
Moderation of teacher assessment at key stage 2 and key stage 3: a review of accuracy and consistency

September 2016

PHOTO REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



The purpose of Estyn is to inspect quality and standards in education and training in Wales. Estyn is responsible for inspecting:

nursery schools and settings that are maintained by, or receive funding from, local authorities;

- ▲ primary schools
- ▲ secondary schools
- ▲ special schools
- ▲ pupil referral units
- ▲ all-age schools
- ▲ independent schools
- ▲ further education
- ▲ independent specialist colleges
- ▲ adult community learning
- ▲ local authority education services for children and young people
- ▲ teacher education and training
- ▲ Welsh for adults
- ▲ work-based learning
- ▲ learning in the justice sector

Estyn also:

- ▲ provides advice on quality and standards in education and training in Wales to the National Assembly for Wales and others
- ▲ makes public good practice based on inspection evidence

Every possible care has been taken to ensure that the information in this document is accurate at the time of going to press. Any enquiries or comments regarding this document/publication should be addressed to:

Publication Section

Estyn

Anchor Court

Keen Road

Cardiff

CF24 5JW or by email to publications@estyn.gov.wales

This and other Estyn publications are available on our website: www.estyn.gov.wales

© Crown Copyright 2016: This report may be re-used free of charge in any format or medium provided that it is re-used accurately and not used in a misleading context. The material must be acknowledged as Crown copyright and the title of the document/publication specified.

Contents	Page
Introduction	1
Background	1
Main findings	3
Recommendations	4
School pre-cluster moderation and standardisation	5
Cluster moderation meetings	10
Schools' actions in response to cluster moderation	12
The role of consortia and local authority	14
Appendix: Evidence base	15
Glossary	16
References	17

Introduction

This report is written in response to a request for advice from the Welsh Government in the Minister's annual remit letter to Estyn for 2016-2017. The report examines the effectiveness of cluster moderation and standardisation procedures in securing reliable end of key stage 2 and 3 teacher assessment.

The report is intended for the Welsh Government, headteachers and staff in schools, local authorities and regional consortia.

The findings of the survey are based on visits to eight cluster moderation meetings, two in each of the regional consortia, followed by visits to two schools in each cluster. During the visits, inspectors considered the effectiveness of cluster moderation meetings and of schools' internal procedures for ensuring accurate teacher assessment. Inspectors focused specifically on the moderation of English, Welsh and Welsh second language. They also evaluated the extent to which providers are familiar with the new statutory requirements for cluster moderation.

Background

Since 2006, when the external marking of standard assessment tests was discontinued, schools have produced teacher assessment levels in a range of subject areas as the key summative measure of pupils' progress. Teachers make these assessments at the end of key stage 2 (Year 6) and at the end of key stage 3 (Year 9). The judgements are made on the full body of work completed by each pupil during the key stage on a 'best fit' basis.

Schools are expected to ensure that exemplification materials, which have been produced by the Welsh Government, are used to inform teachers' understanding of the standards associated with each level. Schools are required to moderate internally to ensure consistency and reliability across the school. Secondary schools work with their partner primary schools in cluster groups to moderate levels across key stages 2 and 3 with the intention of securing consistency in the application of levels. Regional consortia or local authorities oversee these meetings as well as support schools in their assessment of pupils' work.

Estyn reports and a number of Welsh Government publications have raised concerns regarding the accuracy, consistency and reliability of teacher assessments over a number of years. In 2010, the Welsh Government asked Estyn to complete a thematic survey on teacher assessments. The Welsh Government also commissioned a report by the Australian Council for Educational Research (ACER), *An investigation into Key Stages 2 and 3 teacher assessment in Wales*, in 2013. Both reports found that there were questions about the accuracy of judgements on pupils' standards, with inconsistency in cluster moderation arrangements and how schools used guidance.

The Welsh Government has issued a range of documents, both generic and subject specific, to support teacher assessment, including:

- *English Guidance for Key Stages 2 and 3* (2008a)
- *Cymraeg - Canllawiau ar gyfer Cyfnodau Allweddol 2 a 3* (Welsh only) (2008b)
- *Making the most of assessment 7–14*. (2010)
- *Programme of Study for English Key Stages 2–4* (2015a)
- *Programme of Study for Welsh Key Stages 2–4* (2015b)
- *Welsh Second Language in the National Curriculum for Wales* (2015c)
- *Statutory assessment arrangements for the Foundation Phase and end of Key Stages 2 and 3* (2015d).

In 2015, the National Assembly for Wales introduced *The National Curriculum (Moderation of Assessment Arrangements for the Second and Third Key Stages) (Wales) Order 2015*. This order requires all schools to take part in cluster group moderation at the end of key stage 2 and key stage 3. It also makes statutory the participation of all headteachers and/or relevant staff in moderation meetings of the relevant clusters. The order also states that the chief education officer of the local authority may attend school moderation cluster meetings.

