."	(CXC, 2014, p. 1)
England	"If skills and knowledge cannot be assessed through exams, exam boards test them through non-exam assessments"	(Ofqual, 2017b)
Hong Kong	"The main rationale for NEA is to enhance the validity of the public assessment and extend it to include a variety of learning outcomes that cannot be assessed easily through public examinations."	(HKEAA, 2013, p. 1)
Namibia	Syllabi emphasise the assessment of practical skills via NEA.	(eg NIED, 2005)
New Zealand	"Internal assessments are used to assess skills and knowledge that cannot be tested in an exam."	(NZQA, n.d.-f)
Scotland	"Construct validity concerns the extent to which an assessment actually measures what the Unit specification states it is intended to measure.	(SQA, 2017b, p. 8)

	For example, an assessment that asked a candidate to write about a skill rather than demonstrate it would have low construct validity.”	
Singapore	“The rationale for school-based assessment is that teachers would know best about the extent of effort put in by each group member... Manageability in terms of cost and efficiency of the assessment of oral presentation is [also] an important consideration... [as] it is more efficiently carried out by the teachers within the school”	(Chong & Leong, 2014, p. 4)
South Africa (Gauteng)	“School-based assessment (NEA) is a purposive collection of learners’ work that tells the story of learners’ efforts, progress or achievement in given areas. The quality of NEA tasks is integral to learners’ preparation for the final examinations.”	(DBE, 2014, p. 3)
West African Examinations Council (WAEC)	“Students tend to do well on [NEAs] because they are more relaxed when taking tests in their own classrooms... It is therefore generally accepted that the performance of students on [NEAs] better reflects the true performance of the student.” This quote relates to the BECE, as information could not be found in relation to the WASSCE. It has to be assumed that the same rationale exists for both.	(Republic of Ghana Ministry of Education, 2011, p. 1)
Zimbabwe	“Assessment information enables teachers, learners and other key stakeholders to know: what learners can do assisted; what they can do unassisted; what they can do when working in groups; and when working alone... Learner-centred assessments also give learners an	(MOPSE, 2015, p. 52)

	opportunity in making assessment a learning experience – assessment for learning not assessment of learning.”	
--	---------------------------------------------------------------------------------------------------------------	--

Table 3. Purpose of moderation

Country/territory	Statement	Source
Australia (New South Wales)	“All schools use different programs of assessment tasks and they all mark slightly differently. This means that students from different schools experienced different assessment conditions. For this reason, all HSC assessment marks that schools submit are adjusted by the Board using a process called moderation.	(NESA, n.d.-a)
Canada (Ontario)	“Assessment practices can have wide variance from classroom to classroom. Opportunities for professional dialogue about assessment practices bring coherence to those practices, nourish a climate of inquiry that supports student learning, and challenge teachers to focus future instruction on specific learning outcomes.”	(Literacy and Numeracy Secretariat, 2007, p. 1)
Caribbean Examinations Council (CXC)	“CXC moderates the school-based assessment in order to ensure that the assessment of the work of different schools carried out by different teachers using sometimes different tasks, is aligned to the standard of assessment defined by CXC.”	(CXC, 2014, p. 4)
England	“[Moderation should determine whether] the criteria against which learners’ performance is differentiated are being applied accurately and consistently by Assessors in different Centres, regardless of the identity of the Assessor, Learner, or Centre.”	(Ofqual, 2017a, para. H2.2)
Hong Kong	“Teachers know their students well and can reliably judge their performance. However, they are not necessarily aware of the standards of performance across all schools... The HKEAA makes use of appropriate methods to moderate NEA marks submitted by different	(HKEAA, 2013, p. 7)

