Including pupils with special educational needs and disability in national assessment

Douglas, Graeme; McLinden, Michael; Robertson, Christopher; Travers, Joseph; Smith, Emma

DOI:
10.1080/1034912X.2015.1111306

Document Version
Peer reviewed version

Citation for published version (Harvard):
Douglas, G, McLinden, M, Robertson, C, Travers, J & Smith, E 2016, 'Including pupils with special educational needs and disability in national assessment: comparison of three country case studies through an inclusive assessment framework', International Journal of Disability, Development and Education, vol. 63, no. 1, pp. 98-121. https://doi.org/10.1080/1034912X.2015.1111306

Including pupils with special educational needs and disability in national assessment: Comparison of three country case studies through an inclusive assessment framework

Pre-publication copy: article will be published in 2016, Volume 63, Number 1 of the *International Journal of Disability, Development and Education*.

Abstract

The assessment of educational progress and outcomes of pupils is important to all concerned with education. This includes testing which is undertaken for accountability and award bearing purposes. This paper examines how students with special educational needs and disability (SEND) are included in assessment. An 'inclusive assessment' framework is outlined based around three core features: (1) all students are included and benefit from assessment; (2) assessments are accessible and appropriate for the diverse range of children in the education system; and (3) the full breadth of the curriculum is assessed (including curriculum areas of particular relevance to students with SEND). Assessment policies and practice in three countries (England, Ireland and the US) are drawn upon to demonstrate how the framework usefully enables between-country comparisons and within-country analysis. This analysis shows that in comparison to Ireland, the US and England have highly developed system-based approaches to assessment which seek to 'include all' (feature 1) and be 'accessible and appropriate' (feature 2). However, the analysis highlights that a consequence of such assessment approaches is the narrowing of the curriculum around topics that are assessed (most notably literacy and mathematics). Such approaches therefore may be at the expense of wider curriculum areas that have value for all students, but often of particular value for those with SEND (feature 3). It is argued that within such systems there may be a danger of neglecting the third feature of the inclusive assessment framework, i.e. ensuring that the full breadth of the curriculum is assessed. A consequence of such an omission could be a failure to assess and celebrate progress in relation to educational outcomes that are relevant to a diverse range of students.

Key words: inclusive assessment, special educational needs, disability, national assessment

Authors:

Graeme Douglas[1], Reader in Disability and Special Educational Needs
Mike McLinden[1], Professor of Education
Christopher Robertson[1], Lecturer in Inclusive and Special Education
Joseph Travers[2], Director of Special Education
Emma Smith[3], Professor of Education

[1] Department of Disability, Inclusion and Special Needs
School of Education
University of Birmingham
Birmingham, B15 2TT
UK

[2] Department of Special Education
St. Patrick's College
Drumcondra
Dublin 9
Republic of Ireland

[3] School of Education
University of Leicester
21 University Road
Leicester, LE1 7RF
UK

Introduction

Assessment approaches that sensitively measure students' educational progress and outcomes are crucial for those concerned with understanding how educational systems are serving young people. Sheil *et al* (2010) describe national assessments as fulfilling a variety of purposes including the award of qualifications to individual students. They note that such assessments also have an important monitoring function such that 'the focus of interest is on the aggregation of data collected from the students, not on the performance of individual participating students' (p34). This function enables policy makers to use the data to answer specific questions such as: How well are students learning with reference to general expectations, the aims of the curriculum or preparation for life? or Do particular sub-groups in the population perform poorly? Sheil et al (2010) contrast national assessments with international and classroom assessments: international assessments have many of the features of national assessments but aim to provide information on standards of student achievement in a number of countries, and individual countries can compare student performance against that in other countries (e.g. PISA carried out by OECD); classroom assessment includes formative and summative assessments carried out by teachers to support individual student learning.