In summer 2015, the Welsh Government introduced a national programme, 'Securing Teacher Assessment Programme of External Verification', commonly referred to as EV. The programme was designed to strengthen the accuracy and consistency of teacher assessment through high-quality feedback to teachers, headteachers and all key stakeholders. The Welsh Government awarded the contract to a partnership of the four regional consortia and CDSM Interactive Solutions, supported by the Association of Directors of Education in Wales. The programme has undertaken two annual reviews to date. This survey does not evaluate the effectiveness of this programme.

Main findings

- 1 In-school moderation of teacher assessment is effective where there is enough evidence available. However, in a minority of schools, the range of evidence does not cover all areas of the level descriptor or show evidence of drafting or teacher support in completing work. In these cases, schools award the level despite a limited range of evidence. Where the evidence focuses on final versions or selected best pieces, this leads to inflation of levels achieved.
- 2 In cluster moderation, most schools bring either work that illustrates secure or higher-end version of levels. As a result, meetings focus too much on moderating levels of work that teachers already know to be accurate. This means that clusters waste time in focusing on activities that do not add value. There is not enough focus in the cluster moderation meeting on discussing pupils' work that is on the borderline between levels.
- 3 In a few cases, teachers consider the whole range of a pupil's work, including whole workbooks and drafts, during in-school and cluster moderation meetings. This means that teachers can gain a more holistic, 'best fit' view of the pupil's standards, by sampling from the whole range of the pupil's work during moderation. This strengthens the moderation process by reducing unconscious bias in using selected pieces of work, such as final drafts, that make up a learner profile.
- 4 Only a minority of schools refer to their own or to Welsh Government exemplification and standardisation materials to check their decisions during internal moderation meetings. In most cluster moderation meetings, exemplification and standardisation materials are available, but these materials are rarely referred to or used.
- 5 Nearly all schools undertake additional assessment work to prepare learner profiles for the cluster moderation meeting. These include extensive labelling of learners' work, identifying the evidence for each element of the level descriptor, and preparing 'pupil commentaries' to explain the level awarded. While this practice is useful in strengthening teachers' understanding of the criteria for levels, it is often an unnecessary task for cluster moderation meetings, as the profiles submitted generally illustrate a secure level and there is little disagreement about the levels.
- 6 Cluster reports following moderation meetings focus mainly on procedural matters with very few identifying actions to review and amend the levels of learner profiles. While many schools would alter the level of individual learner profiles if directed by the cluster, the extent of agreement found in cluster moderation meetings limits the need for further action. Schools do not routinely review and amend the levels of learner profiles for the rest of the cohort.
- 7 Local authorities and regional consortia support most schools well in the procedures for moderation. However, the role of local authorities and regional consortia representatives in the moderation of standards is unclear regarding their role in ensuring consistency of levels across clusters.

- 8 This year, cluster moderation focused on English and Welsh. The range of evidence to support pupils' standards in writing is stronger than that for reading, especially in secondary schools. In a few secondary schools there is not enough variety or opportunity in pupils' work to assess the different types of reading skills, particularly higher-order skills such as synthesis. In a minority of primary and secondary schools there is either too little recorded evidence or evidence of poor quality to support teacher assessment of oracy.

Recommendations

Local authorities and regional consortia should:

- R1 Develop training opportunities for schools to ensure consistency in the judgements for oracy, sufficiency of the evidence base, better application of the 'best fit' method, and moderation of work on the borderline between levels
- R2 Review their role in ensuring consistency of standards across clusters, authorities and regions

Schools should:

- R3 Take account of a wide range of pupils' work when assessing and moderating levels
- R4 Take appropriate account of the level of support, drafting processes, the impact of teachers' marking and the sufficiency of evidence when awarding a final level
- R5 Focus on pupils' work that is on the lower borderline of levels when moderating in schools and in cluster meetings
- R6 Make sure that all levels are reviewed and adjusted suitably after internal and cluster moderation and before submitting final levels
- R7 Refer to standardised materials when assessing, moderating and standardising in schools and in cluster meetings

