Minimising bias in assessment for students with special educational needs and disabled students: reasonable adjustments in written national curriculum tests at key stage 2 and key stage 3

November 2007

Nick Peacey and Lindy Peacey

SENJIT

Report commissioned by the Qualifications and Curriculum Authority 2007

CONTENTS

Introduction

In January 2007 the Qualifications and Curriculum Authority (QCA) asked SENJIT to undertake a literature review which would explore a number of issues surrounding attempts to minimise bias within national curriculum assessments at key stage 2 (KS2) and key stage 3 (KS3). In English schools, English, mathematics and science attainment is assessed by test for most pupils at the end of KS2 and KS3.

The issues covered

The review covers these issues:

- The history of the modifications, known at different periods as special arrangements, access arrangements and reasonable adjustments, in relation to their development and application across national curriculum assessments (NCAfor those with impairments/special educational needs (SEN)
- Reasonable adjustments and their justifications in relation to disability discrimination legislation and the objectives of the NCA
- The inappropriate use of reasonable adjustments
- The inter-relation between the evaluation of assessment practice in relation to reasonable adjustments and future test development
- The take-up of reasonable adjustments in assessment by schools and their application as part of NCA
- The pupils who are awarded reasonable adjustments for NCA

Note
This review uses the term ‘adjustments’ throughout for all modifications to assessments discussed in the UK literature.
Reasonable adjustments/access arrangements are known in the United States as ‘accommodations’ and that term is used in considering the American literature.
Background

Gareth Foulkes of the Disability Rights Commission gave this classic statement on the use of adjustments to assessment arrangements in 2003.

‘The purpose of an assessment is to determine the student’s achievement or skills. Assessments must be rigorous regarding standards so that all students are genuinely tested against an academic benchmark. However, they must be flexible enough to allow all students an equal opportunity to demonstrate their achievement. In all cases, this means being very clear about what is being assessed so that modifications can be made without compromising standards’ (Foulkes 2003).

Skidmore has emphasised the importance of national assessments in public consciousness and policy.

‘Assessment has enormous symbolic power in sending out a message about the skills and competences that society values as important, and provides important components of the systems used by a wide range of institutions to make decisions about people and resources’ (Skidmore 2003).

If we accept Skidmore's argument, decisions made by authorities responsible for national assessment about the participation of students with disabilities and/or special educational needs give important messages about how society values them and their inclusion in its 'high stakes' processes. At the same time, the assessments need to be seen as rigorous in terms of the standards they set for all. The route to this is the diligent consideration in the design of assessments of Foulkes' two dimensions; clarity about aims and flexibility.

It needs to be stressed that like all inclusion, this is a process of development. There is not an endpoint. In year-on-year assessments like the national curriculum tests, a body such as QCA can embed a cycle of procedures which, over time, ensure that lessons learnt about reasonable adjustments from one set of assessments feed into the design of those to come.

Minimising bias in national curriculum assessments at KS2 and KS3

The common criteria for national curriculum assessments say that they should:

- provide a valid measure of the required knowledge, skills and understanding as defined by the national curriculum and subject criteria
- generate results that provide a reliable measure of pupils' performance
- generate results that provide comparability of standards
- generate results that minimise bias, differentiating on the basis of pupils' ability to meet the requirements
- be manageable (including for pupils to complete within given time constraints and for teachers to administer and mark).

National curriculum assessments: Regulatory framework (QCA 2006)

History (general introduction)

In 1987 the Task Group on Assessment and Testing (DES and Welsh Office 1987) noted: 'Given the extremely diverse nature of special educational needs, it does not seem to us that it would be useful to issue blanket statements about the inclusion or otherwise of such children in the national assessment programme, but rather that the programme should have built into it sufficient flexibility to cope with a wide diversity of needs' (paragraph 169). The Task Group expressed the hope that the tests would be undertaken by children with SEN and that they would be sufficiently wide ranging and sensitive to respond to their needs.

The *National curriculum assessments: Regulatory framework* (QCA 2006), states that: 'The national curriculum assessments must generate results that minimise bias, differentiating on the basis of pupils' ability to meet the requirements' (paragraph 18). The *Regulatory framework* gives eight criteria for this, including the need to take account of all current legislation in relation to equality of opportunity. The National Assessment Agency (NAA) is required to demonstrate 'how it will ensure access for pupils with particular assessment needs, in line with the requirements of current legislation, without invalidating the test or disadvantaging other pupils, and publish arrangements to schools' (paragraph 38).

Designers have in essence two approaches to achieving the purposes described above:

- make the assessment itself as ‘inclusive’ as they can, so that as many participants as possible can use it without modification
- allow specific adjustments for pupils with SEN and/or disabilities to allow them to assessments.

This survey starts by considering the literature on adjustments and then discusses ideas of inclusive assessment.

The journey to the term ‘reasonable adjustments’ in national curriculum assessment

Changing approaches to adjustments since 1988

The bodies responsible for national assessment systems, the School Examinations and Assessment Council and its successors, the School Curriculum and Assessment Authority and the Qualifications and Curriculum Authority, have always sought to allow pupils with special educational needs to take part in assessments and examinations.

Special arrangements

The earliest handbooks setting out the requirements for adjustments (e.g. SEAC 1992) referred to them as ‘special arrangements’. ‘Special educational needs’ had become the accepted term for pupils who required additional support in school ever since the implementation of the 1981 Education Act. At this date there was little challenge to the phrase in the literature: ‘special arrangements’ followed naturally enough as an extension of the term.

Access arrangements

As with all terminology, ‘special arrangements’ carried connotations.

Throughout the 1990s many authors expressed concern that the term ‘special educational needs’ encouraged ‘within-child’, deficit and ‘medical’ models of difference and diversity. QCA recognised this in the revised National Curriculum documentation (DfEE/QCA 1999). The booklets included a statutory Inclusion Statement which emphasised the environmental changes necessary in classrooms to remove barriers to learning and participation.

This shift may have had less effect on the work of assessment designers than the acceptance that that there was not a simple relationship between a pupil’s special educational needs and the adjustment appropriate to such needs. ‘Special arrangements’ tended to suggest a model of modification of assessment in which the special educational need or impairment per se largely determined the arrangement adopted for an individual.

Psychologists have long been clear that this is not the case. For example, Backhouse has explained how she approaches the question of whether an individual will benefit from an adjustment to the assessment conditions (Backhouse et al 2004). She describes checking by repeat testing whether or not a dyslexic pupil will benefit from extra time, a commonly sought adjustment for dyslexics. If her repeat tests suggest that the pupil will not benefit from extra time she does not recommend it.

Similarly, it became more widely recognised that many adjustments may be helpful to a range of pupils with SEN. For example, adjustments typically made for pupils with hearing impairments may assist pupils with receptive language impairment. There is no one-to-one relationship between an identified SEN and the appropriate adjustment for an individual.

The *Assessment and reporting arrangements* have increasingly supported the approaches implied by these understandings. In the 2005 booklets the change in thinking was symbolised by the adoption of the term 'access arrangements'.

Reasonable adjustments

The QCA has now (2007) moved to the use of the term 'reasonable adjustments' in the *Assessment and reporting arrangements*. It is thus pursuing its duty, as a public body under the Disability Discrimination Act 2005, to promote disability equality. The QCA's plans to fulfil that duty are made explicit in its Disability Equality Scheme (QCA 2006b).

'Reasonable adjustments' first appeared in the Disability Discrimination Act 1995. The phrase refers to the steps necessary to avoid discrimination on grounds of disability. 'Reasonable' mediates the duty in relation to the reasonableness of an institution's undertaking a particular adjustment, given its level of capacity. For example, while it may be reasonable to expect a large department store to have lifts to all floors, it might be seen as less reasonable for a small second-hand bookshop to provide the same level of amenity.

When the SEN and Disability Act (SENDA) 2001 added a part 4 dealing with schools and other education institutions to the Disability Discrimination Act 1995, 'reasonable adjustments' became the legal term that applied to modifications that remove barriers for disabled learners.

The disability discrimination legislation makes clear that reasonable adjustments, not just in relation to assessment, should be part of standard classroom practice. Their adoption should be evident in daily teaching and learning processes. Where appropriate, the history and recording of such adjustments can be used as evidence of the need for similar modifications to assessment procedures.

If the QCA retains the term 'reasonable adjustments', with its direct reference to disability discrimination legislation, in the *Assessment and reporting arrangements* in future, it seems likely to present few problems for those involved in assessments. There is a very substantial overlap between pupils identified as having SEN and those who are disabled. Current estimates suggest disabled pupils make up 7% of the school population (Cabinet Office 2005). This is a similar figure to those in recent DfES statistics for SEN (DfES 2007): 7.4% of primary pupils and 7.1% of secondary pupils are recorded on statements of special educational needs or school action plus.

This does not mean that the populations are identical. For instance, cancer patients, even if the disease is in remission, and those with severe disfigurement are statutorily termed 'disabled'. Neither group would necessarily require any modifications to written assessments or be included among those identified as having 'special educational needs'.

Similarly, there will be a population of pupils identified by their schools as having special educational needs, probably those on School Action within the SEN Code of Practice (DfES 2001), who will not be disabled under the disability discrimination

legislation and therefore not technically covered under the law by the duty to make reasonable adjustments, though they are covered by the SEN framework.

However, none of these technical issues is likely to be very significant provided that regulators, assessment designers, teachers, parents and pupils are clear that there is no ‘automatically’ appropriate adjustment to an assessment for an individual simply because of a particular declared disability or identified special educational need. Of course, national agencies, local authorities and schools must also be clear about their duty to ensure that appropriate adjustments to assessments are put in place for all those who need them, whether or not the entitlement comes from disability discrimination legislation or the SEN framework.