	schools, with the aim of ensuring comparability of NEA scores across schools.”	
Namibia	Information could not be found on external moderation, so it has to be assumed that its purpose is similar in nature to the purpose of internal moderation: “When several teachers in a centre are involved in internal assessment, arrangements must be made within the centre for all candidates to be assessed to a common standard.”	(NIED, 2005, p. 23)
New Zealand	“National external moderation provides an assurance that assessment decisions, in relation to assessment standards, are consistent nationally... [and] are at the national standard.”	(NZQA, n.d.-a)
Scotland	“The purpose of external verification is to approve a centre’s assessment approach and... to make sure national standards are being applied consistently by all centres offering internally assessed SQA qualifications.”	(SQA, 2017a, p. 1)
Singapore	“To ensure that the national standards are applied consistently across all schools, the results of the school-based assessment are moderated internally as well as by an external team of moderators.”	(Chong & Leong, 2014, p. 5)
South Africa (Gauteng)	“Moderation is a process of teachers sharing their expectations of the performance of learners and their understanding of standards with each other in order to improve the consistency of their assessment decisions. It is the process of ensuring that the same assessment standards are applied to learners from all schools.”	(Gauteng Province Department of Education, 2012, p. 3)

West African Examinations Council (WAEC)	“The purpose of moderation is to introduce a common standard and to bring the assessments of individual... teachers into line with the standard... The reputation of an examining board rests upon the credibility of certificates issued in its name. It is therefore the responsibility of the West African Examinations Council to ensure that the final scores used in grading candidates are of high reliability and the assessment instrument of high validity.”	(Dery & Addy-Lampitey, n.d., p. 3)
Zimbabwe	“An assessment tool which produces stable, accurate and consistent results is reliable. This requires clarity and consistency in setting, marking, grading and moderation of scripts.”	(MOPSE, 2015, p. 53)

Table 4. Sampling of pupils and centres

Country/territory	Statement	Source
Australia (New South Wales)	All marks submitted by all schools (with the exception of some exclusion criteria) are included in statistical moderation.	(NESA, n.d.-a)
Canada (Ontario)	Consensus moderation is used, and so the sampling approach will presumably depend on the approach adopted within each school.	-
Caribbean Examinations Council (CXC)	Each centre submits one sample of 5 candidates per subject: highest, middle, and lowest marks, plus those mid-way between the highest and median, and between the median and lowest mark. CXC also randomly selects schools to audit NEA materials not included in moderation samples.	(CXC, 2014, sec. 5.1.1) (CXC, 2014, sec. 8.3)
England	Each centre submits a sample of candidates' work, the size of which differs by exam board, and depends on the size of the cohort. Samples should cover a range of attainment. Moderators first review a sub-sample of work, and will then review the full sample [NB still a sample, not whole centre] if issues with the marking are found.	(JCQ, 2017b)
Hong Kong	All marks submitted by all centres undergo statistical moderation. Samples are selected by HKEAA for any subject experts' judgements (size is unclear).	(HKEAA, 2010, secs. 4-5) (HKEAA, 2010, sec. 7.5)

	Samples do not need to include examples of marking from each teacher.	
Namibia	<p>Samples of pupils' work are selected by centres; the number depends on the size of the cohort. Samples cover the highest and lowest mark, with the rest evenly spread across the range of marks. The sample should also include examples of marking from each teacher. Further samples may be requested.</p> <p>For 10 or fewer candidates, all work is sampled. For more than 10 candidates, 10 candidates are sampled for moderation.</p>	(NIED, 2005)
New Zealand	<p>All centres randomly select samples of work for each course. If fewer than 8 samples are available, all work must be submitted for moderation. 4 samples are permitted where the cohort of learners only 'achieved' or 'not achieved' the standards ('merit' or 'excellence' are other possible outcomes).</p> <p>Not all courses are moderated for each centre each year</p> <p>Schools are also selected for an audit at least once every 4 years to check assessment validity/reliability, and that they comply with guidelines.</p>	<p>(NZQA, n.d.-c)</p> <p>(NZQA, 2017)</p> <p>(NZQA, n.d.-d)</p>
Scotland	Centres are sampled each year. A portion are selected randomly, but others are selected based on changes in entry (eg when offering new qualifications), or because their marking was not accepted in previous rounds of moderation. Selected centres are generally moderated for one subject per subject group (1 language, 1 science, etc.).	(SQA, n.d.-a)