School pre-cluster moderation and standardisation

- 9 In nearly all schools, pupils undertake a lot of work, leading to a wide range of evidence for assessment for English and Welsh. While the range of evidence is generally broad for writing and, in most cases, for reading, there are weaknesses in the evidence base for oracy.
- 10 Nearly all primary schools use an extensive range of writing and reading tasks to identify the level of pupils' work. These include writing for different purposes, for example to persuade people to donate to a charity, to inform the general public about a news story, to discuss with the school community issues such as uniform and to narrate an imaginative story. This evidence is collected from pupils' work across the curriculum and not only from their work in Welsh or English lessons. This provides a comprehensive picture of the level of work that pupils achieve.
- 11 Most primary schools collect a range of evidence to evaluate pupils' levels of oracy. These include group discussions on topical issues, individual presentations, news reports and role play conversations, for example. However, in a few schools there is too little recorded evidence and this hinders moderation.
- 12 In most secondary schools, the range of tasks for reading and writing is extensive for Welsh and English. However, in a few secondary schools, there is not enough variety in pupils' work to assess the different types of reading skills. In these cases, teachers rely too much on essays focusing on analysis of writers' style. In these schools, pupils do not have enough opportunity to develop and demonstrate higher-order reading skills such as synthesis and assimilation of independent research, nor is there enough range or sophistication of reading materials.
- 13 In most secondary schools, there are suitable and varied opportunities for pupils to develop and demonstrate appropriate oral skills in English and Welsh. However, in a minority of schools there is too little recorded evidence to standardise and moderate oracy. This lack of evidence, along with the poor quality of recording, limits schools' ability to moderate pupils' levels of achievement. The technical challenges in securing consistent and reliable evidence for oracy are considerable, especially when recording and storing a suitably wide range of examples for large cohorts of pupils.
- 14 In a few primary and secondary schools, the tasks set limit the opportunities for pupils to achieve the higher levels or to make progress to a higher level.
- 15 Primary and secondary schools base the award of pupils' levels on the evidence provided for moderation. In a minority of these schools, where there is not enough evidence to cover all areas of the level descriptor, they do not take action to gather more evidence to confirm a pupil's level of achievement. In these cases, they award the level despite the lack of evidence. This often leads to a distortion of standards achieved.
- 16 The Welsh Government has provided schools with clear guidance that level descriptions should be used only when taking account of a wide range of a pupil's work. More recently, regional consortia have also provided schools with guidance to consider the full range of pupils' work when awarding levels.

- 17 Many schools now focus on pupils' whole books when judging the overall level at the end of a key stage.
- 18 While nearly all primary schools use a wide range of evidence to support assessment and to make judgements of the overall standards at the end of a key stage, a few use only selected examples for writing, reading and oracy to take forward to cluster moderation meetings.
- 19 A minority of secondary schools still focus on using set, assessed pieces of work as the learner profile to level pupils' achievements, rather than the whole range of their work. This limits the range of skills a pupil can demonstrate and can obscure the actual 'best fit' level achieved by pupils.
- 20 Many schools include examples of final versions as part of the learner profile. In many of these cases, schools also provide first and subsequent drafts, which allow the teacher and/or moderator to see pupils' progression over time. The use of the whole book is improving this process. This allows these schools to reach a judgement on final pupil levels by sampling from the whole range of pupils' work. In a few cases, schools include evidence in a range of formats and processes, for example under test conditions, as homework, handwritten and information and communication technology (ICT) produced and single and multiple draft versions. This variety of approach, including using the whole book, provides a more reliable base on which to assess the standards that pupils have reached.
- 21 In many primary schools and in most secondary schools, teachers give levels when marking individual pieces of work, often final versions. This is contrary to advice from the Welsh Government, and means that the levels reported are misleading. In a very few cases, despite schools using pupils' entire workbooks as part of the levelling and moderation process, the focus on levelling final drafts of individual pieces of work has directly led to inflation of overall levels.
- 22 In a minority of cases, drafting and support provided by the teacher is not documented in pupils' workbooks or the learner profiles. This is unhelpful and leads to inaccurate judgements of levels.

School moderation procedures

- 23 The Welsh Government's *Statutory assessment arrangements for the Foundation Phase and end of Key Stages 2 and 3* (2015d) requires headteachers to 'have in place arrangements, using selected learner profiles, so that teachers moderate end of key stage assessments and apply the outcomes from this internal moderation prior to finalising all learners' end of key stage attainment, (p.9).' The Welsh Government has also provided guidance in *Making the most of assessment 7–14* (2010) that schools should not use level descriptions for individual pieces of work. They caution schools to take care to 'avoid an inappropriate, narrow focus for assessment' if they use levels to monitor or track pupils' progress over short periods of time, (p.9).
- 24 In many schools there is some confusion regarding standardisation and the statutory requirements for end of key stage assessment moderation. Most schools hold meetings throughout the year to moderate the accuracy of teachers' assessment,

track pupils' standards by taking indicative levels at key times and make adjustments to assessment for learning practices. Although these activities are productive, on their own they do not replace the moderation requirements outlined in Welsh Government advice.