The current situation: Reasonable adjustments in the *Assessment and reporting arrangements* for KS2 and KS3 2007

The *Assessment and reporting arrangements* for 2007 (QCA 2006c & QCA 2006d) state that it is impossible to list all the different situations where pupils may need reasonable adjustments. The guidance allows a fair degree of flexibility to schools, but they must have documentation showing that a pupil is eligible for such adjustments. There must be evidence that resources are routinely committed to providing support for written tasks in the classroom for those pupils and that the support is the same as is sought for, and allowed in, the national curriculum assessments.

The directions for making adjustments include the following:

- No rewording of questions
- No relatives administering tests
- The possibility of using a separate room if needed for rest breaks, additional time, amanuenses, readers or technological aids

Other general directions include:

- Pupils with a statement of special educational needs can have up to 25% extra time for written tests.
- Pupils using Braille or modified large print (MLP) are automatically entitled to 100% extra time.
- Those using enlarged print versions are awarded no extra time unless they have a statement of SEN when they are allowed 25% extra time.
- Pupils with hearing loss or using sign language are allowed 100% extra time in mental mathematics.
- The spelling test is not timed.

The types of special/access arrangements that may be used include:

- Additional time
- Use of a reader under certain conditions and according to subject
- Use of an amanuensis/scribe
- Use of transcripts (of pupils' writing)
- Use of word processor
- Use of technical or electrical aids
- Vision: modified tests
    - modified large print
    - Braille
    - enlarged print
- Recorded versions of written mathematics and science tests
- Use of coloured paper, coloured overlays and coloured filter lenses
- Enhancing diagrams, use of emphasis in modified large print and Braille, adapting Braille, low-vision aids
- Hearing: Use of communicators or signers; some use of reminders and prompts
- Responses through sign, pointing, Bliss symbolics or eye-pointing
- Using apparatus: mathematics and science: using real objects

- Concentration or fatigue: rest breaks
- Attention problems: Use of prompter
- English as an additional language: use of translators, but not for English except for general instructions
- Mental mathematics test:
    - Pupils using MLP and Braille are given MLP and Braille versions of the answer sheet.
    - Pupils with hearing impairment can use individual full-volume, earphones, or have text read aloud.
    - Modified versions are available for permanent or long-term hearing loss or for those using BSL.
    - Other arrangements available: enlarged answer sheet, extra time, translation, rest breaks.

Reasonable adjustments and their justification

Assessment is typically seen as either formative (assessment for learning), summative (assessment of learning), evaluative (to assess the success of an institution or system) or diagnostic (to understand the learning process in an individual) (e.g. DES/Welsh Office 1987).

National tests assess on a large scale and are normally described as having summative and evaluative purposes. As soon as adjustments for individuals come into play, both diagnostic and formative assessments become part of the assessment process. This can complicate matters. For example, Pitoniak and Royer (2001) note that one of the biggest challenges in the area of test accommodations is the balancing of individualised consideration with the standardisation of the tests.

We cannot, of course, see assessment in isolation from other processes. National curriculum assessment, or other large scale testing that carries a commitment to assess those with SEN and/or disabilities along with everyone else, further commits those involved to the assumption that being tested in specific areas depends on those areas having been taught. While we may understand suggestions that there is a need to 'reduce the 'high stakes' nature of assessment and to separate the need to provide diagnostic information for learners and teachers from schools' league tables and estimates of national standards' (Mortimore 2006), we should also be aware that the involvement of minorities in high stakes testing can be a driver for the improvement of teaching for that group. This, of course, needs to be set against concerns for emotional and other impacts of such testing on vulnerable groups (see below: Skidmore 2003).

'Reasonable access to tests' is defined by Elliott (1999) as a change in the way that a test is administered or responded to by the person tested. The justification for the adjustment is the intention to offset or 'correct' distortions in scores caused by a disability. He suggests thinking about this through two metaphors. His first concerns spectacles ('eyeglasses'). The accommodation (this term is used in the US and is equivalent to 'reasonable adjustments' in the UK documentation) functions as a corrective lens to provide a more valid image of the performance of individuals with disabilities. The second metaphor is that of an 'access ramp'. In other words, the accommodation is part of a package that allows the student access to a test and enables them to demonstrate what they know and can do with regard to the skills or abilities the test is targeting.

Tindal and Fuchs (1999: page 7) describe this in another way. 'Accommodations are changes in standardised assessment conditions introduced to 'level the playing field' for students by removing the construct-irrelevant variance created by their disabilities'.

There is always concern that changes in the administration of a test may affect its validity. The Joint Council for Qualifications' guidance (JCQ 2006) explains their approach to adjustments ('access arrangements'): 'Access arrangements are intended to increase access to assessments but cannot be granted where they will directly affect performance in the skills that are the focus of the assessment'.

The *Assessment and reporting arrangements* (QCA 2006c and QCA 2006d) suggest the following as components of an ideal application of access arrangements for an individual before the test takes place:

- choice of an arrangement which is appropriate to the individual
- regular use of the arrangement /reasonable adjustment in ordinary teaching situations.

This approach is supported by the writings of such British authorities as Backhouse and her colleagues (Backhouse et al 2004).

The literature

Most of the research literature in this field comes from the United States. There is a straightforward reason for this. All the states of the Union have some form of standard assessment arrangements and so, with a field of 50 states to compare and contrast, researchers have plenty to play with.

Inthe US the most commonly allowed accommodations in state policies include individual and small group administration, dictated responses, large print and Braille, extended time, interpreter for instructions, reading and simplifying/clarifying directionscomputer/machine response, reading aloud, writing in test booklet, and testing with breaks. Many of these are allowed without limitations but some have such restrictions specified Elliott and Roach 2002).

It is helpful to understand the differences between the systems on each side of the Atlantic, while remembering the substantial variations from state to state in the United States.

The literature has sharpened its focus over the recent years, not least because federal disability discrimination legislation has been making an impact. This process started with the *Individuals with Disabilities Education Act 2001 (IDEA)* which required all states to include students with disabilities in state-wide assessments and to offer accommodations where appropriate. The *No Child Left Behind Act 2001 (NCLB)* reconfirmed this federal commitment by requiring all students to be assessed in reading and mathematics in grades 3 through 8, with science assessments to be added in 2007.

The NCLB recognises the need for students with significant cognitive disabilities to take alternative assessments. Recent developments include, for example, the allowance of portfolio approaches (Thurlow et al 2005).

All this means that we find ourselves with a wide range of American articles and, more recently, literature summaries (for example, Thompson, Blount and Thurlow 2002, Johnstone, Altman, Thurlow and Thompson 2006) and meta-evaluations (Sireci, Li and Scarpati 2003, Tindal and Fuchs 1999) which succeed one another and explore similar territory.

It should be noted that the American state-wide tests are typically:

- used across a very limited range of curriculum areas
- used with pupils throughout much of their schooling
- annual events
- assessments from which no-one is disapplied (although the results can be disaggregated).

The UK literature

The UK situation is very different. Probably because far fewer tests are going on, there is far less research on the topic in the literature. Twist et al (2006) reported on the paucity of published experimental work on the subject of adjustments to national curriculum assessments and public examinations in the United Kingdom. One paper reports on the administration of the tests for pupils with special educational needs (Bartlett and Peacey 1992), and three on the assessment of students for possible adjustments (Lloyd-Bennett 1994, Hedderly 1996, Marsh 1995). Conti-Ramsden and her colleagues (Conti-Ramsden 2002) explored the experiences of pupils with specific language impairments in KS2 tests. They found that these pupils performed poorly relative to national expectations and echoed Bartlett and Peacey in calling for guidelines relating to this group in the tests.

Hedderly investigated reading accuracy, reading speed, spelling and handwriting speeds in a school population in Years 7, 9, 10, 11, and 12. He put forward norms, including those for handwriting speeds in a sentence completion test, based on those school populations. Backhouse (2004) feels that these norms do not discriminate well at the top end of the age range (for GCSE exams). She cites a test of free writing for this age group (Alcock 2001 cited in Backhouse 2004) as being more appropriate. Hedderly's work offers no separate assessment of students with disabilities such as dyspraxia. Collins (2003) conducted a small scale investigation into the use of amanuensis at A-level. Only two students were observed. Woods (2004) looked at the reading needs of GCSE students. These studies are included below.

Skidmore has written recently on the future of educational assessment. His paper (Skidmore 2003) considers at length the purposes of assessment and how there might be less emphasis on what is learned and more on learning how to learn. He does not mention special educational needs or disability explicitly. He does, however, note the cost of formal assessment to low-achieving children in their self-esteem and mental well-being. A number of states in America now 'base participation decisions at least partly on a student's educational anxiety' (Thurlow et al 2005). This raises the issue of whether, while clearly an entitlement exists, professional and policy development can explore issues of the emotional impact of national curriculum assessment for those with disabilities relating to mental health.

Further searches on the Education Resources Information Centre electronic search base, searches of the web and limited hand searches have not produced any more UK studies involving reasonable adjustments and national curriculum tests or similar

assessments, though some other papers, such as Connelly, Dockrell and Barnett (2005) on writing speed, have obvious relevance.

A search was undertaken of the websites of all the organisations mentioned in the QCA disability equality scheme as consultees (QCA 2006) but yielded few articles that could be regarded as contributing to the literature in the field. These sites tend to contribute practical advice on applying for and managing adjustments, though one (NDCS 2007) discusses a campaign to restore an adjustment to the GCSE in some detail.

Other literature

Perhaps because the US and UK are unusual in having both widespread 'high-stakes' testing and a substantial research culture, the review has not revealed significant literature addressing these issues from any other countries.