	Selected centres are expected to submit a sample of evidence for a minimum of 6 candidates for moderation of Highers (while a sample of 12 is expected overall, this can be split equally between Highers and Advanced Highers qualifications).	(SQA, n.d.-b)
Singapore	Each centre submits a sample of marking for each course. It is unclear what the size of this sample should be.	(Chong & Leong, 2014)
South Africa (Gauteng)	A minimum of 10% of learner evidence is sampled for each centre. At district and provincial level, a minimum of 3-5% of learner evidence must be moderated. Not all centres within each school district are sampled.	(Gauteng Province Department of Education, 2014, para. 4.3.3)
West African Examinations Council (WAEC)	Statistical moderation is used. Assumed to sample all centres and candidates.	(Akuffo-Badoo, 2006)
Zimbabwe	Random sampling at centre, district, and provincial level. Numbers are unclear.	(Success Africa, 2017)

Table 5. Process of moderation

Country/territory	Process	Source
Australia (New South Wales)	<p>Statistical moderation.</p> <p>NEA marks are adjusted so that the NEA mean equals the exam mean, the top NEA mark equals the top exam mark, and where possible, the bottom NEA mark equals the bottom exam mark. A quadratic function (curved line) is used. The rank order of NEA marks is not affected.</p> <p>Outliers can be omitted from moderation, such as students who performed poorly in the exam, relative to their NEA performance.</p>	(NESA, n.d.-a)
Canada (Ontario)	<p>Consensus moderation.</p> <p>Professional discussions are held amongst teachers, in order to align decisions being made within the centre. Alignment between centres comes from working to the same core standards:</p> <p>“The principal will work with teachers to ensure common and equitable grading practices that follow ministry policy and board guidelines.”</p>	<p>(Literacy and Numeracy Secretariat, 2007)</p> <p>(Ontario Ministry of Education, 2010, p. 39)</p>
Caribbean Examinations Council (CXC)	<p>Moderation by inspection with statistical scaling.</p> <p>A sample of each school’s scripts are remarked by moderators. CXC then provides feedback to the teachers who originally marked the work.</p>	(CXC, 2014, sec. 5.1.3)

	<p>All marks for each class are then scaled based upon the relationship between teachers' and moderators' marks. The exact method is unclear.</p> <p>CXC also conducts random audits, to inspect NEA materials not included in moderation samples.</p>	<p>(Griffith, 2016)</p> <p>(CXC, 2014, sec. 8.3)</p>
England	<p>Moderation by inspection with statistical scaling.</p> <p>Moderators review samples of centres' marking. If a moderator disagrees with the marks awarded beyond a specified tolerance, then a further sample for that centre is reviewed. Any necessary adjustments are made by scaling the centres marks to fall on a regression line based upon the relationship between the centre's original marks and the moderated marks. While only a sample of work is reviewed, all candidates for the centre are scaled in the same manner. Where the moderator believes that the centre has got the rank order of marks wrong, then re-marking may be necessary. Outliers (eg clerical errors) can sometimes be removed from scaling where present.</p> <p>In most cases, students' work are posted to moderators, but for performance or artefact submissions, moderators may need to visit centres to moderate in person.</p>	<p>For each of the 3 main exam boards in England:</p> <p>(AQA, 2013; Gill, 2015; Pearson, n.d.)</p>
Hong Kong	<p>Statistical moderation with checks by subject experts, or moderation by inspection with statistical scaling (depending on the subject).</p> <p>Statistical approach – Each centre's NEA marks are adjusted to match the mean and standard deviation of that centre's exam scores. The</p>	<p>(HKEAA, 2010, sec. 4)</p> <p>Formula:</p>