- 25 The normal outcome of moderation is to affirm or amend the final levels achieved by pupils. However, the outcome of the moderation processes undertaken by many schools, although productive, is different. For many, especially earlier in the school year, it results in affirming or adjusting the ways in which teachers are assessing. This is more in line with the role of standardisation. In a minority of cases, the outcome of moderation is used to inform what pupils need to do to achieve the next level, by amending or extending the pupils' learning experiences. This is more in line with developing assessment for learning practices. Most schools also use their moderation procedures to prepare for cluster moderation meetings. Overall, there is not always a clear enough understanding of the role of standardisation and of moderation in schools. While these processes of monitoring assessment are all productive, there is not enough focus on a robust internal moderation of the final levels at the end of the key stage across the whole cohort.
- 26 The approach to moderation varies between primary and secondary schools. In nearly all primary schools, the process of moderation involves Year 6 teachers and senior leaders. In most cases, primary schools also use data analysis or assessment tracking programmes to review internal moderation, checking that indicative and final levels align with expected targets.
- 27 In secondary schools, the main responsibility for moderation lies with middle leaders. In the majority of secondary schools, senior leaders use line management meetings, or undertake data analyses to monitor this process. However, in a few secondary schools, senior leaders do not check the moderation processes closely enough, nor do they check whether levels are changed by departments as a result of moderation.
- 28 Nearly all primary and secondary schools ensure that teachers bring a suitable range of learner profiles at different levels to internal moderation. In many cases, this consists of examples from higher, average and lower ability pupils to demonstrate the range of work across the whole class or secure and borderline examples from a particular level. Only a minority of schools strategically target work that is on the borderline between levels or undertake random sampling. In the majority of cases, this reduces the reliability of moderation in securing the accuracy of teacher assessment across the whole cohort.
- 29 In many primary schools, there is a collaborative approach to moderation, which develops a greater understanding of assessment and levelling. In a few schools, especially where there is only one Year 6 teacher, senior leaders or teachers with responsibility for assessment moderate the work. In secondary schools, there is a similar collaborative approach, but this operates at a departmental level. In a minority of secondary schools, the head of department, often with the teacher with responsibility for key stage 3, adds another level of moderation through additional sampling during book scrutiny exercises. They take account of a broad range of work across the curriculum. This approach strengthens the moderation process and supports more reliable assessment.

- 30 In most primary and secondary schools, teachers focus appropriately on the level descriptors when marking and assessing learners' profiles. This is supplemented with helpful materials from the regional consortia, such as learning ladders, which describe the skills outlined in each level descriptor their rising stages of development.
- 31 In nearly all schools, teachers understand the need to apply a 'best fit' approach when moderating learner profiles. However, there has been confusion and a lack of clarity on how to apply a 'best fit approach'. In some cases, regional consortia and local authorities have added to the confusion in previous years. For example, schools reported they had received messages to award a level if the pupil 'dips their toe into it', or that '75% of the descriptor should be met'. Despite more helpful support recently, these messages may have led to inconsistency and inflated levels.
- 32 Only a minority of schools refer to standardised materials to check or to support their decisions *during* moderation meetings. In most cases, teachers' experience and their existing knowledge of assessment criteria and of standards are considered sufficient to secure accuracy. This makes the moderation process less accurate.
- 33 In many schools, teachers and senior leaders work in pairs or threes to moderate learner profiles. A minority of schools use a mixed economy approach, varying the methodology of moderation to develop more fully the understanding of assessment and to scrutinise learner profiles more comprehensively. Approaches include round robin reviews of learner profiles, reviews of books, reviews of individual tasks, review by level, random sampling, blank reviews of pupils' books and blind copy levelling by all teachers. This multi-layered view of assessment and moderation leads to greater reliability, as outlined in the report commissioned by the Welsh Government, *An investigation into Key Stages 2 and 3 teacher assessment in Wales (ACER, 2013)*.
- 34 Most schools consider that internal moderation shows that classroom teachers accurately assess levels in most cases. Where there is a dispute over levels, the subject or assessment leaders have the final say. In most cases, schools report that when they discover inaccuracies, around half are too generous and around half are too harsh. However, as the selection of learner profiles lacks sufficient rigour, this reported level of accuracy is unreliable.
- 35 As a result of the limited number of inaccuracies found through moderation, and due to the method of ongoing moderation processes, in only a few cases do schools change levels in response to internal moderation. In many of the cases where inaccuracies are found, schools return to the learner profile in the next meeting, follow up with a wider review of the teacher's assessments or ask for more evidence to award a level.

Schools' standardisation procedures

- 36 The Welsh Government published *English Guidance for Key Stages 2 and 3* (2008a) and *Cymraeg - Canllawiau ar gyfer Cyfnodau Allweddol 2 a 3 (Welsh only)* (2008b) to provide schools with standardised examples of pupils' working at different levels. However, despite noting in the synopsis that the examples reflect features characteristic of the level only, these documents use individual assessment pieces to exemplify levels, a practice now discouraged by the Welsh Government.