Differences in the role of individual education planning

Typically, accounts of the American use of accommodations emphasise the creation of the Individualized Education Program (IEP) as the point at which the student's participation in a state-wide assessment is decided (Thurlow et al 2005) and an accommodation is agreed and comes into force (Pitoniak and Royer 2001). The Individuals with Disabilities Education Act mandates that decisions regarding testing accommodations be supported by the student's Individualized Education Plan (IEP). Decisions regarding how a student will be assessed (e.g. alternate assessment, general assessment with accommodations) are typically made by the IEP team using guidelines provided by the state department of education and information about the student and the test.' (Cahalan-Laitusis 2004). This process should enable the student's learning before the test to be supported in the same way as the *Assessment and reporting arrangements* suggest.

There is an assumption within the process described that those creating the individual education program have the ability to choose, with the student, the appropriate arrangements. This is discussed in the section below.

By contrast, no English guidance (for example the SEN Code of Practice (DfES 2001) or its Disability Discrimination counterpart (DRC/DfES 2004)) suggests this approach should be taken during the creation of an individual education plan. The SEN Code of Practice states that pupils should be included in making decisions where possible and be given the opportunity to make choices (chapter 3), but there is no direct reference to the national curriculum assessments and adjustments for these.

Matching modifications to individuals

More seriously from the point of view of this review, much of the earlier American literature leaves out of the discussion the issues of appropriateness for individuals and the regular use of the accommodation in the time before the test. As a result the reader is left with a view which might be seen as either offering a one-to-one relationship between the accommodation and the impairment involved (for example, a diagnosis of dyslexia equals extra time), or that a range of accommodations should be offered as a bundle to pupils with a range of disabilities.

Sireci and his colleagues (Sireci et al 2003) take our understanding beyond this. Their review concluded: 'Many accommodations have positive construct-valid effects for certain groups of students. The remaining challenge is to implement those accommodations appropriately'. They recommend: 1. Local education agencies need to learn how to implement accommodations appropriately 2. More knowledge is needed about how to identify the appropriate accommodations for individual students 3. Better tests should be designed to minimise the need for accommodations

Koretz and Hamilton came at the same issue through concerns about American and other studies (including their own) which rely on 'routine' national or state collection of assessment data Koretz and Hamilton 1999). While the volume of data from such assessments is tempting for reviewers, there are 'limitations of non-experimental analysis ….for determining the effects of accommodations'. In an earlier paper

(Koretz and Hamilton 1997) they called for 'true experimental studies of accommodations'. By 1999, however, they had recognised the limitations on such experimental work, if researchers were trying to work to stringent comparative standards: 'We reiterate here the call for additional experimental studies, but it is important to realise that there may be serious constraints on the feasibility of experimental trials in the context of large scale assessments… [for example] Advocates or policy makers in some jurisdictions may…find a no-accommodations condition unacceptable, even in field trials. If that were to happen, experiments would be limited to comparisons among potentially acceptable types of accommodations.'

Koretz and Hamilton suggest that more sophisticated non-experimental studies would yield benefits. 'It might be feasible to obtain the needed data by piggybacking additional collection on the administration of ongoing large-scale assessments. For example, teacher surveys could obtain additional information about the characteristics of students, *the accommodations offered to them in instruction* [our italics] and on other assessments, and their performance on different measures of achievement. These data would provide a much more complete descriptive view of assessment and accommodations and a stronger basis for hypothesizing about the effects of format, accommodations and other factors'.

The pupils eligible for reasonable adjustments: Could there be bias in the take-up of reasonable adjustments?

In working to minimise bias in national curriculum assessments it is necessary to investigate which pupils are allowed adjustments. As we have seen above, the current regulations say that pupils who have been assessed as having SEN and/or a disability, or who are at School Action or School Action Plus, or who have a statement of SEN will be considered. Pupils with statements are automatically allowed 25% extra time for written tests. Pupils using Braille or MLP versions or with long-term hearing loss or who use sign language are allowed 100% extra time (QCA 2006 ARA

Pupils' areas of need should have been documented over a period of time and acknowledged through provision for them on the SEN Code of Practice (Backhouse et al 2004). The student should have practice and guidance in using the adjustments. (Such provision can also be seen as 'reasonable adjustments' within disability discrimination legislation.)

However, it is not clear whether there is an unbiased route to assessment and provision through the SEN Code of Practice (DfES 2001). Lindsay and his colleagues (Lindsay et al 2006) point out that the identification of SEN is often by no means a scientific process. 'Special educational needs' is a relative term, defined in the SEN Code of Practice (DfES 2001, paragraph 1:3): 'children have special educational needs if they have a *learning difficulty* which calls for *special educational provision* to be made for them' (their italics). One significant factor is the variations between local authorities (LAs) in the percentage of children identified for statements and the percentage of children with SEN but without statements. This is a matter of local policy and can lead to considerable differences, even between neighbouring authorities.

The paper points out that there is an association between SEN and socio-economic disadvantage. For example, in 2005 almost 30% of pupils with SEN were eligible for free school meals in primary schools compared with around 14% of pupils with no SEN. The tables below illustrate the differences in percentages of pupils identified with SEN in six London LAs, either high or low on the Index of Deprivation.

	Index of deprivation 2004	% secondary pupils with SEN statements	% secondary pupils with SEN without statements
Hackney	5	2.4	23.1
Islington	6	2.5	25.9
Tower Hamlets	4	3.6	15.3

Table 1 Three inner London LAs high on the index of deprivation 2004

	Index of deprivation 2004	% secondary pupils with SEN statements	% secondary pupils with SEN without statements
Kingston upon Thames	266	1.5	10.4
Merton	220	3.2	15.4
Richmond	301	3.4	13.1

Table 2 Three outer London LAs low on the index of deprivation 2004

Another factor relates to minority ethnic groups and their over- or under-representation in identified special educational needs categories. Lindsay et al (2006) found that although socio-economic disadvantage and gender had stronger associations with such an identification, and that a greater proportion of children in year 6 were identified than in other years, children from minority ethnic groups were disproportionately represented in different types of SEN compared to White British students. For example, after controlling for year group, gender and socio-economic disadvantage and compared to White British pupils, Black Caribbean and Mixed White and Black Caribbean pupils are more likely to be seen as experiencing behavioural, emotional and social difficulties, while Indian pupils are less likely to be identified as experiencing behavioural, emotional and social difficulties, specific learning difficulties, autism spectrum disorders and moderate learning difficulties.

The authors also found that there was sometimes variation between local authorities in categorisation of ethnic groups in relation to SEN identification. They state: ‘For example, while in general Pakistani pupils and White British pupils do not differ substantially in the likelihood of having an identified SEN, in 10 LAs Pakistani pupils are half as likely as White British pupils to have an identified SEN while in four LAs Pakistani pupils were 1.5 times more likely than White British pupils to have an identified SEN.’ (page 118, Lindsay et al 2006).

These differences in identification rates for groups are indicative of possible pre-existing bias within the SEN framework which would affect pupils’ rights to adjustments.

Definition and disability

A further complication is suggested in the work of such authors as Corker and Shakespeare, cited in Riddell and Weedon (2006). The approach argues that while 'a binary line between disabled and non-disabled people may be useful politically and legally…such a hierarchy may not accord with people's experience, as impairments fluctuate and are experienced in different ways over an individual's life course'.

Riddell and Weedon also note that what is constructed as an impairment across a culture at one point in time may be viewed differently at another. They note as an example the rapidly increasing diagnosis of attention deficit (hyperactive) disorder and might well have included a similar phenomenon in relation to the growth in diagnosis of autism spectrum disorders. An historical survey of the development of the current *Assessment and reporting arrangements* and their predecessors would show the process at work. For example, while dyslexia, deafness and blindness are likely to trigger adjustments right through the series of English guidance booklets, autism spectrum conditions are still not mentioned in the 2007 version (QCA 2006c and QCA 2006d).

Riddell and Weedon point to the 'categorising' arguments which have swirled round many impairments, perhaps most contentiously around dyslexia. This is also illustrated in the case of the identification, assessment and classification of specific language impairment. Different professionals have different views of the definition and assessment of such an impairment (Dockrell, George, Lindsay and Roux 1997). Teachers may be uncertain of how to meet the children's needs (Dockrell and Lindsay 2001), and parents and teachers may have different perspectives on the children's language and behaviour (Lindsay and Dockrell 2000).

Pitoniak and Royer (2003) also note the conceptual and practical problems of defining 'learning disability' (specific learning difficulty) across the United States. The literature on the identification and diagnosis of developmental learning and cognitive disabilities suggests this holds good elsewhere. Backhouse and Morris (2005) come to a similar conclusion in relation to reading and dyslexia.
They go as far as to state 'At this point we have to acknowledge that there is no one universally agreed definition of dyslexia!' (Backhouse and Morris 2005 page 16). The definition in the DfES Pupil Level Annual Schools Census (PLASC) guidance includes: 'Pupils with dyslexia may learn readily in some areas of the curriculum but have a marked and persistent difficulty in acquiring accuracy or fluency in learning to read, write and spell'.

Desforges (in press) argues similarly that the identification of moderate learning difficulties is not straightforward. He notes the difficulty in operating the definition given in the DfES guidance for the Pupil Level Annual Schools Census (PLASC) for students with moderate learning difficulties. 'Pupils with moderate learning difficulties have much greater difficulty than their peers in acquiring basic literacy and numeracy skills and in understanding concepts. They may also have associated speech and language delay, low self-esteem, low levels of concentration and under-developed social skills' (DfES 2005).

Teachers' accuracy in determining who would benefit from reasonable adjustments

There is also the question of the accuracy of teachers' (and students' and parents') assessments and beliefs about who would benefit from reasonable adjustments.

Twist and her colleagues (Twist et al 2006) note that American studies had found that many teachers' decisions were not accurate about which students would benefit from accommodations. They investigated this further in the experimental phase of their study and found that teachers accurately predicted only about one fifth of all the pupils (those with and without SEN) who in fact benefited from extra time in reading and science tests (the tests used were developed for use in the national tests). The team therefore argue that pupils who may benefit from additional time may not be allowed it, while pupils who have to meet specific criteria described primarily in terms of attainment may not need it.

