

Consultation on Accessibility at Source for National Curriculum Tests and General and Vocational Qualifications



Date: June 2011

Product code: Ofqual/11/4869

Contents

Contents	2
Introduction.....	5
Background	5
About this consultation	6
Section 1: Principles for language modification at source for National Curriculum tests.....	7
Section 2: National Curriculum tests accessibility: quality assurance at source	8
Introduction	8
Proposal 1: Special Education Teacher Review Panels.....	9
Current position.....	9
Proposal.....	9
Responding to Proposal 1	10
Proposal 2: Pupil Review Panels	11
Current position.....	11
Proposal.....	11
Responding to Proposal 2.....	12
Proposal 3: Differential Item Functioning Analysis	13
Current position.....	13
Proposal.....	15
Responding to Proposal 3.....	16
Section 3: Principles for language modification at source for general and vocational qualifications	17
Section 4: General and vocational qualifications accessibility: quality assurance at source	18

Introduction	18
Responding to Proposals 4 to 6	18
Proposal 4: National Voluntary Training Programme	19
Responding to Proposal 4	19
Proposal 5: Best Practice Document.....	19
Responding to Proposal 5	20
Proposal 6: Data Collection and Analyses	20
Responding to Proposal 6	22
Section 5: General and Vocational Qualifications - Using Computer Adaptive tests	23
Responding.....	24
Annex 1	25
Accessibility at Source for National Curriculum Tests and General and Vocational Qualifications - Questions on Proposals 1 to 6	25
Proposal 1: Special Education Teacher Review Panels.....	25
Proposal 2: Pupil Review Panels	28
Proposal 3: Differential Item Functioning (DIF) Analysis	30
Proposal 4: National Voluntary Training Programme	33
Proposal 5: Best Practice Document.....	33
Proposal 6: Data Collection and Analyses	35
Annex 2	38
Principles for language modification at source for National Curriculum tests	38
Topic 1: The principles of language accessibility for test developers in charge of writing test items for National Curriculum Assessments.....	38
Topic 2: What affects the readability of test items? What are the characteristics of the target readers of these items?	46

Annex 3	51
Principles for language modification at source for General and Vocational qualifications	51
A. Current arrangements for language modification	51
B. Defining the linguistic needs of different candidates	54
C. Principles for language accessibility	56
D. Building expertise in addressing language issues	57
Annex 4	60
General and Vocational Qualifications - Using Computer Adaptive tests questionnaire.....	60

Introduction

Background

The Office of Qualifications and Examinations Regulation (Ofqual) is the regulator of qualifications (other than degrees), examinations and assessments in England, and of vocational qualifications in Northern Ireland. We aim to make sure that learners get the results their work merits and that the qualifications they receive count, both now and in the future.

In 2010, Awarding Organisations (AOs), Qualifications and Curriculum Development Agency (QCDA) and organisations representing disabled learners indicated they would welcome guidance on the language accessibility of, and modifications to, assessment papers. It was suggested that Ofqual should take a lead in providing such guidance.

We have in the past contributed to guidance such as *Fair access by design* (2010)¹. *Fair access by design* focuses almost exclusively on sentence difficulty and legibility aspects of assessments. New research suggests, however, that readability problems can be caused by other textual factors. Factors to do with textual cohesion and coherence have been especially neglected in the past. We are therefore consulting on guidance on what should represent good practice in this area.

Subject to the views expressed in the consultation, we will revise and publish the guidance. This guidance must be considered in the context of the form of the assessment and the knowledge, skills and understanding being assessed.

We propose that the guidance could be used by test development agencies (for National Curriculum Assessments) and by awarding organisations (for qualifications). However, its use would be optional. We recognise that some of the approaches explored in the consultation, such as Computer Adaptive Tests, would not be suitable or desirable for all assessments.

The Guidance might also be useful for those who train subject experts and others who prepare the assessments.

The present consultation is a result of work on the accessibility of assessments carried out by Ofqual since 2009. In 2009, Ofqual carried out a research project on the *Current Practices and operational aspects of paper modification in England*

¹ http://www.rewardinglearning.org.uk/docs/regulation/fair_access_by_design.pdf

2009/10² (2009) leading to a publication on *The Future of language accessibility* (2010) in the UK.³

About this consultation

This consultation on language accessibility considers National Curriculum Assessment (NCA) tests in sections 1 and 2 and qualifications in sections 3, 4 and 5.

We wish to reach, and receive responses from as many stakeholder groups as possible. These include, but are not limited to: test development agencies (TDAs), principal examiners from awarding organisations who write and quality assure assessment items; awarding organisations, learners and organisations and individuals who represent their interests, speech and language therapists, diversity and inclusion specialists, and other stakeholders representing the diverse equality strands. Our aim is to build up strong relationships with our stakeholders which will last after this consultation has ended.

² <http://www.ofqual.gov.uk/files/2009-10-paper-modification-in-england-2009-10.pdf>

³ <http://www.ofqual.gov.uk/files/2010-10-29-modified-question-papers.pdf?Itemid=142>

Section 1: Principles for language modification at source for National Curriculum tests

Recent research has shown that there are a number of factors which affect a learner's ability to read texts. These include:

- Linguistic difficulties: word difficulty and sentence difficulty
- Cohesion and coherence
- Conceptual difficulty
- Legibility and print issues
- Text organisation.

The *Fair access by design* (2010) document focuses almost exclusively on the sentence difficulty and readability aspects in the above list. Research suggests, however, that readability problems can be caused by other textual factors. Factors to do with textual cohesion and coherence have been especially neglected in the past.

Your views are sought on two topics through a questionnaire:

- The principles of language accessibility for test developers in charge of writing test items for National Curriculum Assessments
- What affects the readability of test items? What are the characteristics of the target readers of these items?

The questionnaire can be completed online⁴ or at Annex 2 of this document. The PDF version⁵ of this document can be obtained online.

If you would like to read more about this subject before answering these questions, please look at the Ofqual commissioned report by language experts David Wray and Dahlia Janan⁶.