- 37 The Welsh Government's *Statutory assessment arrangements for the Foundation Phase and end of Key Stages 2 and 3* (2015d) requires headteachers to 'have in place arrangements by which teachers confirm and maintain a shared understanding of national curriculum standards, using samples of their learners' work to generate a reference set of exemplars (standardisation procedures)(p.8)'.
- 38 In most schools, there is a suitably regular programme of standardisation to support teachers in raising their awareness of National Curriculum standards. The nature of the support for standardisation takes a variety of forms across both sectors. Many schools use previously created cluster or school generated standardised exemplars. A minority of schools support teachers with consortia generated resources or by using the Welsh Government documents. In many schools, materials are distributed or made available in meetings annually. This provides helpful support in the drive for accuracy and consistency.
- 39 In a very few primary schools there has been no history of in-house standardisation. In these schools, regional consortia officers are providing support to develop new processes.
- 40 Many schools have a continuous programme of standardisation throughout the school year that closely mirrors the approach of the moderation process. In many cases, schools use learner profiles that include whole books, whereas a minority still rely on only agreeing the levels of individual pieces of work. Teachers refer closely to National Curriculum level descriptors to make sure that standardisation affirms secure examples of the level. However, too few schools give enough focus on developing a good understanding of where the borderline between levels lies. In addition, in only a minority of cases do teachers cross-reference the standardisation of current work with existing standardised materials.
- 41 As a result, despite the improvements in the procedures for standardisation, in a minority of cases, there remain inaccuracies in awarding of levels.
- 42 In a minority of schools, the process of standardisation has ensured better practice in other areas of school life. For example, a minority of schools report that this process has had a direct impact on effective skills mapping and development of a wider range of writing opportunities to extend pupils of differing abilities.
- 43 In many schools, teachers produce school standardisation materials. This is a helpful process to develop greater understanding of National Curriculum levels and provide more bespoke standardisation materials for schools to use. However, there is no external system to verify the accuracy of these materials.
- 44 Around half of schools use a more extensive range of resources to underpin the standardisation practices. These include the use of cluster moderation profiles, resources from the Learning Wales website, data profiles using commercial assessment programmes or bespoke school tracking systems, to inform this quality assurance. These multi-layered approaches support greater consistency and reliability in standardisation processes.
- 45 A few schools have not reviewed their standardisation materials recently or do not have a process of standardisation in place.

Cluster moderation meetings

Preparation for cluster moderation

- 46 Nearly all schools undertake additional assessment work to prepare learner profiles for the cluster moderation meeting. This preparatory work takes a variety of forms. Nearly all schools complete 'pupil commentaries' to explain why they awarded the level, and to supplement the moderation process. Many schools complete more extensive labelling of learners' profiles, identifying the evidence for each of the elements of the level descriptor. While this is a useful process in developing assessment that is more reliable and in improving teachers' understanding of the criteria for levels, this practice is not sustainable for all profiles, and is an unnecessary task for cluster moderation.
- 47 A few schools pre-select the profiles at the start of the year. This allows for a clear and consistent focus for the school or department and allows for the close annotation and labelling of pupils' books to show how a level has been achieved throughout the year. However, this approach runs the danger of pre-determining pupils' levels or limiting the range of the internal moderation. Many schools mitigate against this by preselecting a suitable range of books, though this adds significantly to the marking workload.
- 48 A minority of clusters undertake earlier 'mock moderation' meetings to inform the final cluster moderation meetings. These meetings allow primary and secondary colleagues to discuss issues of accuracy and evidence together before the moderation meeting. They develop a shared understanding of levels, and reduce the problems faced in having sufficient evidence to moderate effectively. However, in most cases, the same learner profiles are brought to the final cluster moderation meeting and this simply confirms levels.
- 49 This variation of approach creates significant inconsistency in the way cluster moderation informs schools on the accuracy of their final levels.

The evidence considered in cluster meetings

- 50 Nearly all primary schools bring two learner profiles at level 4 and at level 5, and nearly all secondary schools bring two profiles at level 5 and at level 6. In most cases, schools bring work illustrating either secure or higher-end versions of levels to the cluster moderation meetings observed. This contributes to high degree of agreement in cluster moderation meetings. Also, it helps to develop an improved understanding for teachers as to what constitutes a clear level.
- 51 In most cases, cluster moderation meetings do not review learner profiles at the lower borderline of the level. This means that cluster groups do not provide robust challenge to ensure that borderline levels are accurate.
- 52 Many schools bring learner profiles that include pupils' whole books. This is very helpful in providing a suitable context and more holistic evidence for moderators to discuss. However, in a minority of cluster meetings observed, teachers only

reviewed two or three individual tasks for each strand (reading, writing and oracy) rather than whole books. In these cases, teachers rarely refer to other work in pupils' books.

- 53 The coverage of the three strands for English and Welsh taken to cluster moderation reflects the range found in the evidence for in-school moderation. A majority of schools bring a range of recorded oracy evidence to enable moderators to judge this strand. However, in a minority of cases, cluster meetings do not have a sufficient range of evidence to verify oracy levels. In a few cases, moderation is hampered by a lack of context such as the limited pupil commentary or the lack of stimulus material, especially for reading. In most cases when there are gaps in evidence, clusters still agree levels. This is not robust enough and tends to lead to inflation of levels
- 54 In a few clusters, the local authority has worked with schools to provide a common scheme of work for Welsh second language in primary schools. This has ensured more consistency in the evidence base for the cluster moderation process.