This article is concerned with how students with special educational needs and disability (SEND)[1] are included in approaches to national assessment. It draws upon and develops the findings from an earlier review undertaken by some of the authors (Douglas et al, 2012). We begin by considering how educational outcomes for students (including those with SEND) can be usefully conceptualised. We then present a framework of 'inclusive assessment' which is used to consider the assessment practices in three countries during 2012 (England, Ireland and the US). Through this analysis we examine examples of inclusion and exclusion of pupils with SEND from the assessment approaches with reference to each feature of the inclusive assessment framework. Drawing on this analysis we reflect upon the

[1]Different countries use different terms to refer to this broadly defined group of students. In this article we use the term 'special educational needs and disability' (SEND) to navigate this ambiguity, although we accept that different countries not only use different terms, but also have different definitional boundaries (e.g. 'who counts' as having a disability associated with low attainment). It is also worth noting that SEND has gained formal use in recent policy and legal frameworks in England, e.g. Children and Families Act 2014 and the publication of the associated 2014 'Special Educational Needs and Disability Code of Practice: 0 to 25 years'. SEND is used generally in the article, except when we discuss countries specifically when we use the term most commonly used in that country: students with disability in the USA and students with special educational needs (SEN) in England and Ireland.

balances and tensions in assessment approaches that seek to be ‘inclusive’ and consider the implications for assessment practice.

Educational outcomes and SEND

A broad aim of those concerned with educational policies is to design and implement educational systems (inputs and processes) that support the development of the students (outcomes). An example of this conceptualisation is the ‘input-process-outcomes’ model of effectiveness which has been utilised by the European Agency for the Development of Special Needs Education (EADSNE). As part of the EADSNE’s work, Kyriazopoulou and Weber (2009, pp14-15) argue that educational input and resources denote all aspects provided to the system: e.g. financial resources, legislation, qualified teachers and infrastructure. Education processes transform these inputs and resources into outputs and outcomes which include pupil participation rates and curricular achievements: e.g. academic and functional literacy, independence or citizenship. In our previous international review of educational outcomes (see Douglas et al, 2012) we found that outcomes can be usefully grouped into ‘attainment-related’ outcomes (commonly concerned with traditional curriculum areas) and ‘wider curriculum-related’ outcomes (which commonly include wellbeing and independence-related outcomes). While different national systems may emphasise different types of outcomes, and conceptualise their curriculum in different ways, this broad distinction was found to be helpful in our review. Importantly, the review found that different countries gather evidence in relation to these different outcomes in a variety of ways and to different extents drawing upon classroom, national and international assessments.

Of relevance to this article is the extent to which particular assessment approaches include students with SEND, and at the heart of this is the concept of ‘inclusive education’. However, inclusive education itself is a much debated and ambiguous concept. Norwich (2013, p3) reviews definitions of inclusive education presented in the literature and identifies nine key themes which include: not leaving anyone out; accepting/valuing/extending scope to all; active participation in school life; and enhancing equal opportunity. Norwich notes that these themes appear to overlap and can even be inconsistent with one another. In part, it is these theoretic concerns that lead him to focus his analysis upon ‘dilemmas of difference’ (Norwich, 2008, 2013). Central to his theoretical position is how inclusive education must navigate education systems that contain dilemmas: most notably education systems that treat people the same on one hand (but might be insensitive to their differences) and respond to people’s difference on the other (which might stigmatise and hinder them

on that basis) (Minow, 1990 in Norwich, 2013, p7). Norwich applies this analysis to education in relation to:

1. identification-non-identification of some children having SEND;
2. curriculum commonality-differentiation ('inclusive curriculum'); and
3. common-separate teaching and learning settings ('inclusive pedagogy' and 'inclusive schooling').

It is interesting nevertheless that Norwich does not directly extend his analysis to include assessment, and it is the development and exploration of the concept of 'inclusive assessment' that is a key aim of this paper.

Inclusive assessment – who, how and what?

In this section we develop and define an inclusive assessment framework. Drawing upon Watkins' (2007) early development of the term 'inclusive assessment', we unpick three inter-related features of the concept: *who* is assessed, *how* they are assessed and *what* is assessed?

Who is assessed?

Watkins (2007) presents an analysis of assessment processes in mainstream primary school settings in 23 countries. In offering a definition of inclusive assessment, Watkins (2007, p47) notes: 'The overall goal of inclusive assessment is that all assessment policies and procedures should support and enhance the successful inclusion and participation of all pupils vulnerable to exclusion, including those with [SEND].' So a central theme of inclusive assessment is in relation to 'who' assessment serves: i.e. the inclusion of *all* children (including those with SEND) in assessment practice. Inclusion of all in educational practice is in line with the spirit of international declarations (e.g. UN Convention on the Rights of Person with Disabilities 2006, under articles 7 and 24) and national declarations (e.g. EPSEN Act [Ireland], Equality Act [UK], and IDEA [US]). Nevertheless, this is necessarily context-dependent because different countries have different approaches to assessment generally.