⁴ <http://surveys.ofqual.gov.uk/s3/Principles-for-language-modification-at-source-for-National-Curriculum-tests-Annex-2>

⁵ <http://www.ofqual.gov.uk/files/2011-06-16-accessibility-consultation.pdf>

⁶ <http://www.ofqual.gov.uk/files/2011-06-15-principles-of-language-accessibility-test-developers-in-charge-of-writing-test-items.pdf>

Section 2: National Curriculum tests accessibility: quality assurance at source

Introduction

Our aim for this part of the consultation is to ask for comments on the ways in which National Curriculum Test Development Agencies might be able to evaluate the accessibility of the key stage tests that they produce. This self-evaluation guidance is designed to check that all the questions in a set of tests are accessible, without further modification, to the widest possible a range of pupils.

There has always been a strong emphasis on the development of accessible questions for the national tests – stronger than for some other assessments such as GCSEs or GCEs. In part this is because the key stage tests were originally conceived as school tests, and were designed for pupils who were younger than those taking GCSEs. Furthermore, key stage tests have always been statutory and intended for all pupils, so it was never possible to argue that a pupil who could not access them should not be doing them. For this reason, fair access to the key stage tests had to be ensured as far as possible.

There are three proposals in this section we are seeking your response to:

- Proposal 1: Special Education Teacher Review Panels
- Proposal 2: Pupil Review Panels
- Proposal 3: Differential Item Functioning (DIF) analyses.

Proposal 1: Special Education Teacher Review Panels

Current position

The evaluation Test Review which the Qualifications and Curriculum Development Agency (QCDA) conducts each year is very valuable in offering an overall picture of the way tests operated in practice when taken by pupils in the classroom.

Proposal

The evaluation Test Review process might be extended to include a particular focus on issues relating to the accessibility of the materials for pupils with a range of special educational or assessment needs. This might involve the following steps:

- a select group of about twelve teachers who work with pupils with special educational or assessment needs who are due to take the key stage tests are invited to take part in a review
- immediately after their pupils have taken the tests, the teacher will talk to a small group of up to four pupils that offer a representative sample of the types of special educational or assessment needs of pupils in their year group.

A sheet with four or five question prompts is then made available to help teachers to structure their discussions with their pupils.

- The teachers will consider their pupils' comments, and decide whether they indicate any issues relating to the accessibility of the questions.
- The teachers are then invited to a meeting with the test developers that has an Ofqual observer in attendance to report on their pupils' experiences with the tests.

Question prompts that might be included on the prompt sheet include:

- What was your favourite question? Why did you like it?
- What was your least favourite question? Why did you not like it?
- Did you get stuck on any of the questions? Why were you stuck?
- Were there any questions that you didn't understand?
- What could the people who write the tests do to make them better?

The meeting of the *Special Education Teacher Review Panel* to discuss pupils' experiences in the live tests will have to be set up and funded under the

contractual arrangements between commissioners and test developers. This panel approach will provide highly focused qualitative evidence on the accessibility of the tests for pupils with a range of special education and assessment needs.

If you would like to read more about this before answering these questions, please look at the Ofqual commissioned report by language experts Jamal Abedi and Tandi Clausen-May.⁷

Responding to Proposal 1

Please respond to questions 1 to 8 online⁸ or at Annex 1 of this document.

⁷ <http://www.ofqual.gov.uk/files/2011-06-15-the-accessibility-of-the-national-curriculum-assessments.pdf>

⁸ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

Proposal 2: Pupil Review Panels

Current position

All key stage test questions are trialled when they are being developed, and test developers routinely scrutinise pre-test scripts to see how pupils have responded to individual items and to identify any problems of accessibility. In addition, test developers may undertake small-scale informal trialling of the questions.

For a small-scale informal development trial, pupils are normally observed as they work through some draft questions, and are then invited to comment on these. This strategy may be used in the early stages of the development of the questions, and it can provide useful information about their accessibility for a range of test takers.

Proposal

While teachers and expert reviewers may offer valuable insight into the nature of the questions, the people who have the greatest involvement and interest in the tests are the pupils themselves. Input from the pupils might now be extended to evaluate and assure the accessibility of the final tests.

This new approach, which may be described as 'small-scale discursive trialling', might involve the following steps:

- a researcher would work with a Pupil Review Panel of just four pupils with special educational or assessment needs who had taken the current year's tests and were working at an appropriate level
- the researcher would invite the pupils to think about and write their responses to a single question, and note any evidence relating to their methods or reasoning
- the pupils would then discuss the question in detail as the researcher probed their reasoning, checking to see whether they understood what it was asking even if the problem posed was beyond their current knowledge or understanding
- time permitting, the panel would go on to discuss several more questions in detail in the same way.

This discursive trialling would focus on certain questions that would be identified as possibly presenting particular issues of accessibility to the pupils on the *Pupil Review Panel*.

The specific prompting questions that the researcher would ask the pupils would depend on the situation, on the nature of the test item being discussed, and on the researcher's observations of the pupils' responses as they worked through it. The researcher would be likely to address specific prompts to individual pupils.

These prompts might include:

- What is this test question asking us to do?
- How did you go about thinking about the question? Can you explain what you did?
- I see you wrote Can you explain why you wrote that?
- Chris, I see you got the same answer as Ali. Can you explain how you did it? Did you do it in the same way as Ali?
- Josh has explained how he did this question. Did anyone do it in a different way?

Responding to Proposal 2

Please respond to questions 9 to 12 online⁹ or at Annex 1 of this document.

⁹ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

Proposal 3: Differential Item Functioning Analysis

Current position

A well-established statistical procedure that is often used to identify individual questions that may be biased against particular groups of test-takers is the use of Differential Item Functioning (DIF) analysis.

In a DIF analysis the performance on each question of all the members of one group of test-takers is compared with the performance of the members of another group. For example, in a gender-based DIF analysis the results for girls and for boys might be compared for each question in a test. This can help test developers to identify particular items on which members of one of these groups perform in a way that does not match their overall performance on the test. So, for instance, a gender-based DIF analysis might reveal that girls who perform well on the test overall tend to perform less well on a specific question or part of a question. In this case the question should be reviewed to check that it does not have some hidden barriers to accessibility for girls. For example, it might be found that a question set in the context of cricket or motor racing was not as readily accessible to girls as to boys, and so it would be biased against girls and should be amended or rejected. This is a simple and straightforward example which helps to explain the principles behind the use of DIF analysis.