This finding, allied to the work of Backhouse discussed above, suggests that the adjustment of extra time has manageability implications: if teachers are not good at deciding who needs the adjustment, it might be argued either that they need more professional development in that area or that other more specifically trained professionals should undertake such judgements[1]. We return to this area below.

Schulte et al (2001) suggested that pupils needed to be involved in selecting accommodations. Fuchs, Fuchs, Eaton, Hamlett, Binkley and Crouch (2000) studied 4th and 5th grade children with learning difficulties and 4th grade children without learning difficulties and found reason to question teachers' judgements about test accommodations. Teacher judgements were compared with those based on the Dynamic Assessment of Test Accommodations (DATA). The DATA is a tool for assessing students' reading in four conditions: standard time, extended time, large print and with the student reading aloud.

The study found DATA a more accurate way of identifying student need for accommodations than teacher judgement. Teachers over-identified many students who did not benefit and failed to identify some who would have benefited from accommodations. The researchers suggested that teachers were influenced by demographic and performance factors and that they had limited experience in making decisions about accommodations.

Helwig and Tindal (2003) investigated the accuracy of teachers in recommending the 'read aloud' accommodation to students taking a large-scale mathematics test. They found that they were no more successful than chance at predicting students who would benefit. However, in their discussion the authors mention that the students may have been unfamiliar with the video presentation used in the test, and that instruction in the classroom may not have involved this particular format.

[1] The Joint Council for Qualifications(JCQ 2006) recently decided to allow heads of examination centres to select the staff to assess candidates for access arrangements (e.g. for GCSE and GCE). Until recently, those performing that task had to have taken a JCQ-approved specialist qualification. It will be interesting to see what effect this change has on adjustments and professional development at that level (see the following section).

Papers by three educational psychologists (Lloyd-Bennett 1994, Marsh 1995, Hedderly 1996) were concerned with the assessment of pupils, mainly those with specific learning difficulties (dyslexia), at GCSE level and above for reasonable adjustments, from their professional perspective. Hedderly noted the need to collect data about candidates' specific needs over an appropriate length of time, from before the commencement of the course. Lloyd-Bennett was concerned that a more 'cohesive system' of special arrangements should be provided, but also that attention should be given to a student's individual needs. Marsh also addressed educational psychologists' concerns about assessments. She concluded that although much work had been done, the notion of 'special arrangements' in national exams would continue to be contentious. No further recent articles have been identified. This may indicate that such concerns have been satisfactorily addressed.

Finally, we should record an unsettling possibility raised in one study: 'identifying students as disabled does still confer one advantage to educators trying to maximise scores: It permits them to provide them assessment accommodations….'(Koretz and Hamilton 1999). The authors wonder whether 'educators are responding … to classify additional students as disabled in order to raise scores.' We note that there has been no research into this suggestion.

Professional development for teachers on decisions about adjustments

Destefano, Shriner and Lloyd (2001) trialled an approach to training for general and special education teachers making accommodation decisions. They started by exploring students' IEPs, the relevance of the curriculum and accommodations used during instruction. The training then considered accommodations appropriate to the students in tests. The process involved direct training of 10-15 hours plus 'many more hours of informal consultation and feedback'. The authors noted the need for professional development in this area. They also felt that before training the teachers had tended to allow accommodations to all students with disabilities regardless of their assessment needs. The authors concluded that their study highlighted issues of 'knowledge and procedural needs' for special and general educators, maintaining the individuality of assessment decisions in system-wide testing and 'systematic planning for the scope and duration of professional development investments.' The researchers appeared to feel that the training improved the quality of the teachers' recommendations of accommodations although the methodology did not allow them to be certain. The teachers expressed greater confidence in their decision-making.

Similarly, Elliott and Roach (2002) noted that educators in the United States responsible for administering large-scale assessments to students with disabilities needed a wide range of knowledge to successfully select and use accommodations. They found that teachers tended to rely primarily on professional judgement, rather than their own empirical testing of accommodation effects, and rarely requested information on research.

A further American study (Crawford and Tindal 2006) investigated the knowledge and beliefs of education professionals related to the inclusion of students with disabilities in a state assessment in Oregon. They found that although teachers were aware of policies, this did not necessarily change their beliefs and practices. They were unsure of the usefulness of the test data in driving instruction, and whether the test scores

were valid indicators of the students' knowledge and skills. The teachers also believed that parents frequently did not understand the meaning of the test score data. The authors suggest that the next step is to strengthen the connection between educational policy, test data and classroom practice, through the state's education department clarifying the links for teachers 'between the concepts taught and the skills tested'.

The adjustments in the literature: Extra time

Extra time is one of the most commonly used reasonable adjustments in national curriculum assessment. It is best known for being made available to students with specific learning difficulties such as dyslexia. It is equally common in the United States, where this form of reasonable adjustment has secured more lines in the research literature than any other. We take this adjustment first not only because of its frequent appearance in research studies, but also because the issues raised by its use are important for consideration of the whole field.

The QCA *Assessment and recording arrangements* give several criteria for the use of additional time. These include: reading accuracy and speed below the average range for the pupil's age; a writing speed of fewer than 10 words per minute; or a discrepancy between cognitive and performance ability.

Time and the tests

Twist, Donahue, Lewis, Keogh (Twist et al 2006) undertook a major review of the literature on extra time for the National Assessment Agency. This was allied to an investigation, involving comparative experiments in which participants took either reading or science papers with and without extra time. They found that many students will benefit from extra time when large cohorts are involved, though they were not able to feel confident about reasons for this. The experiments suggested that pupils do better on reading and science tests with extra time, but particularly on reading tests, when over 40% gain marks in extra time. The researchers argue that speed is therefore part of what is being tested in the reading test.

A high proportion of pupils (28%) in their study felt they needed more time to complete the reading tests - but higher proportions than that 28% actually increased their marks when given an allowance of extra time. The authors only investigated reading and science assessments, but they note that additional time in relation to mathematics and, especially, writing is important. They argue (following Hedderly 1996) that standardised tests of writing speed with norms are important for the appropriate implementation of adjustments relating to additional time. In a recent small study by Connelly et al (2005) of an undergraduate population in timed conditions the authors were surprised to find that writing speeds were similar to those of 11 year old children. Speed clearly affects the quality of handwritten responses in an assessment. But there are likely to be issues of manageability for those organising assessments in schools as writing speed tests are not often carried out as routine.

Twist and her colleagues also found differential effects between open and closed questions on test results and in relation to the order of items in the test.

Tindal and Fuchs (1999) discuss the fact that there is a history of mental tests being a combination of speed (timed or un-timed) and power (difficulty). Tindal et al argue that the timed conditions may not allow disabled students to reflect their full 'abilities' and may actually introduce error variance. (It should be noted that much of the work on 'speededness' has been done with older students: less has been done with

elementary and middle school pupils (Elliott and Marquart 2003 quoted in Twist 2006)).

In effect, Twist and her team's work suggests that the combination of speed and power is also reflected in the impact of many of the national curriculum tests on disabled and other pupils. Significantly, though, it is not clear that the specifications for the NCA developers include a requirement to create assessments which test speed in completing the assessment over a given time. In effect, the specifications state a given time and then tell the developers to write a test to fit. The quality criteria include the requirement that assessments must 'be manageable for pupils and capable of being completed within any given time constraints, eliciting optimum performance'.

Twist et al note the findings of Elliott and Marquart (2003) and Woods (2000) that the granting of extra time seems to principally allow students to feel less anxious about the assessment situation.

The research on extra time

Extensive research into the use of extended time in assessments has yielded 'mixed results' (Thurlow et al 2005). For example, Johnstone et al (2006) noted that 'several studies found that students with disabilities profit from extended time accommodations'. Buehler (2002) and Elliott and Macquart (2004), however, found no significant effect on scores when students were provided extended time. Elliott and Macquart did, however, find that students reacted positively to the extra time condition. They discuss the feelings of self-efficacy that may be promoted for students by allowing extra time, while noting that in a few cases there is a negative effect.

Johnstone and his colleagues felt that this variety related to the lack of consistency in accommodation research, not least to differences in the administration of the assessments involved. It is not clear whether the age of students or the severity/mildness or type of their impairments affect the results of assessments because these are not usually targeted in the large scale US projects.

The picture is certainly not straightforward. Twist et al (2006), in their experimental study of year 6 pupils at key stage 2 taking reading and mathematics tests, found that 'the provision of additional time appears to have different effects on the performance of pupils working at different levels in reading and writing, depending on the subject.' They further noted that both pupils with special assessment needs and those without gained additional marks from extended time: 40% of those taking the reading tests and about 20% of those taking the science tests.

This raises a major issue in relation to accepted notions of the maintenance of validity in relation to access arrangements. These were set out by Phillips (1994). She argued that three questions need to be answered in the negative if an access arrangement is to be valid in relation to the test outcomes.

- Will alterations in testing conditions change the skill being measured?

- Will taking the examination under altered conditions change the meaning of the resulting scores?
- Would non-disabled examinees benefit if allowed the same accommodation?

The first two stipulations are relatively uncontroversial. If a skill being measured or the meaning of test scores is changed by an access arrangement, few would argue that the arrangement is maintaining the validity of the assessment.

There is less unanimity in relation to the third condition proposed by Phillips. Elliott and Roach (2002) challenge it wholeheartedly. 'This is a questionable criterion that is not feasible for practitioners to apply and is likely to deny students with disabilities some appropriate accommodations.' (page 12).

In fact, the many recent meta-evaluations of the research on various 'accommodations' have shown that there is a great range of adjustments in operation across the English-speaking world. The ways in which they are applied provide an even greater level of variety.