Extending the idea of including all, it is argued that disaggregation of assessment data associated with students with SEND is an important and inclusive practice. First, it is important in understanding how the educational system is working as a whole. By including variables on a range of pupil characteristics including SEND, as well as other characteristics of relevance such as gender, poverty and ethnicity, overall engagement, progress and outcomes can be better understood, and resources and interventions better targeted. Second, there is the social justice

argument. If there is a particular concern for a certain group or groups it seems imperative to seek ways of carrying out necessary monitoring so that levels of inequality can be reduced or eliminated. Third, technical reasons in relation to the design of the assessment procedures might require such disaggregation – certain groups may require particular versions of assessments.

How they are assessed?

Watkins also discusses some technical aspects of 'how' assessment should take place. There is a rich literature associated with modifying assessments to accommodate a diverse range of pupils. In the context of national assessment practice, Greenen and Ysseldyke (1997, pp226-227) provided an early analysis of US accountability systems in terms of their relative inclusion of students with SEND. Adjusting the vocabulary to include more recent literature, different techniques to the inclusion of students with SEND in national assessment can be defined as follows:

- *Total exclusion:* in which assessment does not take place for these students.
- *Alternative systems:* in which *alternative assessments* are developed to include students with SEND. An alternative assessment approach aims to ensure that all pupils, irrespective of their ability, can be assessed appropriately by creating a range of assessments with different assessment criteria. Using the US as an example, Cameto et al (2009) described an alternative achievement standard as 'an expectation of performance that differs in complexity from a grade-level achievement standard' (pp2-3). The measurement of student performance against these standards requires an appropriate alternative assessment. Lowrey et al (2007) defined this as 'an assessment tool for students with disabilities that is used in place of the statewide assessment' (p245). In considering its use they noted that most of these students have need of an alternative assessment (note: 'alternate' is used for 'alternative' in the US):

  > … because of their inability to respond to the format and content of the statewide assessment. That is, the required response mode, context, and content of the statewide assessment may be too challenging or may be inappropriate for students with severe disabilities. Alternate assessment allows for different modes of responding, a different context of assessment, and different content that is still linked to statewide standards. (p245)

- *Accommodated systems:* in which assessment can be modified to include students with SEND. Modifications, or accommodations, are required because standard assessment formats and procedures can present barriers to pupils with SEND, which means they may not be able to demonstrate their abilities under normal assessment conditions. These assessment accommodations seek to make an assessment accessible while maintaining the same assessment criteria (e.g. Hopper, 2001; QCA, 2007; Lazarus et al, 2009). Accommodations can take various forms, e.g. in relation to assessment presentation, response method, setting and scheduling (Hopper, 2001).

- *Universally designed systems*: in which a single assessment method is suitable for all students. The approach argues that a careful attention to assessment design will include all and reduce the need for accommodated and alternative versions (e.g. Lazarus et al, 2009). This seems an important aspiration – with attention to such principles, the need for accommodated and alternative assessments could be reduced; although, we could not find examples of any universally designed national assessments that include all.

What is assessed?

Watkins (2007) also highlights that inclusive assessment 'should aim to "celebrate" diversity by identifying and valuing all pupils' individual learning progress and achievements' (p48) and that 'a wide range of assessment methods are necessary in inclusive assessment in order to make sure that there is a wide coverage of areas (non-academic as well as academic subjects) assessed' (p49). Importantly, this goes beyond the technical analysis of *how* assessment should take place as offered in the previous section. Rather it focuses upon 'what' should be assessed and this has a key relationship with the curriculum.

This relationship between assessment and curriculum is obviously crucial, but defining what should be assessed in an international context is somewhat ephemeral because curricula are (1) linked to country and cultural context, and (2) particular to, or of particular concern to, specific SEND groups. With regards to the first point, national education systems will to some extent define outcome priorities by specifying the curriculum to be taught. Different countries do this to different extents, e.g. England defines a national curriculum; Ireland and the US do not have a national curriculum in the same way but national examinations (Ireland) and learning standards (US) have similar impacts. As the country case studies that follow demonstrate, system-based data collection commonly focuses on attainment-related outcomes regarding specific parts of the curriculum (especially literacy, numeracy and science). A concern then is that assessments reflect the full breadth of a curriculum, rather than just part of it.