	<p>rank order of NEA scores is not changed. Outliers are excluded from the model, but are still scaled. Samples of work are reviewed by subject experts, to check that the outcomes of moderation are appropriate. For example, some centres' NEA marks may be legitimately different from their exam marks.</p> <p>Expert judgement approach – This has similar aims to the above (ie to match scores by mean and standard deviation), but uses subject experts' marks as the reference, instead of examination marks. Where large difference exist between a school's original marks and moderated marks, these marks will undergo further review.</p>	<p>(HKEAA, 2010, pp. 24–25)</p> <p>(HKEAA, 2010, sec. 5)</p> <p>Formula: (HKEAA, 2010, pp. 26–27)</p>
Namibia	<p>Inspection/verification.</p> <p>The documentation refers to the work of 'moderation teams' and so an inspection/verification approach is assumed. Further detail could not be found.</p>	<p>(Republic of Namibia Ministry of Education, 2016)</p>
New Zealand	<p>Verification.</p> <p>Subject experts evaluate assessment materials and whether teachers' marking is consistent with the standard. Centres are responsible for addressing feedback.</p> <p>Schools are also audited at least once every 4 years to check assessment validity/reliability, and that they comply with guidelines.</p>	<p>(NZQA, n.d.-b)</p> <p>(NZQA, n.d.-d)</p>
Scotland	<p>Verification.</p>	<p>(SQA, 2017e)</p>

	<p>Most moderation is conducted remotely, although some is done via moderation visits to the centre</p> <p>Moderators review teachers' assessments, and must decide whether to accept or not accept them. Feedback for improvement may still be given for the former. For the latter, moderators work with the centre to develop a plan for corrective action, but ultimately the centre is responsible for carrying this action out.</p>	(SQA, 2014)
Singapore	<p>Inspection/verification.</p> <p>Moderators check marking for consistency and leniency/severity. It is unclear what the outcomes of the moderation process is, and who has responsibility for making corrective action: "Where school-based assessment is deemed too lenient or severe, the external moderation ensures that the school-based assessment is brought in line with the national standards."</p>	(Chong & Leong, 2014, p. 5)
South Africa (Gauteng)	<p>Moderation by inspection.</p> <p>Districts are responsible for moderating for consistency between schools, and provinces are responsible for moderating for consistency between districts. Moderators remark the samples and provide feedback to teachers. Where appropriate, moderators "recommend adjustments of marks".</p> <p>Mark adjustments, however, are made by 'subject advisors', following the recommendations of moderators.</p>	<p>(Gauteng Province Department of Education, 2012, p. 5)</p> <p>(Gauteng Province Department of Education, 2014, para. 5.5.5)</p>

West African Examinations Council (WAEC)	<p>Statistical moderation.</p> <p>In most cases, scaling is used: each centre's marks are adjusted so that they align with the mean and standard deviation of exam marks for the centre.</p> <p>When the entry size is small, mapping is used. The top NEA mark is aligned with the top mark on the exam. The next top exam score is awarded to the next top NEA candidate, and so on.</p> <p>The teachers' rank order for NEA marks remains unchanged in both cases.</p>	(Akuffo-Badoo, 2006, pp. 39–41) ⁴
Zimbabwe	<p>Inspection/verification.</p> <p>Details on the exact approach could not be found. As random sampling is used, one can assume that moderation is based upon expert judgements (statistical moderation usually targets all schools).</p>	(Success Africa, 2017)

⁴ Akuffo-Badoo describes the moderation process for the Basic Education Certificate Examination (BECE). Dery & Addy-Lampitey (n.d.) implies that scaling and mapping is also used for the WASSCE (which we are focussing upon here) – it has to be assumed that the same scaling and mapping processes are used for the WASSCE as for the BECE.