It is likely that many of these adjustments, even if we are identifying the pupils with SEN and/or disabilities accurately, will benefit some of those outside that group in one way or another. For example, it is commonly suggested (e.g. QCA 2006) that allowing an assessment to be taken in sections (i.e. with rest breaks) is an appropriate adjustment for pupils experiencing emotional difficulties. This adjustment has not been extensively researched (see below); even so, it is hard to believe that it would not benefit the scores of other students than those who currently receive it.

As we have seen, Twist et al found that extra time in reading tests benefited more than just the pupils with SEN. Sireci et al (2003) confirm this pattern. 'The vast majority of studies …showed that all student groups ….had score gains under accommodation conditions.' In other words, if we were to accept Phillips' third condition at face value, we would probably have to discard many of the adjustments/accommodations now in use.

In the United States, the position now (2007) is that in some states students with and without disabilities are eligible to have testing accommodations. Examples quoted by Elliott and Roach include Oregon, Washington and Rhode Island (Elliott and Roach 2002). In their work they have found that non-disabled students will benefit from accommodations almost as much as disabled students, while some students from both groups did not benefit.

They suggest that students should be tested in accommodated and non-accommodated conditions to assess the effects. They conclude that the validity of accommodations needs to come from multiple sourcesstudent factors, test factors and the accommodations themselves (Elliott and Roach 2002).

Sireci et al (2003) proposed an amended version of the 'pure' form of Phillips' third condition. They critiqued a wide range of studies of accommodations against the condition. While the results in many cases were highly variable, the researchers note 'a fairly consistent finding was that the accommodation of extended time improved the performance of students with disabilities more than it improved that of students without disabilities.' This led them to what they call the 'revised interaction

hypothesis'. They argue that where all groups benefit, and when pupils with learning disabilities benefit substantially more than those without, it is not that the accommodation is invalid, but that the standardised test conditions are 'too stringent for all pupils'. This takes them to the consideration of universal design options (see below).

Applications for additional time in 2006

The total number of applications for extra time received by QCA for 2006 was 25,091. The percentage approved for key stage 2 was 87% (total applications 16,738) and for key stage 3 (total applications 8,393) was 88%. The majority of applications received were therefore considered appropriate.

Scheduling, timing and settings

Alternative settings, individual or small group administration

Alternative settings and individual or small group administration are generally not perceived as a problem in the literature, provided there is appropriate resource and supervision. Individual or small group administration would be more likely to be used with children with behavioural or attention difficulties, and with those who require read aloud or computer based administration, or dictated responses.

The *Assessment and reporting arrangements* (QCA 2006) suggest allowing pupils to use a separate room if they are using access arrangements such as rest breaks, additional time, amanuenses, readers or technological aids. Other settings are allowed in exceptional circumstances, such as in a hospital, home, or pupil referral unit.

There is very little research from the United States (and none identified from the UK) on the effects of individual administration or small group administration as a separate factor. According to Elliott and Roach (2002) quoting from Thurlow and Bolt (2001) 44 states allow individual administration without limitations and 41 allow small group administration without limitations. Five studies are cited in Thompson, Blount and Thurlow 2002 as including small group or individual setting as part of a group of accommodations. Generally such accommodation packages helped. In contrast, Koretz and Hamilton (2001) found that disabled students scored lower than non-disabled when using extra time and a separate location.

Tindal and Fuchs 1999 cite two studies, both of which found that setting and group size affected, in one, the performance of oral reading (3rd and 4th grade students at three reading levels) and in the other, motor imitations (eight pre-school children with disabilities). However, these studies were also looking at the effects of examiner familiarity.

Extended time and testing breaks

Breaking up an assessment into chunks

The *Assessment and reporting arrangements* (QCA 2006c and QCA 2006d) allow rest breaks to pupils who find it difficult to concentrate or who are likely to experience fatigue by splitting the tests into sections or stopping the clock. There is no reference to small group administration.

Four studies relating to this are mentioned in Thompson, Blount and Thurlow 2002. One found that participants scored 12 scale points higher for divided time administration than for the single session administrations (DiCerbo et al 2001). Schulte et al (2001) found students with disabilities receiving accommodation packages other than extra time and read test items experienced significant and differential impact on mathematics scores (these included frequent test breaks). General education students did not perform as well if given multiple days for assessments. Sireci (2003) states that these results did not support multiple-day accommodation. Elliott et al (1999) found that extended time resulted in higher performance than other accommodated conditions. Waltz et al (2000) found that multiple-day test accommodation did not improve the scores of pupils with disabilities.

The evidence suggests that this adjustment does not normally alter test constructs unless ‘speededness’ in undivided time is important, and it is therefore probably not controversial. It does, of course, make some demands on assessment managers in terms of its careful administration.

The adjustments in the literature:

Presentation

Oral administration (read aloud), read/clarify directions, Braille, sign interpretations, audio or video cassette, computer-based presentations

Readers
The *Assessment and reporting arrangements* (QCA 2006c and QCA 2006d) allow the use of a reader at a school's discretion, when the use of a reader is normal classroom practice and there is evidence to show that resources are routinely committed to providing this support. A reader can only be used on a one-to-one basis. In most cases this will be with a student whose reading age is much lower than their actual age (a reading age of nine or lower). Use of a reader for test instructions is permitted except in an English paper where the instructions form part of the competences assessedUse of a reader for general instructions that do not form part of the competences assessed is permitted in English. In mathematics and science tests a reader may help a pupil to read any part of the tests.

The right to read-aloud accommodations in the United States reported in Elliott and Roach (2002 p 6) was defined by a legal case which involved the Hawaii State Department of Education (1990). The Office of Civil Rights ruled that the State could not deny a student with a learning disability read-aloud accommodations on the non-reading portions (i.e. those where reading was not being assessed) of the Hawaii-mandated graduation test (Phillips 1993). It also determined that accommodations must be judged on a case-by-case basis (Phillips 1994), and that read-aloud accommodations may not be appropriate for all students with learning disabilities. In its ruling, the Office of Civil Rights conceded that allowing a reader for the reading portion of the test would 'defeat the purpose of the test and that denying it would not be discriminatory' (Phillips 1994).

Woods (2004) investigated the provision of a reader to 40 year 11 candidates during history and mathematics examinations. A reader was available on request during the practice examinations. In fact, the invigilator responded to 20 requests from four students in the history exam, and four reading requests from three candidates in the mathematics exam. The responses were dealt with quickly by reading words, phrases or sentences. Although the awarding bodies suggest a reading age of under ten years to determine eligibility for this provision, Woods found that neither reading age nor candidates' self-predictions were reliable indicators of need for a reader.

Building on this, Woods argues that if reading age is not a reliable indicator of the need for a reader, the use of direct observation and recording of candidates' actual examination behaviours would be better. He further suggests that widening the provision would be feasible and the benefits could include reduced student disaffection and reduced examination anxiety. It would also reduce some of the administrative burden of special examination arrangements and increase satisfaction across user groups (teachers, parents and candidates).

The students in this study were GCSE candidates so the findings should be interpreted with caution for younger groups.

Sireci et al (2003) conclude that about half the studies they reviewed on 'oral accommodations' found a positive effect (they reference four which found positive effects) but several found either no gains for pupils with SEN (they reference one) or similar gains for students without SEN (they reference two studies). The methods of presentation included the teacher reading, the student reading aloud and the use of screen-reading software. They suggest that the benefits of oral presentation are unclear, partly because only certain subsets of students with SEN benefit from this type of accommodation and partly because administering it to a larger group of students obscures its effects.

Tindal and Fuchs (1999) reported six studies in which students had the test read to them. They note that often the 'read-aloud' accommodation is one of several administered at the time, but that five of the studies concluded that read aloud had a positive effect for students with disabilities.

However, one study from Kentucky which they quote (Koretz 1997) concluded that oral reading (as well as rephrasing, cueing and dictation) in mathematics and science tests given to students with moderate cognitive and learning disabilities, was biased by these accommodations. The disabled students achieved scores near the mean of students without disabilities who did not receive the accommodation. Koretz concluded that inappropriate use of the accommodations (see below) was the most likely explanation for the finding.

A recent American study by Fletcher et al (2006) tested accommodations which were chosen to minimise the barriers faced by dyslexic students with word decoding and to avoid affecting the typical readers. The subjects were a group of Grade 3 pupils identified as dyslexic and a group of 'average decoders'. They were assessed in standard conditions and with accommodations, including the oral reading of proper nouns, the comprehension stems and the possible responses to the (multiple choice) questions.

They found that the dyslexic students improved their scores with these accommodations, although their scores were still not comparable to those of the 'average' students. They noted that many students have decoding difficulties and that these types of accommodations would benefit a large number of them.

They also noted that a single accommodation would probably have only a weak impact on performance and that a 'bundle' is more effective, although it is of course less easy to identify the effects of the individual components in such an experiment.

Interestingly, in the literature review and rationale for this study the authors suggested that reader accommodations were more appropriate for dyslexic pupils than extra time.

Overall, the evidence suggests that the considered use of readers in assessments is both effective and equitable. There are, however, substantial manageability issues, particularly in timetabling staff to implement the adjustment properly.

Braille, Modified Large Print (MLP) and enlarged print

Students with visual impairments who have a statement of special educational need or are at School Action Plus (or have a medical condition based on visual impairment) are allowed the provision of Braille, modified large print (MLP) or enlarged print papers (QCA 2006c and QCA 2006d). Enlarged print is also available for students with other special educational needs. Schools may record the tests, for mathematics and science (but not for English tests). In America 35 state policies allow the use of Braille without restrictions and 14 allow its use with restrictions. The evidence suggests that such access arrangements are largely unproblematic when used by pupils with visual impairments. They are likely to be in current use in the classroom and are of benefit only to the students with visual impairments who regularly use them. However, Bolt and Thurlow (2004) note two earlier studies which suggest that certain mathematics items (eg. some figures and graphs) may be particularly difficult for students using Braille.