It is important to note that a DIF analysis compares the performance of test-takers in the focus group on each individual question in the test in comparison with their overall performance in that particular test. It might show that question 4, say, is biased for or against the members of a particular group in comparison to the rest of the other questions. But a DIF analysis cannot show whether the whole test is biased against the members of the group. So, for example, if all the members of a particular group tend to do worse than other pupils on every question in the test then the mean score of the group will be lower than that of the mainstream pupils, but the DIF analysis itself will not suggest that there are any issues. For this reason DIF analysis is not useful in identifying problems with a test as a whole, only with specific questions within the test. A DIF analysis cannot say 'This test is biased against pupils in this group'. It can only say 'Questions x, y and z are biased against these pupils'.

National Assessment development agencies have extensive experience of carrying out some types of DIF analyses during the development phase of their projects. The outcomes of these are used to identify any questions which show bias against particular groups, such as boys or girls, and to guide amendments to these questions. If the sample size allowed, similar DIF analyses could be carried

out to identify any questions that might be biased against pupils from different ethnic groups or with different social backgrounds.

However, for the key stage tests a question that was set in a context that was so obviously likely to be biased as cricket or motor racing (or fashion or ballet) would almost certainly be rejected by the test developers at an early stage, so it would not even reach the first pre-test. Only questions that do not carry any obvious cause of bias are selected and trialled, so when the statistical results show that a question is biased against girls (or against boys) there may be no apparent reason for this, with other very similar questions showing no such bias. For example,

'the following question was trialled for a mathematics test:

Put one number in each gap to make the sentences true.

Example

Multiplying by **2** and then by **6** is the same as multiplying by **12**.

- a) Multiplying by **3** and then by **2** is the same as multiplying by _____.
- b) Multiplying by **4** and then by **6** is the same as multiplying by _____.

The wording, presentation and layout of the two parts of the question are identical. None the less, girls did significantly better than boys (at the one per cent level) in part a), but not in part b). This could have been a random statistical effect, but it still gave the test developers some cause for concern.'

(Clausen-May, 2001, pp31-33).

So the statistical results of a DIF analysis showing which items are biased may, or may not, be useful for the selection of questions for a test.

Furthermore, effective DIF analyses rely upon there being enough test-takers in the category which is the focus of the analysis to allow for statistically robust conclusions to be drawn. Zieky (1993) argues that there must be at least a hundred people in the smaller group, and at least five hundred altogether for DIF analyses to be used at the development phase – and more for evaluation purposes (Zieky, 1993, pg 346).

For a gender DIF analysis the sample size is not likely to be a problem, as there are approximately equal numbers of boys and girls taking any key stage test. Similarly, there may be enough pupils who have English as an additional language (EAL) to offer a viable sub-sample of the total. But where only a relatively small

number of pupils belong to a group there may not be enough to provide an adequate sample on which to base any meaningful statistical conclusions. So while the sub-sample of all pupils with EAL may be large enough, there may be too few pupils from each specific ethnic group to allow test developers to carry out a differential item analysis.

For example, a test development agency might want to use a DIF analysis to compare the performance of pupils from different ethnic backgrounds in order to identify particular items on which, say, Caribbean boys or Chinese girls tend to do particularly badly – or particularly well – in relation to their overall performance on the test. However, there may not be enough pupils from these groups to provide an adequate sub-sample for the analysis. Similarly pupils with specific forms of special educational or assessment need may not be adequately represented to allow for a valid DIF analysis for their particular condition. Furthermore, as has been noted above, if the pupils in the focus group tend to do less well or better than mainstream pupils overall in the test then the DIF analysis will not show this.

It should also be noted that even the pupils with one particular condition or disability, such as hearing impaired pupils or dyslexic pupils, may not form a homogeneous group. There may be so much variation between individuals that classing them together is not useful as it ignores their very significant differences.

Key stage test developers routinely carry out at least gender and English as an Additional Language (EAL) DIF analyses as part of the process of developing draft tests. None the less, to Ofqual's knowledge there has been little published work on this use of DIF analyses.

Ofqual is now considering the possibility of using DIF analyses *after* the tests have been taken, using the results of the live tests to confirm that none of the questions were biased against a particular group of pupils, or, if they were, to identify these in order to guide the development of questions for later tests. The proposed study would look at the functioning of test items in Key Stage 2 tests from 2008-2010 with respect to a range of pupil background factors, including gender, ethnicity, eligibility for free school meals (FSM), special educational needs (SEN), English as an Additional Language (EAL), Income Deprivation Affecting Children Index (IDACI), and school type. The purpose of this study would be to establish for quality control purposes whether any of the questions were, in fact, biased against pupils in any particular group.

Proposal

To allow DIF analyses to be used in a similar way for evaluation purposes in the future, item-level data (the number of marks awarded for each part of each question to each pupil) would need to be collected, along with relevant pupil background data, for at least a sample of the total cohort of test-takers. However,

to make the collection of DIF data worthwhile we must identify the purposes to which any evidence that a particular item may be biased would be put. It would not be possible to amend or remove the item at this stage as the test would already have been taken by the cohort of pupils. Two possible purposes might be:

- to guide future item development by providing test developers with robust data relating to previously developed items;
- to establish a bank of biased items for further research.

Your views are now sought on whether the use of DIF analyses using live data from the released tests, rather than pre-test data, as part of an evaluation procedure would be worthwhile, and if so, which groups should be identified for data collection and analysis and what use should be made of the outcomes.

Responding to Proposal 3

Please respond to questions 13 to 17 online ¹⁰ or at Annex 1 of this document.

¹⁰ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

Section 3: Principles for language modification at source for general and vocational qualifications

In response to the perceived need to ensure that the language of assessments is user-friendly to all, through this questionnaire we are seeking your views on the development of guidance that aims to:

- bring together existing guidance about principles of language accessibility
- propose that such principles are applied during question setting, rather than later in the process
- provide a reference point for training in the specific expertise needed to review assessment questions in collaboration with subject experts and AOs
- clarify for teachers some of the issues in preparing candidates for qualifications and for setting in-house papers.

The new guidance resulting from this consultation must be considered in the context of the form of the assessment and the knowledge, skills and understanding being assessed.

The questionnaire can be completed online¹¹ or at Annex 3 of this document.¹² The questionnaire has four parts:

- Part A. Current arrangements for language modification
- Part B. Defining the linguistic needs of different candidates
- Part C. Principles for language accessibility
- Part D. Building expertise in addressing language issues.