Table 6. Appealing moderation outcomes

Country/territory	Statement	Source
Australia (New South Wales)	Candidates can lodge an appeal with their centre if they are unhappy with their school assessment outcomes, and can further appeal to the New South Wales Education Standards Authority if they are not happy with the centre's response. If the appeal is upheld, the centre is responsible for making corrective action.	(NESA, n.d.-b)
Canada (Ontario)	Information not found.	-
Caribbean Examinations Council (CXC)	Information not found.	-
England	Centres can submit for a 'review of moderation'. A second moderator reviews the same work as the original moderator. Individual marks cannot be challenged after moderation, but rather all marks for the centre. An 'automatic protection' rule is currently in place, which means that marks cannot be reduced as a result of a review of moderation. A fee is charged only if the centre's original marks are not reinstated (ie the moderator's decisions were upheld)	(JCQ, 2017c)
Hong Kong	Centres are expected to resolve any issues before submitting marks for moderation. After results are released, students can appeal to HKEAA	(HKEAA, 2013)

	for re-checking of their marks, but not a re-assessment of their performance.	
Namibia	Learners can lodge an appeal with the NQA. Following which, an independent evaluator will re-evaluate assessments. If unhappy with the outcome, the learner can submit a further appeal.	(NQA, n.d.)
New Zealand	Centres can appeal moderation outcomes through NZQA, who reviews the assessment materials.	(NZQA, n.d.-e)
Scotland	Centres can appeal assessment outcomes, including moderation decisions. Appeals are considered by subject experts who were not involved in the original decision.	(SQA, 2017c, 2017d)
Singapore	For a fee, candidates can appeal their A level results via their school. Presumably, this includes the group project.	(SEAB, 2017)
South Africa (Gauteng)	Candidates can appeal to re-mark/re-check examination scripts, but whether there is a similar process for NEA is unclear.	(DBE, 2017)
West African Examinations Council (WAEC)	School candidates can send any 'complaints' (not necessarily relating to NEA) to WAEC via their school principals.	(WAEC, n.d.)
Zimbabwe	A "queries and appeals" service seems to be available (see http://www.zimsec.co.zw/senior-management), but details on this could not be found.	-

References

- Akuffo-Badoo, F. (2006). *The effect of moderation of continuous assessment scores on the performance scores of candidates at the Basic Education Certificate Examination level*. University of Cape Coast. Retrieved from <https://erl.ucc.edu.gh/jspui/handle/123456789/1615>
- AQA. (2013). *Moderation of internal assessments*. Retrieved from <http://www.aqa.org.uk/exams-administration/coursework-and-controlled-assessment/moderation>
- Chong, K. K. K., & Leong, S. C. (2014, May). *Assessment of 21st Century Skills: the Singapore Experience with Group Project at Grade 11*. Paper presented at International Association for Educational Assessment Conference. Retrieved from http://www.iaea.info/documents/paper_226dc2ffd.pdf
- Cuff, B. M. P. (2017). *An exploratory investigation into how moderators of non-examined assessments make their judgements*. (Ofqual Report 17/6252). Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from <https://www.gov.uk/government/publications/exam-and-assessment-marking-research>
- Cuff, B. M. P., Howard, E., Mead, R., & Newton, P. E. (2018). *Key stage 2 writing moderation: Observations on the consistency of moderator judgements*. (Ofqual Report 18/6358). Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from <https://www.gov.uk/government/publications/observations-on-the-consistency-of-moderator-judgements>
- CXC. (2014). *School-Based Assessment Manual for Principals: Caribbean Advanced Proficiency Examinations (CAPE)*. Retrieved from <https://www.cxc.org/examinations/cape/manuals/>
- Daly, A., Billington, L., Chamberlain, S., Meyer, L., Stringer, N., Taylor, M., & Tremain, K. (2011). *Principles of moderation of internal assessment*. Manchester, UK: Centre for Education Research and Policy. Retrieved from <https://cerp.aqa.org.uk/research-library/principles-moderation-internal-assessment>
- DBE. (2014). *Geography: School-based assessment exemplars - CAPS Grade 12 teacher guide*. Republic of South Africa: Department for Basic Education. Retrieved from <https://www.education.gov.za/SchoolBasedAssessmentTasks2014/tabid/611/ItemId/6131/Default.aspx>
- DBE. (2017). FAQs on Examinations. Retrieved December 20, 2017, from [https://www.education.gov.za/Curriculum/NationalSeniorCertificate\(NSC\)Examinations.aspx](https://www.education.gov.za/Curriculum/NationalSeniorCertificate(NSC)Examinations.aspx)
- Dery, R., & Addy-Lampitey, W. (n.d.). Effects of classroom assessment scores on the final scores used in grading students at senior high schools in Ghana. Retrieved from http://www.iaea.info/documents/paper_4d53997.pdf