The discussion of profiles in cluster meetings

- 55 Almost all schools ensure that appropriate staff, including senior leaders, attend cluster moderation meetings. Regional consortia or local authority officers attend a majority of meetings.
- 56 In many clusters, the chair provides clear guidance on the procedures for the meeting. In many cases, meetings are well informed by training courses or updates on new arrangements from the regional consortia or by the cluster's ongoing development. In these cases, this leads to robust and well-run meetings. Pairs of teachers from different schools review each school's learner profiles. They review the evidence with a clear and consistent focus on finding evidence to support the skills prescribed in the level descriptor. In these meetings, cluster leads write a useful commentary explaining the judgement of the moderators.
- 57 In a minority of cluster meetings, the procedures for moderation are less robust. In these meetings, cluster members do not approach the task of moderation in a systematic manner and meetings lack rigour. For example, in a few clusters, there is a whole-group discussion of learner profiles and a continuous passing around of the evidence to all colleagues to discuss a number of learner profiles at the same time. This leads to less robust challenge of school's learner profiles.
- 58 In nearly all cluster meetings, teachers focus well on reviewing learner profiles by examining the evidence in pupils' work against the level descriptors. When discussing the evidence, most teachers refer closely to National Curriculum levels. Most cluster members take suitable account of the level of support and the drafting process, where provided, when discussing the accuracy of levels.
- 59 In many cases, discussions are open and professional, and group members offer appropriate challenge. However, in a minority of meetings, the inclusion of the school that brought the learner profile in these discussions inhibits the independence of scrutiny. In a few cases, it results in justification of the levels awarded rather than moderation.

- 60 In most cluster meetings, the cluster lead provides exemplification materials. However, these materials are rarely used during cluster moderation meetings.
- 61 In many of the cluster meetings, the judgements made on oracy are not as strongly supported as the judgements for reading and writing. There are three main causes of this. There is insufficient evidence to moderate accurately. The recordings are too poor in quality to judge pupils' performance accurately. There is uncertainty over how far to take into account the use of scripts, reading or the inability to see the non-verbal communication cues where the evidence was aural not visual. This problem is particularly prevalent in Welsh second language, especially for schools who are delivering Welsh GCSE courses during Year 9. Secondary schools prepare the evidence in line with GCSE requirements and therefore produce aural recordings. This is in contrast with the evidence usually produced by primary schools in these cases.
- 62 In most cases, cluster moderation agreed with school's judgements on the levels of learner profiles. In the few cases where levels are changed, around half are moved up and around half are moved down.
- 63 When cluster meetings challenge the judgement on learner profiles, many open the discussion to the whole group to review further and to reach an agreed position. In nearly all cases, this leads to the school accepting the changes for the individual learner profile. The cluster lead makes a note of the changes in the report to the school, and schools are required to review the awarding of levels for non-sampled levels.

Schools' actions in response to cluster moderation

- 64 The National Curriculum (Moderation of Assessment Arrangements for the Second and Third Key Stages) (Wales) Order 2015 requires headteachers to, 'review and if necessary revise any preliminary determination made during the current school year to take account of the determinations and decisions of the school moderation cluster group in relation to the standardisation and moderation of the teacher assessment of pupils for the current school year, (5(b))'. In addition, it requires them to 'implement the determinations and decisions of the school moderation cluster group in relation to the standardisation and moderation of the teacher assessment of pupils during the school year immediately proceeding the current school year, (5(c))'.
- 65 After moderation, each cluster writes a report summarising the outcome of moderation and the actions schools should take as a result of moderation. The majority of cluster reports highlight key areas for improvement. However, actions for reviewing learner profiles or elements of a strand within a subject area appear in only a few cluster reports. Most reports focus only on procedural matters, such as the collation of evidence, or the inclusion of sources for reading assessments.
- 66 Most schools have clear processes to report the main messages from cluster moderation to their departments and teachers. In a few cases, schools and clusters make use of Hwb to share reports and exemplar materials. In many cases, this leads to suitable developments in practice.