If you would like to read more about this subject before completing the questionnaire, please look at the Ofqual commissioned report by language experts Maureen Jefferson and Janet White.¹³

¹¹ <http://surveys.ofqual.gov.uk/s3/Principles-for-language-modification-at-source-for-National-Curriculum-tests-Annex-2>

¹² <http://www.ofqual.gov.uk/files/2011-06-16-accessibility-consultation.pdf>

¹³ <http://www.ofqual.gov.uk/files/2011-06-15-general-and-vocational-consultation-on-principles-for-language-modification.pdf>

Section 4: General and vocational qualifications accessibility: quality assurance at source

Introduction

The research undertaken for Ofqual indicates that language used in assessments designed for test takers from different test taking groups such as gender, race/ethnicity, socioeconomic status, and disability groups can involve the following modifications.

Change in the *characteristics of the test tasks*

This includes linguistic modification or simplified language, native language or bilingual presentation, as well as other forms of modification such as increase in font size, use of Braille, use of black and white instead of colour diagrams, etc. Other modified papers include A3 modified paper at 24 point bold, Braille papers and tactile diagrams with print labels.

Change in *administration of test tasks*

These include extended time, frequent breaks during the test, individual testing, translation and/or repetition of test items directions, use of sign language, oral reading of written directions and test items, individual or small group administration, separate location, for example.

Change in *response to test tasks*

These include use of a scribe, calculator, spelling aid, use of sign language, response in native language, for example.

To take investigation of the effects of such modifications forward Ofqual is seeking your views on three proposals and associated topics.

If you would like to read more about this subject before answering these questions, please look at the Ofqual commissioned report by language experts Anthony Kunnan and Martin Walker.¹⁴

Responding to Proposals 4 to 6

Please complete the response form online¹⁵ or at Annex 1 of this document.

¹⁴ <http://www.ofqual.gov.uk/files/2011-06-15-proposal-to-the-language-accessibility-working-group.pdf>

Proposal 4: National Voluntary Training Programme

(Please note that Ofqual does not intend to run or fund any national training programmes.)

Based on an understanding of how assessments are used, develop a national voluntary training programme for test designers/senior examiners. The training programme would focus on the effects of using various language forms and constructions on candidates. The training programme should be short enough to be manageable but substantial enough to provide test designers/senior examiners with a clear insight into the ways that language can influence the outcomes of assessment. This could be a one day course in the first instance with a possible follow up day and some form of accreditation.

Responding to Proposal 4

Please respond to question 18 online¹⁶ or at Annex 1 of this document.

Proposal 5: Best Practice Document

Circulate a “best practice document” for a test design process based on accepted procedures. The document needs to:

- (a) be consistently used in Awarding Organisations
- (b) generate agreement on the adoption of strategies to implement item analysis procedures during the preparation stage of each examination.

Content analysis for planning language modifications

Detailed content analysis is necessary to examine test tasks and their suitability for the test takers who are likely to be affected by the test. Content analysis can reveal a number of aspects including unnecessary use of terminology or difficult-to-comprehend language (which applies *to all weak readers*). If such difficult-to-comprehend items or task are used in tests, appropriate language modifications would be needed to make the test more language-accessible.

¹⁵ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

¹⁶ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

Developing a check list for test item review

Checklists have been used on a regular basis by test review panels (of two or three test reviewers). These reviewers check all items or tasks for all aspects of fairness in testing such as offensive or insulting language or topics or unfair penalisation based on topics. Based on such reviews, items or tasks are green-lighted on to be part of a test (if no problems are detected) or are revised or deleted from the test.

Responding to Proposal 5

Please respond to questions 19 to 24 online¹⁷ or at Annex 1 of this document.

Proposal 6: Data Collection and Analyses

Develop a format for AOs to collect and analyse data on previous rounds of assessments. Encourage AOs to use known item response data to inform test developers/senior examiners at the next round of qualification design and preparation.

Language accessibility for whom

There should be a clear sense of which test taking groups are affected due to some tests not being written in accessible language. Research evidence needs to document which groups are affected so that appropriate language modifications can be made to the test tasks, test administration and/or test responses.

Addressing this research topic will help focus on examining language accessibility for these test taking groups.

Differential Item Functioning (DIF) analysis of test performance

DIF analysis is performed on test performance data of sub-groups to examine their data. Such analysis is expected to show whether test takers from different sub-groups of comparable ability levels (by total score of the test) perform differently on certain test items or tasks (or, in other words, whether the items function differentially).

¹⁷ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

Empirical research on interaction hypothesis

The interaction hypothesis is the assumption that test accommodations or modifications will improve test scores for the test takers who need the accommodation but not for those test takers who do not need the accommodation.

Empirical research data needs to be collected from test-taking sub-groups in terms of how they performed with and without the test accommodation in order to undertake the required analyses.

Empirical research on the effects of language modifications on test taking group performance

Specific modifications should be examined in terms of whether the test taking groups performed better with modifications or not. Empirical research on the effects of different language or test modifications for the test taking groups can examine:

- if the modifications helped the test-taking groups to gain higher scores
- if the general candidate population benefit from better-designed assessment.

Validation studies across sub-groups

The main question that needs to be addressed here is whether the scores received through accommodated or modified tests are equivalent in terms of meaning and interpretation to scores received through non-accommodated or modified tests. Examining score comparability between test takers who took an accommodated or modified test and test takers who took a standard unmodified test is not sufficient. Test takers' scores in both types of administrations need to relate to external criteria (such as other grades, admission test scores, etc.). Further, it is necessary to establish that the test accommodations provided did not change the construct that was being measured

It is proposed to set up a double-blind trial¹⁸ of modified and non-modified items using representative candidates from the identified sub-groups. The double blinding would be particularly important due to the effects that examiners could

¹⁸ Double blind trial refers to an especially rigorous way of conducting an experiment usually on human subjects, in an attempt to eliminate subjective bias on the part of both experimental subjects and the experimenters. In a double-blind experiment, neither the individuals nor the researchers know who belongs to the control group and the experimental group.

have on the study if they were to announce the “easy” and “hard” versions of the test. Develop an online model which would assign the different versions of the test without teachers being involved.

Responding to Proposal 6

Please respond to questions 25 to 32 online¹⁹ or at Annex 1 of this document.

¹⁹ <http://surveys.ofqual.gov.uk/s3/Accessibility-at-Source-for-National-Curriculum-Tests-and-General-and-Vocational-Qualifications-Proposals>

Section 5: General and Vocational Qualifications - Using Computer Adaptive tests

This last section of the consultation seeks your views on the practicalities of developing a computer adaptive test with special attention given to methods used to mitigate test and question biases that impact members of different population subgroups. Your views are sought through a brief questionnaire.