Landau (2003), cited in Johnstone, Altman, Thurlow and Thompson (2006) found that students performed better on five of eight items when using the Talking Tactile Tablet (TTT) (a hybrid between a Braille paper-based test and laptop computer) and scored equally well on the remaining three, with or without the tablet; use of the TTT also yielded item difficulties that more closely resembled the item difficulties obtained by general education students during testing.

Sireci, Li and Scarpati (2003 page 14) found that use of Braille, large-print and oral accommodations on mathematics tests seemed to improve the performance of students with disabilities but not the performance of students without disabilities. Only one study specifically noted the use of Braille which was among several other accommodations used in a large-scale state assessment (Johnson, Kimball, Brown and Anderson 2001).

Tindal and Fuchs (1999) mention eight studies that used large print or Braille. Large print is often used with other accommodations and with generally learning disabled pupils, not only those with visual impairments. They found some evidence that Braille and large print do indeed help visually impaired pupils.

The findings suggest that overall such accommodations are likely to be unproblematic.

Sign interpretation

Current guidance (QCA 2006c and QCA 2006d) suggests adjustments for pupils with hearing impairments, including the use of communicators and signers. In the United States sign interpretation is allowed in 37 states without restrictions, and in eight with restrictions. Five states do not mention this accommodation specifically.

Johnson et al (2001) (cited in the summary of research by Thompson et al (2002) and in Sireci et al (2003)) carried out a study investigating the use of American Sign Language as an accommodation. They found that it presented political, practical and psychometric challenges. Their data showed that sign language translation may result in the omission of information required to answer a test item correctly.

Sireci, in considering Johnson's work, noted that the differences appeared as a function of the type of translation and grade of the pupil (there was a loss of information at grade 4 but not at grade 7) and whether or not a native American Sign Language speaker did the translation: in that case, there was no loss of information. If this research is corroborated by future research there would obviously be implications for the specification of adjustments for deaf students using sign language, particularly taking into account the age of the students and the competence of the sign interpreter.

Huynh (2002) cited in Johnstone et al (2006) found that students with disabilities, including those with hearing impairments, using appropriate accommodations, including administration in sign language, performed as well as students without disabilities.

Sign language interpretation has been a feature of the English system for many years. While it is generally regarded as unproblematic, the literature suggests that this is dependent on the quality of the interpreters and this factor should always be taken into account when seeking to minimise bias for deaf pupils in assessments.

The British Association of Teachers of the Deaf has produced useful documentation to guide developers in removing barriers to written assessment for learners with hearing impairments

Computers, videos, audio tapes

Computer-assisted presentations necessitate some familiarity with their use on the part of the students. The majority of the studies quoted in Tindal and Fuchs (1999) examined the assessment of older students and young adults. They found that the lack of large studies, the changing computer technologies, and the use of computer presentations with other accommodations made it difficult to draw conclusions. Their discussion first described computer-assisted testing in which the computer responds to the students' responses, with more difficult items if the response is correct and easier items if the response is incorrect. Computer-based testing offers a medium which allows the presentation of a 'conventional' test, but can remove barriers for students with disabilities because changes can be made in the manner in which items are displayed, paced, sequenced, or presented. Items can be presented singly or grouped.

Tindal and Fuchs report almost no research on accommodations for students with physical disabilities, although computer-assisted testing may be the only viable accommodation for them without modifying the test.

Thompson et al (2002) found nine studies using computer administration to support students with disabilities. Four found a positive effect on scores, three studies found no significant effect, and two studies found that computer presentation altered item comparability, affecting the construct the assessment was intended to measure. The authors report that more research is necessary into familiarity with computer use, screen navigation, screen readers, and use of speech recognition software.

Much more recently, Johnstone et al (2006) found five papers reporting computer-presented tests used by students with disabilities. Three found no significant effects on scores, one found a positive effect, and one identified differential item functioning, that is to say that items were in fact changed in validity as a result of format differences. The types of disabilities were either wide-ranging or not reported.

The adjustments in the literature: Response

Amanuenses/scribes

The current arrangements (QCA 2006c & QCA 2006d) state that the use of an amanuensis must be normal classroom practice and used routinely if a student is to be given permission to use one in an assessment. The arrangements give precise instructions for the role of an amanuensis and suggest that using a word processor or transcription should be considered first.

Sireci's review (Sireci 2003) noted a study on the use of scribes which found that among many other accommodations for pupils with special education needs there was no unfair advantage in using them. Earlier, Tindal and Fuchs (1999) had found six studies relating to the use of scribes for children with mainly cognitive or physical difficulties. Their evidence is overall somewhat inconclusive. Although the use of a scribe appears to boost performance there is little information about their selection and training and the conditions under which they are used.

An English study by Collins (2003) on the use of a scribe with two A level students found that the adjustment was not 'straightforward', even though one candidate had considerable previous experience of dictating. As the author notes, a student needs practice to use this method in timed essays and examinations. Other methods such as voice recognition software may in time become more acceptable alternatives. (Tape recording, though it may appear to have attractions and will certainly be appropriate for some pupils with visual impairments, is less likely to be generally popular as an adjustment because the absence of visual feedback is a disadvantage.)
This study only briefly comments on the effects on results: the students in the comparison groups felt that their dictated essays were inferior to their written essays, although the author found that this was not the case. Collins concluded:

'The difficulties and anxieties registered by the subjects of this study were mainly related to the social context. The working relationship between scribe as adult to the student is an evolving process involving a shift of control from the scribe as adult to the student as dictator. However, getting to this stage is a process that requires careful preparation and sympathetic handling in order to allow candidates to do justice to their potential. This raises serious concerns about the need for funding the provision of a scribe, not only for the examination period, but also for the duration of the course preceding the examinations.'

The research evidence suggests that this adjustment, while it must be used as the *Assessment and reporting arrangements* instruct, requires practice by the student and training for the scribe to ensure correct verbatim responses with no cueing or prompting (Bolt and Thurlow 2004). It is also more demanding in the provision of personnel.

Calculators/abacus

Johnstone et al (2006) identified only one study that considered calculator use in assessment. This found that calculator use had no significant effect on the scores of students taking the SAT (US).

Dictionaries and spell checkers

Johnstone et al (2006) identified only one study on the use of dictionaries. This study using Israeli students found that dictionaries were not an effective accommodation as they interrupted thought patterns.

The inappropriate use of reasonable adjustments

The appropriate or inappropriate use of reasonable adjustments is the focus of many of the studies in this review in the general sense of whether the validity of a test is compromised by an adjustment. But studies discussing inappropriate use in the less refined sense of adjustments being used to give students an unfair advantage are much less frequent.

Two examples have been identified by this review. Sireci (2003) and Koretz et al (1999) reported an instance in Kentucky. In this state-wide assessment, accommodations such as extra time, scribes, oral presentation, paraphrasing, technology, interpreters and separate sessions were widely used. Many students received more than one of these accommodations. It was felt that some of the scores of students were implausible and probably based on the misuse of accommodations.

The other instance is a newspaper article quoted in Twist et al (2006). The article suggested that the use of extra time was being abused by schools and parents and giving an unfair advantage to some students.

These two examples illustrate the need for regulators to ensure clarity about who may need adjustments and how these will be appropriately used to maintain the validity of a test.

Towards inclusive assessment

Universal design

In the United States, Thompson, Johnstone and Thurlow (2002) at the National Center on Educational Outcomes (NCEO) have developed the ideas of universal design, which were first applied to architecture, to large-scale assessments. Their seven elements of universal assessment design are:

1. Inclusive assessment population
2. Precisely defined concepts
3. Accessible, non-biased items
4. Amendable to accommodations
5. Simple, clear, and intuitive instructions and procedures
6. Maximum readability and comprehensibility
7. Maximum legibility

Universally designed assessments are based on the premise that each child in school is a part of the population to be tested, and that testing results should not be affected by disability, gender, race, or English language ability (Thompson et al 2002).

The equivalent approach in England is represented by the specifications from the National Assessment Agency (NAA). These include the requirement that the tests can be manageable for pupils and capable of being completed within any given time constraint. They must also, in standard or modified form, be capable of being used with all pupils working within the targeted levels, including those with special educational needs and those for whom English is an additional language. The test material must provide equality of access and opportunity through minimising ethnic, gender and cultural bias, avoiding disadvantaging pupils from particular backgrounds, and freedom from covert or overt discrimination, either through wording or content. Test development agencies must take account of all current legislation in relation to equality of opportunity. The tests are to be developed in a modified form (large print, Braille, enlarged print and for hearing impairment) by modified test agencies and, in addition, advice from special needs experts must be sought.

The NAA ensure that the tests are trialled, pre-tested, and reviewed. The test material should ensure equality of access for all eligible pupils. Arrangements for access for pupils with particular needs must be in line with the requirements of current legislation, without invalidating the test or disadvantaging other pupils, and must be published for schools.

Discussion

This review has reinforced the position of Twist and her colleagues (Twist et al 2006) that research into adjustments is overall somewhat limited. This is particularly so in relation to England where few, if any, experimental studies have been carried out on the effects of reasonable adjustments in large-scale high stakes assessments.

Furthermore, we have noted that research in this field is not easy (see, for example, Koretz and Hamilton 1999). However, there is enough in the literature to draw some tentative conclusions. Most of the adjustments set out in the *Assessment and reporting arrangements* (QCA 2006c and QCA 2006d) have been examined by studies in the United States which, although they do not allow precise comparability, suggest that for the most part such adjustments will benefit the pupils with SEN and/or disabilities for whom they are designed and will not prejudice test validity. For example, the literature suggests that the use of vision and hearing aids in assessments is largely unproblematic. On the whole they do not alter the test construct and are based on individual need. They also have differential effects in that they help those who need them and do not help those who do not require them.

The literature gives virtually no evidence for the deliberately inappropriate use of accommodations in America and none in the United Kingdom.