Regarding the second point, and overlapping with the previous paragraph, this 'full breadth' includes progress and outcomes identified as relevant to people with SEND. To this extent some assessments do not include all people in the education system. Instead some assessments are applied to those for whom it is particularly relevant, to ensure that their broad and diverse needs are recognised and monitored. More controversially, perhaps, is what Norwich (2013, pp63-65) describes as the common versus differentiated curriculum tension; and related to this the concept of additional

or expanded curricula, sometimes called alternative curricula. This latter term is specifically avoided in this article as it can be associated with exclusion from a mainstream/core curriculum (e.g. a child with learning disabilities being denied access to a mainstream curriculum because he or she is not seen as capable). Here the interest is in the teaching of specific skills that are seen as important (e.g. mobility for children with visual impairment, social skills to children with autism, independent-living skills for children with learning disabilities), and as such may warrant particular assessment and monitoring. Douglas et al (2012, pp37-39) present a summary of additional curriculum areas of relevance to particular SEND groups. Even so, the concept can also be hard to clearly define because different countries may have differing definitions of what constitutes a common educational curriculum (and therefore, what is defined as 'additional'). It can also be difficult to define because of the wide range of additional curriculum provision that may be regarded as beneficial to the learning of particular student groups who may be included under the SEND umbrella.

Inclusive assessment framework

From the above analysis, inclusive assessment is seeking to be inclusive in relation to 'who is assessed', 'how they are assessed' and 'what is assessed'. Our proposed inclusive assessment framework incorporates these three features (see Figure 1) – assessment should: (1) include all; (2) be accessible and appropriate; and (3) assess and report areas of relevance.

[Insert Figure 1 here]

1) *Assessments should include all children and young people.* Different countries assess and collate pupils' educational progress and outcomes in different ways. Within an inclusive assessment approach, assessments should be carried out for *all* children and young people. The data generated from such assessments should be appropriately disaggregated as required and as is useful (e.g. to show outcomes for different SEND groups).
2) *Assessments should be accessible and appropriate for those being assessed.* Procedures should be designed to include the diverse range of pupils within the educational system. For children with SEND, assessments should be *accessible* through suitable accommodation and *appropriate* through suitable breadth of assessment and range of criteria.
3) *Assessments should measure and report areas of relevance.* Assessments should seek to measure progress and outcomes on the full breadth of the curriculum that an education system offers. The inclusion of a diverse range of students within the educational system means it will be necessary to assess areas of specific relevance to people with SEND (in some countries this is

referred to as a wider or additional curriculum). Therefore, systems also need to be in place to record educational progress on these areas that may be of particular interest or concern to given stakeholders. Examples include aspects of mobility and the use of specialist technology for those with physical and sensory disabilities.

Method of analysis

Our review had four stages. First, the authors constructed case studies for each of the three countries (England, Ireland and the US) based upon policy context; key national approaches to measuring outcomes and progress; and details of the assessment approaches with a focus upon accommodation and alternative assessment policy, and breadth of the curriculum assessed. The countries of England and the US were chosen because they have traditions of national assessment policy, are English speaking and have significant associated literature. Ireland was chosen because it was the country of interest for the original project funder (see acknowledgements) and was in the process of undergoing significant development in educational assessment. Literature was gathered through various sources, but most notably ERIC database searches; internet-based policy documents; research and administrative reports, often government funded or carried out by government departments; academic literature (generally policy oriented); and telephone and email enquiries with relevant contacts. Table 1 presents a summary of the key terms, policies, and prevalence and school placement figures in relation to SEND for each country. Table 2 presents the education year group labels for each country.

[INSERT TABLE 1 AND 2 ABOUT HERE]

In the second stage, the initial review was reorganised with attention to the inclusive assessment framework presented in the previous section. Table 3 (England), Table 4 (Ireland) and Table 5 (US) present a summary analysis of each country's inclusive assessment approaches. These tables present key information about national assessments in the three countries; specific information about student age and timings of assessment can be gathered from Table 2.