A computer adaptive test (CAT) is an assessment administered on a computer that adapts the difficulty level of each question or item to the ability level of the candidate. In computer adaptive testing, the difficulty level of the test items is determined by the ratio of the number of past candidates who answered the item incorrectly to the total number of candidates who viewed the item. An item that many candidates get incorrect is determined to be difficult. An item that many candidates get correct is determined to be easy.

A candidate who answers correctly items that many candidates answer incorrectly will get a higher score than a candidate who answers items that nearly all candidates answer correctly. While this may seem reasonable, it is a departure from the practice of using subject-matter experts to determine the difficulty level of an item. Using computer adaptive testing models, there is no subjective measure of an items difficulty. Difficulty is strictly a statistical parameter.

If you would like to read more about this subject before completing the brief questionnaire, please look at the Ofqual commissioned report by language expert Michael Birdsall.²⁰

The questionnaire can be completed online²¹ or at Annex 4 of this document.²²

²⁰ <http://www.ofqual.gov.uk/files/2011-06-15-implementing-computer-adaptive-testing-to-improve-achievement-opportunities.pdf>

²¹ <http://surveys.ofqual.gov.uk/s3/General-and-Vocational-Qualifications-Using-Computer-Adaptive-tests-questionnaire-Annex-4>

²² <http://www.ofqual.gov.uk/files/2011-06-16-accessibility-consultation.pdf>

Responding

Thank you for taking the time to consider responding to our consultation on the accessibility for tests and qualifications.

You can respond to the proposals raised and questions asked in this document using one of the following methods:

- completing the appropriate online response form access from the Ofqual consultation platform²³
- printing and completing by hand the proposal response form at Annex 1 or questionnaires at annexes 2, 3 and 4 of the consultation document.

Please feel free to respond to as few or as many questions as you wish.

The deadline for receipt of responses is no later than 5pm on **Thursday 15th September 2011**.

Responses not submitted through the online platform can be returned by email to consultations@ofqual.gov.uk or by post to Accessibility Consultation, Ofqual, Spring Place, Herald Avenue, Coventry CV5 6UB.

Once the consultation has closed, Ofqual will collect the data and finalise non-statutory guides on the importance of linguistic accessibility for all National Curriculum tests and General and Vocational qualifications. The guides will support greater uniformity in how test developers and awarding organisations formulate and quality assure the language of test and assessment items, thus avoiding item bias. The guides should be completed by January 2012 and will be published on the Ofqual website.

²³ <http://comment.ofqual.gov.uk/>

Annex 1

Accessibility at Source for National Curriculum Tests and General and Vocational Qualifications - Questions on Proposals 1 to 6

If you would like to read more about proposals 1 to 3 before answering these questions, please look at the Ofqual commissioned report by language experts David Wray and Dahlia Janan.²⁴

Proposal 1: Special Education Teacher Review Panels

1. Please rate your responses to the following statements.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
A group of teachers should be invited to join a Special Education Teacher Review Panel to evaluate each year's tests	[]	[]	[]	[]	[]	[]
Small groups of pupils should be interviewed immediately after they have taken a set of tests	[]	[]	[]	[]	[]	[]
A Question Prompt sheet should be made available to help teachers to structure their discussions with the pupils	[]	[]	[]	[]	[]	[]

2. Please comment on the suggested question prompts (listed below). We welcome suggestions for alternatives.

²⁴ <http://www.ofqual.gov.uk/files/2011-06-15-principles-of-language-accessibility-test-developers-in-charge-of-writing-test-items.pdf>

- **What was your favourite question? Why did you like it?**
- **What was your least favourite question? Why did you not like it?**
- **Did you get stuck on any of the questions? Why were you stuck?**
- **Were there any questions that you didn't understand?**
- **What could the people who write the tests do to make them better?**

3. A group of teachers should be invited to join a Special Education Teacher Review Panel to evaluate each year's tests.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree
- Don't know

4. Please add any further comments on the formation of Special Education Teacher Review Panels.

If you are not a test development agency please go to Question 9.

The following questions are directed towards test development agencies.

5. Do you hold meetings with Special Education teachers during the development of the assessment materials?

Yes

No

If 'Yes', how often and at what stage in the test development cycle?

6. If post-release Special Education Teacher Review Panels were set up to review the accessibility of the questions, how would this be used within your organisation?

7. How would the findings of a Special Education Teacher Review Panel feed into the development of the test items for following years?

8. Do you currently use a checklist to check the accessibility of the questions during the development of the tests?

Yes

No

If 'Yes', how often and at what stage in the test development cycle? Please help us by submitting a copy or the text of any checklist that you use.

Proposal 2: Pupil Review Panels

9. Do you agree that Pupil Review Panels should be used to assure the validity and accessibility of National Assessment test questions?

Strongly disagree

Disagree

Neutral

Agree

Strongly agree

Don't know

10. Please add any further comments on the formation of Pupil Review Panels

If you are not a test development agency please go to Question 13.

The following questions are directed towards test development agencies.

11. Do you routinely carry out informal trialling during the development of the assessment materials?

- Always
- Sometimes
- Rarely
- Never
- Don't know

12. If post-release Pupil Review Panels were introduced to help to assure the accessibility of the questions, how would this work within your organisation?

Proposal 3: Differential Item Functioning (DIF) Analysis

13. Do you agree that each of these DIF analyses should be routinely carried out for evaluation purposes using data from the live key stage tests?

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
Gender	<input type="checkbox"/>					
English as an Additional Language (EAL)	<input type="checkbox"/>					
Any registered special educational or assessment need	<input type="checkbox"/>					
Race/Ethnicity	<input type="checkbox"/>					
Other	<input type="checkbox"/>					

Other, please specify:

14. What purposes might this data serve if it were collected?

- Guide future item development
- Establish a bank of biased items for further research
- Other

Other, please specify:

15. Please add any further comments on DIF analyses

16. Do you routinely carry out the following DIF analyses during the development of the assessment materials?

- Gender
- EAL
- SEN
- Race/Ethnicity
- Other (please specify below)

Other, please specify:

17. Do you agree that the collection of DIF data from the live test results for evaluation purposes would be worthwhile?

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree
- Don't know

If you would like to read more about this subject before answering these questions, please look at the Ofqual commissioned report by language experts Anthony Kunnan and Martin Walker.²⁵

²⁵ <http://www.ofqual.gov.uk/files/2011-06-15-proposal-to-the-language-accessibility-working-group.pdf>

Proposal 4: National Voluntary Training Programme

(Please note that Ofqual does not intend to run or fund any national training programmes.)