While there is some conflicting evidence about the use of scribes and readers, the literature seems to us to suggest that careful monitoring of the implementation of the arrangements for these adjustments is a more appropriate response than further research. In other words, if pupils have been making regular use of these reasonable adjustments in their lessons, there is little reason for concern about their being permitted in assessments. It is, of course, absolutely necessary for scribes and readers to stick to the rules during the assessments. Indeed, as far as readers for questions are concerned, we might consider Woods' argument (Woods 2004) that use of readers is not contentious and could be made available, on request, to all those taking the assessments. Technology can perhaps offer other ways of addressing the manageability of this adjustment (see below).

The granting of extra time is an exception to this generally optimistic picture. This adjustment has been the subject of extensive research and there is doubt as to its effects. It appears to benefit some with and without disabilities and/or SEN. Furthermore, accurate decisions on those likely to benefit from it are hard to achieve without some sophistication in the methods used and consideration for the time allowed for those decision-making. There is also some evidence that a major effect of extra time is motivational: it gives confidence to those worried about being rushed to complete an assessment.

In fact national curriculum tests may well appear to be 'speeded' (i.e. the time taken to complete the tasks is part of the assessment) to a proportion of those taking them. But it is not clear that this is intended or necessary. If it is decided that national curriculum assessments are not meant to assess pupils' speed in taking them, a longer overall time limit, made available to all, could remove the need for almost all extra time allowances and save the resources expended by the substantial bureaucracy that surrounds the granting of this adjustment to 25,000 pupils a year.

Trialling could establish whether the extension of all national curriculum assessment times in this way would be manageable and would not substantially disadvantage pupils with SEN and/or disabilities compared with the current situation. The argument for such trialling, in essence, is this:

- The literature suggests that much of the granting of extra time is likely to be 'inaccurate'. If the present arrangements are maintained, the system should be modified to take account of the research evidence on this area.
- To bring 'accuracy' to the level required either requires substantial professional development of the many teachers concerned or the use of trained professionals, such as educational psychologists, as an additional resource.
- There are substantial 'manageability' issues in guaranteeing such a system across the country including the finding of time by all the professionals concerned, the costs of professional development and the expense to schools and parents of psychologists and educational psychology services.
- There are substantial 'manageability' issues for QCA and NAA, given their responsibilities under the Disability Discrimination Act 2005, in overseeing the equitable roll-out of such developments.
- The testing of speed of response is not a specified aim of the assessments
- Allowing all pupils the possibility of extra time might offer a way forward. The Irish authorities already have such a system for 16+ assessments, though there is some anecdotal concern that this benefits able pupils at least as much as those with special educational needs.

It may be that, like democracy, the use of extra time in its present form as an adjustment is the 'least worst' option. But it would be useful to explore the options through well-designed studies.

Another way forward could be offered by the growth of computer-assisted assessments. If these are in standard use in the future it would be relatively easy to allow students to use earphones connected to their computer to hear, as often as they needed to, general instructions for the English (reading) tests and general and specific instructions for mathematics and science tests. A recommendation has gone to the authorities responsible for the Irish Leaving Certificate (Costello 2007) that permitted instructions should be available in a format appropriate to a pupil's iPod or similar MP3 player. Without disturbing other pupils, the candidate could listen to the instructions on their earphones as often as they wish.

Research in the many American states which are trying standard computer-assisted assessments has established that whether or not students benefit from the medium depends on whether they feel more comfortable with it than with handwriting. We are far from an adequate understanding of the inter-relation of the standard use of computer-assisted assessment for all and the needs of pupils with SEN and/or disabilities. This does not, of course, apply to the use of assistive technology which is used routinely in class and assessments by some pupils. This is an area where further research could usefully delineate more of a rather murky picture.

There is evidence of concern about the emotional pressures created by national curriculum assessments. In general terms, this is beyond the scope of this review.

There will, however, be pupils with mental health disabilities taking the assessments. They have entitlements within disability discrimination legislation which the QCA will wish to ensure it is addressing appropriately. This area could be conveniently considered alongside issues relating to those with autism spectrum conditions, such as Asperger syndrome, about whom little has appeared in this context.

Conclusion

Most of the reasonable adjustments now in use in National Curriculum Assessments are unproblematic and the guidance addresses them appropriately.

The evidence suggests further consideration would be helpful in relation to:

- Extra time as an adjustment (see above)
- Areas such as autism, speech, language and communication and mental health which have had less attention in the past
- Adjustments where an adult provides the interface to the presentation and response. Studies comment on the importance of professional development for teachers implementing reasonable adjustments and an opportunity to embed this exists with the Government's commitment to accreditation for all Special Educational Needs Co-ordinators in schools. If DCSF and the Training and Development Agency for Schools were to include reasonable adjustments in relation to assessment in any accredited professional development programme, all schools would have a senior leader who had received training in that area.

Finally, the spirit of the disability discrimination legislation suggests that the QCA should strive towards making all its assessments as inclusive as possible, rather than emphasising 'access' through reasonable adjustments.

Reporting on reasonable adjustments each year could usefully include a section on desirable changes to the tests themselves, as well as any necessary development of the adjustments.

References

Allcock P (2001). *Assessment of Handwriting Speed,* Patoss (cited in Backhouse G and Morris K (eds) (2005). *Dyslexia? Assessing and reporting: the Patoss Guide.* London: Hodder Murray.
Backhouse G and Morris K (2005). *Dyslexia? Assessing and reporting: the Patoss Guide.* London: Hodder Murray.
Backhouse G, Dolman E, Read C (2004). *Dyslexia: Assessing the need for Access Arrangements during Examinations: a Practical Guide*. Evesham: Patoss.
Bartlett D and Peacey N (1992). Assessments – and Issues – for 1992. *British Journal of Special Education 19 (2) 94-97.*
Bolt SE and Thurlow ML (2004). Five of the Most Frequently Allowed Testing Accommodations in State Policy, *Remedial and Special Education* 25 (3)141-152.
British Association of Teachers of the Deaf (BATOD) (2007) Sign Advice.
Buehler KL (2002). Standardized group achievement tests and the accommodation of additional time (doctoral dissertation, Indiana State University 2001) cited in **Johnstone CJ, Altman J, Thurlow ML, Thompson SJ** (2006). *A Summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45) Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Tech45
Cabinet Office (2005). *Improving the Life Chances of Disabled People.* London: Cabinet Office.
Cahalan-Laitusis (2004*) Accommodations on High-stakes Writing Tests for Students With Disabilities ETS* Princeton NJ
Conti-Ramsden G, Knox E, Botting N, Simkin Z (2002) Education guidelines and National Curriculum Key Stage 2 test outcomes of children with a history of specific language impairment *British Journal of Special Education* 29(2) 76
Collins E (2003). 'It's really hard, this dictation business'; observations on the use of an amanuensis in examinations. *Support for Learning* 18 (2) 66-70
Connelly V, Dockrell JE and Barnett J (2005). The Slow Handwriting of Undergraduate Students Constrains Overall Performance in Exam Essays *Educational Psychology* 25 (1) 99-107.
Corker M and Shakespeare T (2002). Mapping the terrain in M Corker and T Shakespeare (eds) *Disability/postmodernism,* London: Continuum.
Costello, M (2007). Personal communication: National Educational Psychological Service, Ireland.
Crawford L and Tindal G (2006). Policy and Practice: Knowledge and Beliefs of Education Professionals Related to the Inclusion of Students with Disabilities in a State Assessment. *Remedial and Special Education* 27 (4) 208-217.
Desforges C (in press). *Children in the secondary school age range with moderate special education needs (MLD): a research review of aspects of provision and experience.* The Esmee Fairbairn Foundation.
Destefano L, Shriner JG, and Lloyd CA (2001). Teacher Decision Making in Participation of Students With Disabilities in Large-Scale Assessment, *Exceptional Children* 68 (1) 7-22.
DES and Welsh Office (1987). *National Curriculum Task Group on Assessment and Testing: a Report*
DfEE/QCA (1999). *The National Curriculum: handbooks* London: Qualifications and Curriculum Authority

DfES (2007). *Statistical First review: Special Educational Needs SFRSEN 2007*
DfES (2005). *Data Collection by Type of Special Educational Need.* Annesley: DfES Publications
DfES (2001). *Special Educational Needs: Code of Practice.* Annesley: DfES Publications.
DiCerbo KE, Stanley E, Roberts M and Blanchard J (2001). Attention and standardized reading test performance: Implications for accommodation cited in Thompson S, Blount A, Thurlow M (2002). *A Summary of Research on the Effects of Test Accommodations: 199 through 2001.* (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Technical34.html
Dockrell J and Lindsay G (2001). Children with specific speech and language difficulties – the teachers' perspective. *Oxford Review of Education* 27 (3) 369-394.
Dockrell JE, George R, Lindsay G and Roux J (1997). Problems in the Identification and Assessment of Children with Specific Speech and Language Difficulties. *Educational Psychology in Practice* 13 (1) 29-38.
DRC/DfES (2004). *Disability Discrimination Code of Practice for Schools* Disability Rights Commission
Elliott SN (1999). *Valid Testing Accommodations: Fundamental Assumptions and Methods for Collecting Validity Evidence.* Wisconsin Center for Education Research: University of Wisconsin: Madison.
Elliott SN and Roach AT (2002). *The Impact of Providing Testing Accommodations to Students with Disabilities.* World Wide Web: www.wcer.wisc.edu/testacc
Elliott SN and Marquart AM (2004). Extended Time as a Testing Accommodation: Its Effects and Perceived Consequences. *Exceptional Children* 70 (3) 349-367.
Elliott SN and Marquart AM (2003) *Extended Time as an Accommodation on a Standardized Mathematics Test: an Investigation of Its Effects on Scores and Perceived Consequences for Students With Varying Mathematical Skills* (WCER working paper) cited in Twist L, Donahue B, Lewis K, Keogh N (2006). *Report on the impact of additional time in National Curriculum tests.* NFER.
Fletcher JM, Francis DJ, Boudousquie A, Copeland K, Young V, Kalinowski S, Vaughn S (2006). Effects of Accommodations on High-Stakes Testing for Students With Reading Disabilities. *Exceptional Children* 72 (2) 136-150.
Foulkes G (2003). 'SENDA and Equal Access to Assessment and Qualifications'. Access to Assessment and Qualifications Conference, Cardiff, 21.3.03.
Fuchs LS, Fuchs D, Eaton SB, Hamlett C, Binkley E, Crouch R (2000). Using Objective Data Sources to Enhance Teacher Judgments About Test Accommodations, *Exceptional Children* 67 (1) 67-81.
Hansen EG, Lee MJ, Forer DC (2002). A 'self-voicing' test for individuals with visual impairments. *Journal of Visual Impairment and Blindness* 96 (4) 273-275 cited in **Johnstone CJ, Altman J, Thurlow ML, Thompson SJ** (2006). *A Summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45) Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Tech45
Hedderly R (1996). Assessing Pupils with Specific Learning Difficulties for Examination Special Arrangements at GCSE, 'A' Level and Degree level. *Educational Psychology in Practice 12 (1) 36-44.*