- 67 In one cluster group visited, when profiles are not agreed, a further meeting takes place to look more extensively at learners' profiles. Schools in this cluster consider the levels determined in the cluster meeting as automatically binding. This triggers a full review of learner profiles in the school with appropriate changes made to similar learner profiles. In one instance, this resulted in an 11 percentage point reduction in the award of level 5 for a core subject in the secondary school. This is a robust form of cluster moderation.
- 68 However, most schools do not have the same level of rigour in applying the findings of cluster moderation. While many schools would alter the level of individual learner profiles if directed by the cluster, the extent of agreement found in cluster moderation meetings limits the need for further action. Schools do not routinely review and amend the levels of learner profiles for the rest of the cohort.
- 69 A minority of clusters hold their moderation meeting a full term before the levels are due to be finalised with the Welsh Government. In these cases, most schools use moderation to support assessment and allow pupils to respond to any deficits in their learner profile.
- 70 Most schools find the cluster moderation process very helpful, with many describing their cluster as a 'critical friend'. It has become a more rigorous and stronger process. It has begun to improve the trust in levels between the sectors, although a minority of secondary schools remain concerned about inflated levels at transition.
- 71 For most schools, cluster moderation has improved assessment practice. In a majority of clusters, secondary schools offer strong support in developing the confidence of primary colleagues in their assessment of Welsh second language and higher levels. A minority of secondary schools have increased their expectations and strengthened their provision as a result of the understanding they have gained as to the level pupils achieve in primary school, especially in Welsh second language.
- 72 In a minority of clusters, the cluster moderation meeting is part of a strong, developmental working relationship involving tracking the progress of pupils as they make the transition across schools.
- 73 Nearly all secondary schools use the assessment information from primary schools as part of their arrangements for target setting and intervention programmes as pupils make the transition from primary into secondary schools.
- 74 A majority of schools use the moderation process to inform and develop teaching and learning practices, as well as to identify or develop opportunities for more able pupils. Nearly all schools find the process helpful in ensuring that teachers have a good grasp of individual pupils' abilities and to inform how pupils can make progress.
- 75 A minority of school leaders expressed concerns about the impact that accountability pressures have on the validity and accuracy of teacher assessment. In particular, they are concerned that accountability linked to teacher assessment leads to inflated levels, which do not directly relate to a genuine improvement in standards.

The role of consortia and local authority

- 76 The National Curriculum (Moderation of Assessment Arrangements for the Second and Third Key Stages) (Wales) Order 2015 states, 'The chief education officer of each local authority which has a maintained school in its area that is a school member of the school moderation cluster group, or that person's representative, may attend any meeting of a school moderation cluster group for the purpose of seeking to ensure consistency in school moderation and standardisation of teacher assessment practices across all school moderation cluster groups in that local authority's area, (6(3)).'
- 77 Most schools are well supported in the procedures for moderation and standardisation by regional consortia. For example, regional consortia have designed useful assessment frameworks and resources, developed opportunities to share practice across schools and assisted in the administration of moderation processes.
- 78 However, the role of local authority and regional consortia representatives in the moderation of standards remains unclear regarding their role in ensuring consistency of levels across clusters. The requirement that they 'may' attend, produces an inconsistency in their attendance at cluster meetings. Most agree that their representatives at moderation meetings are observers of the moderation process. In a very few cases, local authority or regional consortia representatives take an active or expert role in guiding decisions on the final level of learner profiles. However, a few schools assume that the local authority or regional consortia's presence in cluster moderation meetings verifies the levels agreed.

Appendix: Evidence base

The findings and recommendations in this report draw on visits to eight cluster moderation meetings, nine primary schools and seven secondary schools. The sample draws on practices from all four regional consortia and takes account of socio economic background, size of school and linguistic contexts. Inspectors also considered the range of guidance published by the Welsh Government. In addition, inspectors considered recent Estyn inspection reports for primary and secondary sectors.

In cluster meetings, inspectors:

- observed the discussions undertaken by teachers and senior leaders
- reviewed the evidence provided in the learner profiles and other documentation used to support moderation
- discussed the procedures with the consortia representative, cluster leads and teachers when available

In school visits, inspectors:

- met with senior leaders, middle leaders and teachers to discuss the school's processes for moderating and standardising pupils' work
- reviewed a sample of learners' profiles

List of clusters and schools visited

Cluster	Schools
Cathays High, Cardiff	Cathays High, Cardiff Gladstone Primary, Cardiff
Maesydderwen, Powys	Maesydderwen, Powys Ysgol Y Cribarth, Powys
Newport High, Newport	Newport High, Newport Crindau Primary, Newport
Pen Y Dre, Merthyr Tydfil	Pen Y Dre, Merthyr Tydfil Dowlais Primary, Merthyr Tydfil
Tredegar Comprehensive, Blaenau Gwent	Tredegar Comprehensive, Blaenau Gwent Georgetown Primary, Blaenau Gwent
Ysgol Ardudwy, Gwynedd	Ysgol Cefn Coch, Gwynedd Ysgol Y Traeth, Gwynedd
Ysgol Bryn Elian, Conwy	Ysgol Bryn Elian, Conwy Ysgol Iau Hen Golwyn, Conwy
Ysgol Gyfun Gwyr, Swansea	Ysgol Gyfun Gwyr, Swansea Ysgol Gynradd Gymraeg, Pontybrenin