18. Please indicate your support for the described National Training Programme.

Yes

No

Undecided

Proposal 5: Best Practice Document

19. Please indicate your support for the production of a Best Practice Document as described.

Yes

No

Undecided

If you are not an awarding organisation not all the following questions may be appropriate to you.

20. Is your awarding organisation willing to participate in the design and delivery of a study to establish areas for content analysis?

Yes

No

21. Is your awarding organisation willing to collect and release data as part of such as study on the design and delivery of a study to establish areas for content analysis?

Yes

No

If Yes, for which qualifications? (Please list)

22. Is your awarding organisation interested in collecting and releasing data on difficult to comprehend language as part of a nationwide study?

Yes

No

Developing a check list for test item review

23. What additional checklist items could be included in the question paper and mark scheme analysis for written papers, internal assessments, etc.? Please list.

24. Should criteria to determine which items should not be allowed to become part of an item bank be developed by Ofqual with awarding organisations and be incorporated in Ofqual's code of practice?

Yes

No

Proposal 6: Data Collection and Analyses

25. Please indicate your support for the Data Collection and Analyses described.

Yes

No

Undecided

Language accessibility for whom

26. Please tell us which test taking groups you consider are currently affected and in which tests? (Please list)

27. Does your organisation have research evidence available to support the claim that particular groups are affected?

Yes

No

If Yes, please tell us more about the evidence?

Differential Item Functioning (DIF) analysis of test performance

28. Please tick your responses to the following statements.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
DIF analyses should be routinely used to monitor a few key assessments as well as to inform the development of test items or tasks	<input type="checkbox"/>					
DIF analysis by sub-groups, age and gender could be carried out regularly as part of the on-going test development process	<input type="checkbox"/>					

Which groups should be considered for DIF analyses?

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
Candidates with dyslexia	<input type="checkbox"/>					
Candidates with hearing impairment (and low reading comprehension scores)	<input type="checkbox"/>					
Candidates with visual impairment (and in receipt of access arrangements)	<input type="checkbox"/>					

29. Please list any other groups which should be considered for DIF analyses.

Empirical research on interaction hypothesis

30. Does your organisation have research evidence available to support the claim that particular groups are affected?

Yes

No

Empirical research on the effects of language modifications on test taking group performance

31. Is your awarding organisation interested in collecting and releasing data to assist a study on the effects of language modifications on candidate performance?

Yes

No

Validation studies across sub-groups

32. Is your awarding organisation interested in supporting and/or participating in this validation study through double-blind trialling?

Yes

No

Thank you for providing responses on the proposals.

Annex 2

Principles for language modification at source for National Curriculum tests

Topic 1: The principles of language accessibility for test developers in charge of writing test items for National Curriculum Assessments

Linguistic difficulties: Word length

1. Are you aware that difficulties for test takers could be caused by the word length of test items?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

2. In your opinion to what degree is a consideration of word length in tests of importance to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

Please give reasons for your answer to Question 2.

3. Please tell us what strategies you think will help bring the importance of word length to the attention of test developers?

Linguistic difficulties: Word familiarity

4. Are you aware that test item difficulties could be caused by test takers' word unfamiliarity?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

5. In your opinion to what degree is a consideration of test takers' word familiarity important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

6. Why do you think a consideration of test takers' word familiarity might or might not be important to test developers?

7. Please tell us what strategies you think will help test developers consider the importance of word familiarity to test takers?

8. Do you agree that a standard word list, according to the subject matter of a particular test, might be useful to test developers?

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree
- Not Applicable

9. Why do you think that such a standard word list might or might not be important to test developers?

Linguistic difficulties: Sentence difficulty

10. How far were you aware that test item difficulties could be caused by sentence complexity?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

11. In your opinion to what degree is a consideration of sentence complexity important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

12. Why do you think a consideration of sentence complexity might or might not be important to test developers?

13. Please tell us what strategies you think will help bring sentence complexity to the attention of test developers?

14. The sentence difficulty issue is confused by the fact that sometimes long sentences can be easier to read than short, and sometimes the opposite is true. In your opinion what guidance might be given to test developers to help them address this issue?

Cohesion and coherence

15. Are you aware that test item difficulties could be caused by cohesion and coherence difficulties?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

16. In your opinion to what degree is a consideration of cohesion and coherence difficulties important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

17. Why do you think a consideration of cohesion and coherence difficulties might or might not be important to test developers?

18. Please tell us what strategies you think will help bring cohesion and coherence to the attention of test developers?

Content structure

19. Are you aware that test item difficulties could be caused by content structure difficulties?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

20. In your opinion to what degree is a consideration of content structure difficulties important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

21. Why do you think a consideration of content structure difficulties might or might not be important to test developers?

22. Please tell us what strategies you think will bring content structure to the attention of test developers?

Legibility and print issues

23. Are you aware that test item difficulties could be caused by legibility issues?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

24. In your opinion to what degree is a consideration of legibility issues important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

25. Why do you think a consideration of legibility issues might or might not be important to test developers?

26. Please tell us what strategies you think will help bring legibility issues to the attention of test developers?

27. In your organisation, please tell us who decides the size of font, typeface, paper quality and colour of the test item, and why?

Text organisation

28. Are you aware that test item difficulties could be caused by text organisation issues?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

29. In your opinion to what degree is a consideration of text organisation issues important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

30. Why do you think a consideration of text organisation issues might or might not be important to test developers?

31. Please tell us what strategies you think will help bring text organisation issues to the attention of test developers?

Topic 2: What affects the readability of test items? What are the characteristics of the target readers of these items?

Physical capabilities

32. Are you aware that test item difficulties could be caused by test takers' physical capabilities?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

33. In your opinion to what degree is a consideration of test takers' physical capabilities important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

34. Why do you think a consideration of test takers' physical capabilities might or might not be important to test developers?

35. Please tell us what strategies you think will help bring test takers' physical capabilities to the attention of test developers?