- Gauteng Province Department of Education. (2012). *Examination Instruction No.10 of 2012*. Retrieved from http://www.gauteng.gov.za/government/departments/education/ExamInstructions/2012/exam_instruction_10_of_2012.pdf#search=moderation
- Gauteng Province Department of Education. (2014). *Examination Instruction No.01 of 2014*. Retrieved from [http://www.gauteng.gov.za/government/departments/education/Documents/Assessment Guideline 01 of 2014.pdf#search=moderation](http://www.gauteng.gov.za/government/departments/education/Documents/Assessment%20Guideline%2001%20of%202014.pdf#search=moderation)
- Gill, T. (2015). The moderation of coursework and controlled assessment: A summary. *Research Matters: A Cambridge Assessment Publication*, 19, 26–30. Retrieved from http://www.cambridgeassessment.org.uk/our-research/all-published-resources/research-matters/index.aspx?filter=year:2015%7Cpublish_date#
- Griffith, S. A. (2016, March). *Quality Assurance in the School Based Assessment Component of a Public Examination*. Paper presented at Association of Commonwealth Examinations and Accreditation Bodies(ACEAB) Conference. Retrieved from <https://tinyurl.com/y6vyegq9>
- HKEAA. (2010). *Moderation of School-based Assessment Scores in the HKDSE*. Retrieved from http://www.hkeaa.edu.hk/en/sba/moderation_sba_scores/
- HKEAA. (2013). *Hong Kong Diploma of Secondary Education Examination: Information on School-based Assessment*. Retrieved from http://www.hkeaa.edu.hk/en/sba/sba_hkdse
- JCQ. (2017a). *Instructions for conducting non-examination assessments (new GCE & GCSE specifications): 1 September 2017 to 31 August 2018*. London, UK: Joint Council for Qualifications. Retrieved from <https://www.jcq.org.uk/exams-office/non-examination-assessments/in>
- JCQ. (2017b). *Legacy GCE unitised AS and A-level qualifications: ELC and Project qualifications: Instructions for conducting coursework*. London, UK: Joint Council for Qualifications. Retrieved from <https://www.jcq.org.uk/exams-office/coursework/instructions-for-conducting-coursework-2017-2018>
- JCQ. (2017c). *Post-Results Services*. London, UK: Joint Council for Qualifications. Retrieved from <https://www.jcq.org.uk/exams-office/post-results-services/post-results-services-june-2017-and-november-2017>
- Johnson, S. (2011). *A focus on teacher assessment reliability in GCSE and GCE*. (Ofqual Report 11/4807). Coventry, UK: Ofqual. Retrieved from <https://www.gov.uk/government/publications/a-focus-on-teacher-assessment-reliability-in-gcses-and-a-levels>
- Klenowski, V., & Wyatt-Smith, C. (2014). *Assessment for education: Standards, judgement and moderation*. London, UK: SAGE.
- Linn, R. L. (1993). Linking Results of Distinct Assessments. *Applied Measurement in Education*, 6, 83–102. http://doi.org/10.1207/s15324818ame0601_5