Glossary

Accuracy	The accuracy of levels relates to how the characteristics of the learners' profiles reflects the characteristics in the National Curriculum level descriptors.
Best fit	'Best fit' judgements require teachers to allocate the level descriptor that most closely matches a pupil's work.
Consistency	Consistency is considered in relation to the similarity of the approach taken within a school, across schools and across clusters. This may relate to the processes of standardisation and moderation, the evidence base and the awarding of levels for learners' profiles.
Cluster	A secondary school and its main partner primary schools
Learner profile	All available evidence from an individual learner that demonstrates their understanding and independent use of skills across a range of contexts
Moderation	The process of reaching an agreement (internally within the school, or externally across the cluster) on the best-fit level that a learner is working at
Reliability	The reliability of the levelling process will depend on the range of the evidence base and the thoroughness of the moderation and standardisation processes. It is a measure of whether the learner profile includes the range and challenge of work to justify the awarding of a level. Also, it is a measure of whether there is sufficient information to make an accurate judgement on a level (for example, the context in which the learner completed the work, the level of support received etc.).
Secure level	This describes work that is firmly within a particular National Curriculum level.
Standardisation	The process whereby samples of previously completed work are used to enable teachers to reach an agreement on levels of attainment by confirming a shared understanding of the characteristics of a level
Verification	The Welsh Government has introduced a national programme, 'Securing Teacher Assessment Programme of External Verification', commonly referred to as EV, to externally verify the levels awarded by schools and confirmed in cluster groups. This is achieved through sampling.

References

Australian Council for Educational Research (2013) *An investigation into Key Stages 2 and 3 teacher assessment in Wales*. Cardiff: Welsh Government. [Online].

Available from: <http://gov.wales/docs/caecd/research/130718-investigation-key-stages-2-3-teacher-assessment.pdf> [Accessed 4 August 2016].

Estyn (2010) *Evaluation of the arrangements to assure the consistency of teacher assessment in the core subjects at key stage 2 and key stage 3*. Cardiff: Estyn. [Online]. Available from:

<https://www.estyn.gov.wales/sites/default/files/documents/Evaluation%20of%20the%20arrangements%20to%20assure%20the%20consistency%20of%20teacher%20assessment%20in%20the%20core%20subjects%20at%20KS2%20%26%20KS3%20-%20March%202010.pdf> [Accessed 4 August 2016].

National Assembly for Wales (2015) *The National Curriculum (Moderation of Assessment Arrangements for the Second and Third Key Stages) (Wales) Order 2015*. SI 1309: 113. London: The Stationery Office Ltd. [Online]. Available from:

<http://www.legislation.gov.uk/wsi/2015/1309/article/5/made> [Accessed 4 August 2016].

Welsh Assembly Government, Department for Children, Education, Lifelong Learning and Skills (2008a) *English: Guidance for Key Stages 2 and 3*. Cardiff: Welsh Assembly Government. [Online]. Available from:

<http://learning.gov.wales/docs/learningwales/publications/130508-english-guidance-en.pdf> [Accessed 4 August 2016].

Welsh Assembly Government, Department for Children, Education, Lifelong Learning and Skills (2008b) *Cymraeg - Canllawiau ar gyfer Cyfnodau Allweddol 2 a 3 (Welsh only)*. Cardiff: Welsh Assembly Government. [Online]. Available from:

<http://learning.gov.wales/resources/browse-all/welsh-guidance-key-stages-2-3/?cr=7&ts=1&lang=en> [Accessed 4 August 2016].

Welsh Assembly Government, Department for Children, Education, Lifelong Learning and Skills (2010) *Making the most of assessment 7–14*. 034/2010. Cardiff: Welsh Assembly Government. [Online]. Available from:

<http://learning.gov.wales/docs/learningwales/publications/130425-making-the-most-en.pdf> [Accessed 4 August 2016].

Welsh Government (2015a) *Programme of Study for English: Key Stages 2–4*. Cardiff: Welsh Government. [Online]. Available from:

<http://learning.gov.wales/docs/learningwales/publications/150717-nc-english-en-v2.pdf> [Accessed 4 August 2016].

Welsh Government (2015b) *Programme of Study for Welsh: Key Stages 2–4*. Cardiff: Welsh Government. [Online]. Available from:

<http://learning.gov.wales/docs/learningwales/publications/150717-nc-welsh-en-v2.pdf> [Accessed 4 August 2016].

Welsh Government, Department for Education and Skills (2015c) *Welsh Second Language in the National Curriculum for Wales*. Cardiff: Welsh Government. [Online]. Available from: <http://learning.gov.wales/docs/learningwales/publications/150707-welsh-second-language-cy.pdf> [Accessed 4 August 2016].

Welsh Government, Curriculum Division (2015d) *Statutory assessment arrangements for the Foundation Phase and end of Key Stages 2 and 3*. 185/2015. Cardiff: Welsh Government. [Online]. Available from: <http://learning.gov.wales/docs/learningwales/publications/141030-statutory-assessment-arrangements-en-v2.pdf> [Accessed 4 August 2016].