Reading abilities

36. Are you aware that test item difficulties could be caused by test takers' reading abilities?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

37. In your opinion to what degree is a consideration of test takers' reading abilities important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

38. Why do you think a consideration of test takers' reading abilities might or might not be important to test developers?

39. Please tell us what strategies you think will help bring test takers' reading abilities to the attention of test developers?

Preferred reading strategies

40. Are you aware that test item difficulties could be caused by test takers' preferred reading strategies?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

41. In your opinion to what degree is a consideration of test takers' preferred reading strategies important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

42. Why do you think a consideration of test takers' preferred reading strategies might or might not be important to test developers?

43. Please tell us what strategies you think will help bring test takers' preferred reading strategies to the attention of test developers?

Engagement/motivation

44. Are you aware that test item difficulties could be caused by test takers' engagement or motivation?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

45. In your opinion to what degree is a consideration of test takers' engagement or motivation important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

46. Why do you think a consideration of test takers' engagement or motivation might or might not be important to test developers?

47. Please tell us what strategies you think will help bring test takers' engagement or motivation to the attention of test developers?

Prior knowledge

48. Are you aware that test item difficulties could be caused by test takers' prior knowledge?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

49. In your opinion to what degree is a consideration of test takers' prior knowledge important to test developers?

- Not important at all
- Not important
- Neutral
- Important
- Very important
- Don't know

50. Why do you think a consideration of test takers' prior knowledge might or might not be important to test developers?

51. What strategies should test developers use for considering the importance of test takers' prior knowledge when developing tests?

Thank you for providing responses to this questionnaire.

Annex 3

Principles for language modification at source for General and Vocational qualifications

If you would like to read more about this subject before completing this questionnaire, please look at the Ofqual commissioned report by language experts Maureen Jefferson and Janet White.²⁶

A. Current arrangements for language modification

In England, Wales and Northern Ireland, detailed non-statutory guidance on language accessibility is issued by the regulators: *Fair access by design: Guidance for qualifications regulators and awarding bodies on designing inclusive qualifications* (CCEA, DCELLS, Ofqual, 2010).²⁷

Organisations who represent particular groups of candidates also publish guidance or training courses:

BATOD (British Association of Teachers of the Deaf)²⁸

NATED (National Association for Tertiary Education for Deaf People)²⁹

Signature (Level 3 award in Modifying Written English Tests for Deaf People MET3)³⁰.

Guidance on Oral Language Modification is published by the Chartered Institute of Educational Assessors (CIEA).³¹

Commercial organisations such as the Plain English society³² likewise offer guidance and training on language accessibility for a range of end users, including awarding organisations.

²⁶ <http://www.ofqual.gov.uk/files/2011-06-15-general-and-vocational-consultation-on%20principles-for-language-modification.pdf>

²⁷ http://www.rewardinglearning.org.uk/docs/regulation/fair_access_by_design.pdf

²⁸ <http://www.batod.org.uk/content/articles/resources/training-materials/language-modification/Training%20materials%20for%20language%20modification.pdf>

²⁹ <http://www.nated.org>

³⁰ <http://www.signature.org.uk/documents/qualifications/qualification-specs>

³¹ http://www.ciea.org.uk/training_and_qualifications.aspx

³² <http://www.plainenglish.co.uk>

1. To what extent are you aware of any of the guidance mentioned?

- Not aware
- A little aware
- Neutral
- Aware
- Very aware
- Don't know

Please name the guidance with which you are familiar.

2. If you are aware of any of the guidance mentioned, how would you rate your use of it?

- Never use guidance
- Seldom use guidance
- Sometimes use guidance
- Often use guidance
- Actively use guidance
- Don't know

3. To what extent do you think that the principles set out in *Fair access by design* address issues of linguistic accessibility for you/the groups of candidates you work with?

- Would hardly address any issues
- Would seldom address any issues
- Would sometimes address any issues
- Would often address any issues
- Would address all issues
- Don't know

Please comment, explaining your views with reference to specific examples/candidates.

4. To what extent do you think that the current system of modification on demand ensures accessibility for all candidates?

- Does not ensure accessibility
- Somewhat ensures accessibility
- Moderately ensures accessibility
- Very much ensures accessibility
- Ensures complete accessibility
- Don't know

Please explain your views.

5. How far do you agree that the situation would be improved if language specialists worked with examiners at the point of paper setting?

- Would make no improvement
- Would make a slight improvement
- Would make a moderate improvement
- Would make a significant improvement
- Would make a very significant improvement
- Don't know

Please explain your views.

B. Defining the linguistic needs of different candidates

With the overall aim of 'plain language and clear presentation', some of the needs of specific candidates are identified. However, 'getting the language right' must take into account the design of the question paper as a whole, ranging from considerations of legibility, to the scope and range of cultural referents in the paper, particularly those not mandated by the subject content.

In the characterisation of different groups of candidates it is important to emphasise that there are overlaps between them. A candidate may be deaf and also learning English as an additional language; a blind candidate may be dyslexic, and so on. 'Accessible language' is not a hard and fast issue. For example, *Fair access by design* (2010) sets out overlaps in a 'common core' of language variables in the form of guidance to examiners and teachers.

Candidates who would benefit most from consideration of universal accessibility of tests and assessments are those with: dyslexia or specific learning difficulties (SpLD); speech, language and communication needs (SLCN); hearing impairments; who have English as an additional language (EAL); underperforming White British working class boys; travellers and the visually impaired.

6. Do you agree that the written examination papers with which you are familiar show an awareness of the diversity of lifestyles, cultural beliefs and developmental activities of the candidates within your educational establishment?

- Very poor awareness
- Poor awareness
- Fair awareness
- Good awareness
- Very good awareness
- Don't know

Please explain your views with reference to specific issues of diversity and inclusion.

7. In your view, how accurately and helpfully are the linguistic needs of the different groups characterised? (As characterised in Section B of the paper by Maureen Jefferson and Janet White³³)

- Not at all accurately and helpfully
- Slightly accurately and helpfully
- Fairly accurately and helpfully
- Quite accurately and helpfully
- Very accurately and helpfully
- Don't know

Please explain your view, commenting on both the language needs and the ways in which the groups are characterised.

8. To what extent do you think that access arrangements for pupils with English as an additional language (EAL) in national curriculum key stage tests (in England) should be extended to candidates with EAL in public examinations?