- Literacy and Numeracy Secretariat. (2007). *Teacher Moderation: Collaborative Assessment of Student Work*. Retrieved from <http://www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/capacityBuilding.html>
- Martínez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgements, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment*, 14(2), 78–102. <http://doi.org/10.1080/10627190903039429>
- Maxwell, G. (2002). *Moderation of teacher judgments in student assessment*. Brisbane, Australia: Queensland School Curriculum Council. Retrieved from https://www.qcaa.qld.edu.au/downloads/publications/research_qscs_assess_report_2.pdf
- McCurry, D. (2013). *Overview of Senior Assessment and Tertiary Entrance in Australia and other countries*. Australian Council for Educational Research. Retrieved from <https://www.acer.org/queensland-review/informing-papers-and-research>
- MOPSE. (2015). *Curriculum Framework for Primary and Secondary Education 2015-2022*. Zimbabwe Ministry of Primary and Secondary Education. Retrieved from <http://www.mopse.gov.zw/index.php/downloads-key-resources/>
- NESA. (n.d.-a). Moderation. Retrieved December 4, 2017, from <https://tinyurl.com/y9mgmnqh>
- NESA. (n.d.-b). School assessment. Retrieved December 7, 2017, from <http://educationstandards.nsw.edu.au/wps/portal/nesa/11-12/hsc/about-HSC/school-assessment>
- NESA. (2013). HSC assessment. Retrieved December 4, 2017, from http://www.boardofstudies.nsw.edu.au/hsc_assessment_policies/
- NIED. (2005). *Agriculture Syllabus*. Republic of Namibia: The National Institute for Educational Development. Retrieved from http://www.ibe.unesco.org/curricula/namibia/sx_us_ag_2005_eng.pdf
- NQA. (n.d.). Appeals. Retrieved December 14, 2017, from <http://www.namqa.org/about-us/Appeals/90/>
- NZQA. (n.d.-a). External moderation. Retrieved December 4, 2017, from <http://www.nzqa.govt.nz/providers-partners/assessment-and-moderation/managing-national-assessment-in-schools/secondary-moderation/external-moderation/>
- NZQA. (n.d.-b). External moderation outcomes. Retrieved December 4, 2017, from <http://www.nzqa.govt.nz/providers-partners/assessment-and-moderation/managing-national-assessment-in-schools/secondary-moderation/external-moderation/external-moderation-outcomes/>
- NZQA. (n.d.-c). *Guidelines for the Selection of Student Work for Moderation*. Retrieved from <http://www.nzqa.govt.nz/providers-partners/assessment-and-moderation/managing-national-assessment-in-schools/secondary-moderation/external-moderation/external-moderation-outcomes/>

moderation/managing-national-assessment-in-schools/secondary-moderation/external-moderation/

NZQA. (n.d.-d). MNA review. Retrieved December 4, 2017, from <http://www.nzqa.govt.nz/providers-partners/assessment-and-moderation/managing-national-assessment-in-schools/mna-review/>

NZQA. (n.d.-e). Query or appeal a moderation report. Retrieved December 7, 2017, from <http://www.nzqa.govt.nz/providers-partners/assessment-and-moderation/managing-national-assessment-in-schools/secondary-moderation/external-moderation/appeal-a-moderation-report/>

NZQA. (n.d.-f). Standards. Retrieved December 7, 2017, from <http://www.nzqa.govt.nz/ncea/understanding-ncea/how-ncea-works/standards/>

NZQA. (2017). Assessment (including Examination) Rules for Schools with Consent to Assess 2017. Retrieved December 4, 2017, from <http://www.nzqa.govt.nz/about-us/our-role/legislation/nzqa-rules/assessment-including-examination-rules-2017/6/5/>

Ofqual. (2017a). Ofqual Handbook: Section H - From marking to issuing results. Retrieved December 13, 2017, from <https://www.gov.uk/guidance/ofqual-handbook/section-h-from-marking-to-issuing-results>

Ofqual. (2017b). Regulating GCSEs, AS and A levels: guide for schools and colleges. Retrieved December 1, 2017, from <https://www.gov.uk/guidance/regulating-gcses-as-and-a-levels-guide-for-schools-and-colleges/grading>

Ontario Ministry of Education. (2010). *Growing Success: Assessment, evaluation, and reporting in Ontario schools*. Ontario, Canada. Retrieved from <http://www.edu.gov.on.ca/eng/policyfunding/success.html>

Pearson. (n.d.). *Coursework moderation and mark adjustments (an explanation for centres)*. Retrieved from <http://qualifications.pearson.com/en/support/support-topics/results-certification/moderator-reports.html>

