

NAA response to the 2007 QCA Regulation and Standards report on level setting

December 2007

QCA/07/3420

Contents

Summary of the National Assessment Agency response.....	3
Section 1: Draft level setting meetings	4
Section 2: Script scrutiny meetings.....	5
Section 3: Final level setting meetings	6
Conclusions and implications for future regulation	7
Issues requiring action by the National Assessment Agency.....	7
Improvements and minor issues	8

Summary of the National Assessment Agency response

The National Assessment Agency (NAA) is pleased to note the findings of the Qualifications and Curriculum Authority (QCA), the regulator of external qualifications and national curriculum assessments, in relation to level setting in 2007. QCA has noted that general improvement was evident at all stages of the process this year, with a great deal of good practice observed by monitors. The NAA is content to accept most of the issues identified by the regulator and to detail the actions it has put in place to address them. However, the NAA has concerns about the issues raised by QCA in relation to script scrutiny and the use of the level setting exercise data.

Section 1: Draft level setting meetings

Compliance issues identified by QCA	NAA response
<p>'...paragraph 293 of the 2007 <i>Code of practice</i>¹ states that the draft level setting meeting "must consider the thresholds and associated ranges in turn, starting with the target level for each key stage (and tier, where relevant), then progressing onto the higher levels before finishing with the lower levels". Yet in one meeting (key stage 3 mathematics) the level thresholds were dealt with via a 'lowest to highest' approach that contradicted the requirement to start with the target level (level 5 in this instance). This is a minor issue and seemed unlikely to affect the security of the meeting's outcomes.'</p>	<p>Accepted.</p> <p>While we appreciate that this is a minor issue with no detrimental consequences, the NAA endeavours to ensure that the correct procedures are followed at all times and will address this in the future.</p>
Other issues identified by QCA	NAA response
None identified.	N/A

¹ *National curriculum assessments: Code of practice* (QCA/07/2828) was published in January 2007 and is available on the QCA website: www.qca.org.uk.

Section 2: Script scrutiny meetings

Compliance issues identified by QCA	NAA response
None identified.	N/A
Other issues identified by QCA	NAA response
<p>'Some script scrutiny meetings set very narrow threshold zones of one or two marks, which were, on occasion, exactly the same as the final threshold recommendations. While this might occasionally occur in any key stage / subject, this year there were some key stage / subjects where it seemed to occur for every threshold. Scrutineers are asked to establish both a zone and a threshold, suggesting that it is not a requirement to have threshold zones that exactly match the thresholds.'</p>	<p>Not accepted.</p> <p>Scrutineers must base their decisions on the evidence in the scripts, even if this means that the zones are narrow. It would be wrong for the NAA to stipulate wider zones.</p>

Section 3: Final level setting meetings

Compliance issues identified by QCA	NAA response
None identified.	N/A
Other issues identified by QCA	NAA response
<p>'...national curriculum assessments observers still raised some concerns about what the exact roles and responsibilities of some participants were, particularly when participants seemed to be advocating certain pieces of evidence. It often seemed, for example, that a representative from the NAA was present to support the test development agency without a similar advocate being present to support the marking programme leader.'</p>	<p>Not accepted.</p> <p>The role of the NAA is not to support the test development agency or the marking programme leader, and so it would be inappropriate for a member of the NAA to support either. NAA staff argue in support of whichever strand of evidence they judge to be the stronger.</p>
<p>'In some subjects there are good historical reasons for favouring equating evidence. There are also other subjects where this is the first time for several years that equating evidence has been this secure, and where advocacy of script scrutiny would be more in line with past practice.'</p>	<p>Not accepted.</p> <p>The final level setting meeting enables judgements to be made based on the strength of the evidence, and while one measure of the strength of evidence will be what has happened historically, this cannot be the overriding consideration. Further, as noted in the QCA report, equating evidence that had been less secure in previous years was improved in 2007. The fact that one form of evidence has been advocated in the past is no reason for the NAA to continue to advocate it, especially in the light of improvements in other forms.</p>

Conclusions and implications for future regulation

The NAA welcomes the conclusion that the results of the level setting process can be taken as secure and that standards have been maintained.

Issues requiring action by the National Assessment Agency

Issues identified by QCA	NAA response
<p>'...past experience indicates that these two strands [script scrutiny and equating evidence] cannot be relied upon to align every year, and while the level setting exercise data should not be treated in the same way as the script scrutiny or equating evidence, it should provide potentially useful information on the impact of the other types of evidence.'</p>	<p>Accepted.</p> <p>The NAA accepts that level setting exercise data is used to predict the impact of the decisions at final level setting meetings and that a larger sample, which is more representative of the population, would be beneficial. There are practical reasons that dictate the early timing of the level setting meetings, which results in level setting exercise data being a relatively small sample. However, this will be changed in 2008, as the level setting meetings will be held later in the process.</p> <p>This will mean that the meeting will have access to a substantial sample of 50 to 80 per cent of live data. This is one of several improvements being made to the marking process.</p> <p>However, the NAA maintains its position that level setting exercise data does not provide compelling evidence that can contribute to the level setting decision. Its clear and stated purpose is to inform attendees at the level setting meeting of potential outcomes of decisions that are based on other forms of evidence. The function of the meeting is to maintain standards, not to maintain equivalent statistical performance year on year.</p>

Improvements and minor issues

Issues identified by QCA	NAA response
<p>‘...next year there is a transfer of responsibilities to a different test operations agency and this will need to be carefully managed by the NAA.’</p>	<p>Accepted.</p> <p>The NAA has strong programme management expertise and processes in place to manage this transfer.</p>
<p>The final results from script scrutiny do seem to show an increasing precision, which is a perception that is not always supported by scrutineers’ individual results. Threshold zones of two marks do not always seem adequately to account for the differences in scrutineers’ judgements. For example, on one occasion a threshold zone was chosen which did not include a mark point at which six scrutineers had observed threshold performance. In such instances there does not seem to be a reason why such small zones have been identified and it might be appropriate if wider zones, which more accurately account for the judgements of all the scrutineers, were agreed.’</p>	<p>Accepted.</p> <p>The NAA accepts that in some meetings the script scrutiny panel automatically indicates the threshold zone as being from the highest mark on which threshold performance has been observed by at least one scrutineer to the lowest mark on which threshold performance has been observed by at least one scrutineer. In other meetings a decision is taken to set the threshold zone within this wider range, which can lead to instances of observed threshold performance outside the threshold zone.</p> <p>The NAA will amend its procedures to ensure a more consistent approach in future, with the majority of threshold performance included in the threshold zone.</p>
<p>‘Some minor concerns were raised relating to roles and procedures at these [final level setting] meetings. However, these are not significant enough to affect the final outcomes and will be assuaged by a review of the NAA’s document on formal procedures for the meetings.’</p>	<p>Not accepted.</p> <p>As detailed above.</p>