Helwig R and Tindal G (2003). An Experimental Analysis of Accommodation Decisions on Large-Scale Mathematics Tests, *Exceptional Children* 69 (2) 211-225.
Huynh H, Meyer JP, Gallant-Taylor D (2002). Comparability of scores of accommodated and non-accommodated testings for a high school exit examination of mathematics. Paper presented at the annual meeting of the National Council on Measurement in Education cited in Johnstone CJ, Altman J, Thurlow ML, Thompson SJ (2006). *A Summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45) Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Tech45
Jackson LM (2003). The effects of testing adaptations on students' standardized scores for students with visual impairments in Arizona (doctoral dissertation, University of Arizona 2003) cited in Johnstone CJ, Altman J, Thurlow ML, Thompson SJ (2006). *A Summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45) Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Tech45
Johnstone CJ, Altman J, Thurlow ML, Thompson SJ (2006). *A Summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45) Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Tech45
Joint Council for Qualifications (2006). *Access Arrangements and Special Consideration Regulations and Guidance Relating to Candidates who are Eligible for Adjustments in Examinations.* Available: http://www.jcq.org.uk/access-arrangements/ [13.3.07].
Koretz D (1997) *The assessment of students with disabilities in Kentucky (CSE Technical Report No 431)* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing cited in Pitoniak MJ and Royer JM (2001). Testing Accommodations for Examinees With Disabilities: A Review of Psychometric, Legal, and Social Policy Issues. *Review of Educational Research* 71 (1) 53-104.
Koretz D and Hamilton L (1999) *Assessing students with disabilities in Kentucky: the effects of Accommodations, Format and Subject (CSE Technical Report No 498)* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing
Koretz D and Hamilton L (2001). The Performance of students with disabilities on New York's Revised Regents Comprehensive Examination in English (Technical Report 540) Los Angles cited in Thompson S, Blount A, Thurlow M (2002). *A Summary of Research on the Effects of Test Accommodations: 199 through 2001.* (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Technical34.html
Landau S, Russell M, Gourgey K, Erin JN, Cowan J (2003). Use of talking tactile tablet in mathematics tsting. *Journal of Visual Impairment and Blindness* 97 (2) 85-96 cited in Johnstone CJ, Altman J, Thurlow ML, Thompson SJ (2006). *A Summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45) Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Tech45

Lindsay G and Dockrell J (2000). The Behaviour and Self-esteem of Children with Specific Speech and Language Difficulties. *British Journal of Educational Psychology* 70, 583-601.
Lindsay G, Pather S and Strand S (2006). *Special Educational Needs and Ethnicity: issues of over- and under-representation: research report 757* London: Department for Education and Skills
Lloyd-Bennett P (1994). Research Initiatives and Current Thinking on Special Arrangements in Examinations for Pupils with Special Educational Needs. *Educational Psychology in Practice* 10 (2) 75-81.
Marsh L (1995). Educational Psychologists' Reports Concerning Special Arrangements in GCSE Examinations for Candidates with Specific Learning Difficulties. *Educational Psychology in Practice* 11 (1) 19-28.
Mortimore P (2006). Education reform: which way forward? *Education Review* 19 (6).
National Assessment Authority (NAA) (2007). *KS3 English specification 2008,* National Assessment Authority.
Phillips SE (1994). High stakes testing accommodations: validity versus disabled rights, *Applied Measurement in Education* 7 (2) 93-120.
Pitoniak MJ and Royer JM (2001). Testing Accommodations for Examinees With Disabilities: A Review of Psychometric, Legal, and Social Policy Issues. *Review of Educational Research* 71 (1) 53-104.
Qualifications and Curriculum Authority (2006). *National Curriculum Assessments: Regulatory Framework.* London: QCA.
Qualifications and Curriculum Authority (QCA 2006b). *Disability Equality Scheme* London: QCA
Qualifications and Curriculum Authority (2006c). *Assessment and reporting arrangements* 2007 Key Stage 2. London: QCA.
Qualifications and Curriculum Authority (2006d). *Assessment and reporting arrangements 2007 Key Stage 3*. London: QCA.
Qualifications and Curriculum Authority (2006e). *National Curriculum Assessments: Code of Practice 2007*. London: QCA.
Riddell S and Weedon E (2006). What counts as a reasonable adjustment? Dyslexic students and the concept of fair assessment. *International Studies in Sociology of Education* 16 (1) 57-73.
Sireci SG, Li S, Scarpati S (2003). *The Effects of Test Accommodation on Test Performance: A Review of the Literature.* Center for Educational Assessment Research Report no 485. Amherst MA: University of Massachusetts.
SEAC (1992) *School Examinations and Assessment Council: Handbook Key Stage 1 1992* London: SEAC
Skidmore P (2003). *Beyond Measure: why educational assessment is failing the test.* London: Demos.
Schulte AA, Elliott SN and Kratochwill TR (2001). Effects of testing accommodations on students' standardized mathematics test scores: an experimental analysis, *School Psychology Review* 30, 527-547.
Tindal G and Fuchs L (1999). *A Summary of Research on Test Changes: An Empirical Basis for Defining Accommodations.* Mid-South Regional Resource Center: University of Kentucky.
Twist L, Donahue B, Lewis K, Keogh N (2006). *Report on the impact of additional time in National Curriculum tests.* NFER.

Thompson SJ, Johnstone CJ, Thurlow ML (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 16.3.07. from the World Wide Web: http://education.uumn.edu?NCEO/OnlinePubs?Synthesis44.html
Thompson S, Blount A, Thurlow M (2002). *A Summary of Research on the Effects of Test Accommodations: 199 through 2001.* (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Technical34.html
Thurlow ML and Bolt S (2001). Empirical support for accommodations most often allowed in state policy (Synthesis Report 41) University of Minnesota cited in Elliott SN and Roach AT (2002). *The Impact of Providing Testing Accommodations to Students with Disabilities.* World Wide Web: www.wcer.wisc.edu/testacc
Thurlow ML, Lazarus SS, Thompson SJ, Morse AB (2005). State Policies on Assessment Participation and Accommodations for Students with Disabilities. *The Journal of Special Education* 38 (4) p232-2240.
Walz L, Albus D, Thompson S, Thurlow M (2000). Effect of a multiple day tst accommodation on the performance of special education students (Minnesota Report No 34) Minneapolis cited in Thompson S, Blount A, Thurlow M (2002). *A Summary of Research on the Effects of Test Accommodations: 199 through 2001.* (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 15.3.07. from the world wide web: http://education.umn.edu/NCEO/OnlinePubs/Technical34.html
Woods K (2000). Assessment Needs in GCSE Examinations: some student perspectives. *Educational Psychology in Practice* 16 (2) 131-140.
Woods K (2004). Deciding to provide a 'reader' in examinations for the General Certificate of Secondary Education (GCSE): questions about validity and 'inclusion'. *British Journal of Special Education* 31 (3) 122-127.

Appendix 1: Notes on the law

Disability Discrimination Act (1995) amended 2001: not to treat disabled pupils less favourably, to take reasonable steps to avoid putting disabled pupils at a substantial disadvantage. A disabled person has a physical or mental impairment which has an effect on his or her ability to carry out normal day to day activities.

The SEN Framework (Education Act 1996) places duties on schools to make provision for the special educational needs of individual children.

Glossary

ARA	Assessment and reporting arrangements
ASD	Autism spectrum disorder
BESD	Behavioural, emotional and social difficulty
DDA	Disability Discrimination Acts (1995 revised 2001, 2005)
DES	Department of Education and Science
DfEE	Department for Education and Employment
DfES	Department for Education and Skills
DCSF	Department for Children, Schools and Families
IDEA (US)	Individuals with Disabilities Education Act (1997)
IEP (US)	Individualized Educational Program
IEP (UK)	Individual Education Plan
JCQ	Joint Council for Qualifications
KS	Key Stage (1, 2, 3)
MLD	Moderate learning difficulty
MLP	Modified Large Print
NAA	National Assessment Authority
NCA	National curriculum assessments
NCEO (US)	National Center on Educational Outcomes
NCLB (US)	No Child Left Behind Act (2001)
NDCS	The National Deaf Children's Society
QCA	Qualifications and Curriculum Authority
RA	Reasonable adjustments
SAT (US)	Scholastic Assessment Test
SCAA	School Curriculum and Assessment Authority
SEAC	School Examinations and Assessment Council
SEN	Special educational needs
SENDA	Special Educational Needs and Disability Act 2001
SpLD	Specific learning difficulty
TDA	Test development agency