Republic of Ghana Ministry of Education. (2011). *Teachers' handbook on school based assessment for junior high schools: English language*. Accra, Ghana: Ministry of Education. Retrieved from <https://mingycomputersgh.wordpress.com/school-based-assessmentsba-corner/>

Republic of Namibia Ministry of Education. (2016). *Report on the examinations NSSCO*. Retrieved from http://www.nied.edu.na/assets/documents/07Reports/NSSCO_Examiners_Report_2016.pdf

SEAB. (2017). Frequently Asked Questions. Retrieved December 7, 2017, from http://www.ifaq.gov.sg/seab/apps/fcd_faqmain.aspx?FAQ=93303

SQA. (n.d.-a). How we select centres for verification - National Qualifications. Retrieved December 5, 2017, from <https://www.sqa.org.uk/sqa/77172.html>

SQA. (n.d.-b). *NQ verification 2017–18*. Retrieved from

- https://www.sqa.org.uk/sqa/files_ccc/Generating_the_evidence_sample.pdf
- SQA. (2014). *Handbook for NQ External Verifiers*. Glasgow, Scotland: Scottish Qualifications Authority. Retrieved from https://www.sqa.org.uk/sqa/files_ccc/NQ_Handbook_EV.pdf
- SQA. (2017a). *Evidence required for external verification of units (including Added Value units) at verification events*. Retrieved from https://www.sqa.org.uk/sqa/files_ccc/Evidence_required_for_verificationevents.pdf
- SQA. (2017b). *Guide to Assessment*. Glasgow, Scotland: Scottish Qualifications Authority. Retrieved from https://www.sqa.org.uk/files_ccc/Guide_To_Assessment.pdf
- SQA. (2017c). *Quality Assurance for Higher National and Vocational Qualifications 2015–18: Enhanced guidance on meeting quality assurance criterion 4.8*. Glasgow, Scotland: Scottish Qualifications Authority. Retrieved from https://www.sqa.org.uk/sqa/files_ccc/Internal-Assessment-Appeals-Procedure-guidance.pdf
- SQA. (2017d). *The Appeals Process: Information for Centres*. Glasgow, Scotland: Scottish Qualifications Authority. Retrieved from https://www.sqa.org.uk/files_ccc/Appeals_Process.pdf
- SQA. (2017e). *Type of verification for units (including added value units)*. Retrieved from https://www.sqa.org.uk/sqa/files_ccc/TypeofverificationforUnits.pdf
- Stanley, G., MacCann, R., Gardner, J., Reynolds, L., & Wild, I. (2009). *Review of teacher assessment: Evidence of what works best and issues for development*. Coventry, UK: Qualifications and Curriculum Authority. Retrieved from http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/01/2009_03-Review_of_teacher_assessment-QCA.pdf
- Success Africa. (2017). ZIMSEC finally introduces continuous assessment. Retrieved December 6, 2017, from <https://success.africa/education/zimsec-finally-introduces-continuous-assessment/>
- The World Bank. (2016). GDP per capita (current US\$). Retrieved December 15, 2017, from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?year_high_desc=true
- The World Bank. (2017). Total population. Retrieved January 8, 2018, from https://data.worldbank.org/indicator/SP.POP.TOTL?year_high_desc=true
- WAEC. (n.d.). FAQ. Retrieved December 7, 2017, from <http://www.waecnigeria.org/FAQ.aspx>
- Williamson, J. (2016). Statistical moderation of school-based assessment in GCSEs. *Research Matters: A Cambridge Assessment Publication*, 22, 30–36. Retrieved from <http://www.cambridgeassessment.org.uk/Images/322752-research-matters-22-summer-2016.pdf>
- Wilmot, J., & Tuson, J. (2005). *Statistical moderation of teacher assessments*.

London, UK: Qualifications and Curriculum Authority. Retrieved from
<http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/06/Wilmut-2004-review-of-TA.pdf>

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



© Crown copyright 2018

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.

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.

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