- Not in favour of extending any access arrangements
- In favour of extending very few access arrangements
- In favour of extending some access arrangements
- In favour of extending most access arrangements
- In favour of extending all access arrangements
- Don't know

Please explain your views, for example in relation to specific qualifications or examinations.

³³ <http://www.ofqual.gov.uk/files/2011-06-15-general-and-vocational-consultation-on%20principles-for-language-modification.pdf>

C. Principles for language accessibility

A number of researchers and reviewers have identified features of language that are perceived to have a significant effect on the accessibility of examination papers. Underlying this work is the aim of reducing unnecessary linguistic complexity in the way examination questions are written and presented. The aim is that an examination tests knowledge of subject content rather than proficiency in the English language, unless proficiency in English is part of the subject content.

Concepts considered for easier accessibility include: word length, familiarity/frequency of words, ambiguity of meaning in everyday words, abstract nouns, command words, sentence structure and complexity, and sentence length.

9. To what extent do you find the indicative guidance and examples useful to your own understanding of accessible language in writing questions for candidates? (As outlined in Section C of the paper by Maureen Jefferson and Janet White³⁴)

- Not useful at all
- Slightly useful
- Fairly useful
- Quite useful
- Very useful
- Don't know

Please explain your view with reference to particular examples/explanations.

10. To what extent do you agree that the application of these principles would enhance the accessibility of the language used in qualifications?

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Don't know

³⁴ <http://www.ofqual.gov.uk/files/2011-06-15-general-and-vocational-consultation-on%20principles-for-language-modification.pdf>

Please explain your view, especially if you have reservations about these principles.

D. Building expertise in addressing language issues

The shortcomings in the present system for developing the professionalism of language specialists have been identified as:

- lack of a central register of accredited language specialists (BATOD have a list which has been circulated to some awarding organisations)
- awarding organisations are not obliged to publish lists of people with whom they work
- work of the language specialists is seen as a 'cottage' industry by some awarding organisations and not as a professional body
- feedback not given to modifiers, and suggested changes sometimes ignored without explanation
- attendance at Question Paper Evaluation Committee (QPEC) meetings not required or sought. Instead, most modifiers work remotely and in isolation
- accredited language specialists are not used to their full potential
- item level data not available to assess the impact of modified language.

Possible ways of addressing these shortcomings could be to develop a training programme for language specialists, taking this name as description for people with expertise in improving the accessibility of language in test items (NCA) and examinations (General and Vocational Qualifications). The use of the term language specialist, as opposed to modifier, emphasises the need to place a greater stress on the fact that linguistic expertise is a necessary component in writing questions rather than the current situation of modifying papers once they are written. These courses could vary in length.

A minimum programme might be awareness training conducted as a weekend or intensive one day course or a longer-term, accredited training course aimed at teachers and other professionals. This would be a post-graduate qualification developed and accredited by a university or other accreditation body. A modular course of 120 hours could include online study, 'face to face' at a weekend or on separate days plus a final written examination. This course could be offered to newly trained teachers or teachers who are retraining as a special needs specialist, or examiners.

11. How far do you agree that the proposed courses would create a professional, well-trained workforce capable of reviewing and producing accessible examination papers at Key Stages 4 and 5?

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Don't know

Please explain your views.

12. How far do you agree that an accredited language specialist should be present at QPEC meetings for general qualifications, or their equivalent, to work with subject experts so that 'modification' is carried out at source?

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Don't know

Please explain your views, especially if you:

- have reservations about this proposal?
- think that it should be a mandatory requirement?

13. Would you be in favour of a nationally accredited programme of training for language specialists?

Yes

No

14. In what other ways could changes to the working practices within AOs, and in relation to centres, help to ensure more accessible language in qualifications?

Thank you for providing responses to this questionnaire.

Annex 4

General and Vocational Qualifications - Using Computer Adaptive tests questionnaire

If you would like to read more about this subject before completing this questionnaire, please look at the Ofqual commissioned report by language expert Michael Birdsall.³⁵

1. Is there a genuine interest in your awarding organisation (AO) to develop CAT programs?

Yes

No

Not sure

If Yes, please tell us what research your AO has done in light of this interest?

If Yes, please tell us what processes you have in place for mitigating test and question biases that impact members of different population subgroups.

³⁵ <http://www.ofqual.gov.uk/files/2011-06-15-implementing-computer-adaptive-testing-to-improve-achievement-opportunities.pdf>

2. How much of a concern for your AO is the financial investment required for CAT programs?

- A significant concern
- Somewhat of a concern
- A moderate concern
- A slight concern
- No concern at all
- Don't know

3. How significant does your AO anticipate the technical challenges are in developing CAT programs?

- Very significant challenges
- Quite significant challenges
- Fairly significant challenges
- Slightly significant challenges
- No challenges
- Don't know

4. Please rate the scale of the logistical and test delivery barriers your AO anticipates in delivering CAT?

For example, fixed testing centres, mobile testing, in school testing or some other testing centre arrangement.

- Very significant barriers
- Quite significant barriers
- Fairly significant barriers
- Slightly significant barriers
- No significant barriers
- Don't know

5. Please rate the communications challenges your AO anticipates in delivering CAT?

For example, communicating the change internally, to employees, and externally, to the public at large.

- Significant challenges
- Somewhat significant challenges
- Moderate challenges
- Few challenges
- No challenges
- Don't know

6. Please add any further comments below.

7. If your AO does not have the expertise to develop CATs, do you have the expertise to evaluate the benefits of a CAT program?

- Yes
- No
- Not sure

Thank you for providing responses to this questionnaire.

We wish to make our publications widely accessible. Please contact us if you have any specific accessibility requirements.

First published by the Office of Qualifications and Examinations Regulation in 2011

© Crown copyright 2011

You may re-use this publication (not including logos) free of charge in any format or medium, under the terms of the [Open Government Licence](#). To view this licence, [visit The National Archives](#); or write to the Information Policy Team, The National Archives, Kew, Richmond, Surrey, TW9 4DU; or email: psi@nationalarchives.gsi.gov.uk

This publication is also available on our website at www.ofqual.gov.uk

Any enquiries regarding this document/publication should be sent to us at:

Office of Qualifications and Examinations Regulation	
Spring Place	2nd Floor
Coventry Business Park	Glendinning House
Herald Avenue	6 Murray Street
Coventry CV5 6UB	Belfast BT1 6DN

Telephone 0300 303 3344

Textphone 0300 303 3345

Helpline 0300 303 3346