[INSERT TABLE 3, 4 AND 5 ABOUT HERE]

Third, the authors compared the approaches taken across the three countries to draw out contrasting inclusive assessment practice. Given the different nature of the three countries, some of this discussion was linked to the contrasting national

assessment policies that the countries had (or did not have) in place, while some discussion was focused upon the different implementation of these policies. This analysis cross references to Tables 3, 4 and 5.

Finally, the discussion draws out overarching themes that emerged from the analysis. The discussion also reflects upon the utility of the proposed inclusive assessment framework.

Cross-country comparison of inclusive assessment practice

In the following section we make comparisons across the three countries, cross-referencing to evidence presented in Tables 3, 4 and 5.

Educational standards assessments

The standards-based education approach adopted in the US and England, and the related culture of national testing, differs from the approach taken in Ireland. Indeed, in their review of standardised testing in Ireland, Sheil *et al* (2010) excluded the US and England from their analysis 'because the kind of high stakes testing being carried out in those countries did not seem appropriate, or acceptable, in an Irish context' (p19). Nevertheless, on the inclusion of students with SEND, the US and England offer an interesting and relevant insight. Both countries draw on standards-based education principles in which all pupils are assessed against national educational standards at various points during their school career. England uses standard attainment tests (SATS) [Table 3: 1, 2] and individual US states use standardised assessments [Table 5: 1]. It is the inclusion of all or most students that is relevant here because it has led to innovative development of accessible and appropriate assessment for students with SEND. Both countries provide accommodated and alternative versions of national assessments [Table 3: 7, 8; Table 5: 5, 6]. In England, performance (P) scales are used in schools for recording the achievement of pupils with SEND who are working towards the first level of the English national curriculum (level 1) [Table 3: 8]. In the US, individual states must provide alternative assessments for students working below that state's 'grade-level achievement standard' [Table 5: 6]. Therefore, in each country the progress on these attainment-related measures is recorded for students with SEND, including those with the most complex cognitive disabilities. Both countries also link this assessment data to demographic data so that results can be disaggregated by different SEND groups [Table 3: 5, 6; Table 5: 3].

Ireland has traditionally not carried out such assessments at a national level. However, the National Literacy and Numeracy Strategy (DES, 2011) includes proposals for more assessment at various levels. There are associated requirements for schools to report pupil results to parents and to report aggregated results to the DES. The first implementation (May/June 2012) of the standardised tests did not include all children with SEND [Table 4: 1, 2].

Award-bearing assessments

England and Ireland have award-bearing nationally set external examinations which are broadly equivalent (GSCE and Junior Certificate respectively). Each country collates national data on these [Table 3: 2, 5; Table 4: 3]. While England collects and presents data in a form that offers the possibility for disaggregation on different SEND groups, Ireland does not [Table 3: 5, 6].

The US operates a system by which individual schools assess pupil performance against the national or state standards. Individual states are obliged to report high-school graduation rates disaggregated by category of SEND, although breakdown into sub-groups is not possible in all states (Altman et al, 2010) [Table 5: 2, 3, 8].

National sample-based assessments

In contrast to England, both the US and Ireland conduct sample-based national assessments on different curriculum areas. The US carries out these assessments annually and focuses upon a range of curriculum areas. Although some assessment accommodations exist to enable many students with SEND to participate, recording and subsequent analysis of data in relation to SEND does not appear to take place. Further, the standardised assessments used in these national assessments do not include alternative versions to assess broader ranges of abilities, and therefore are inappropriate for many pupils with learning disabilities [Table 5: 4, 7].

Ireland carries out smaller-scale and periodic sample-based national assessments, e.g. in 2009 in mathematics and English reading in primary second (7+ years) and sixth (11+ years) classes. The sample for each age group was about 4,000 pupils, although this excluded some students with SEND (e.g. special schools and classes were excluded at the sampling stage) [Table 4: 4, 8].

These sample-based national assessments are conceived for specific purposes (e.g. the 2009 national assessments in Ireland were to ‘establish current reading and

mathematics standards of second and sixth class pupils', Eivers *et al,* 2010). Often the assessments themselves are norm-based and therefore their purposes and usefulness are closely linked to the structure of the sample used, which may explicitly exclude people with more severe learning disabilities. Improving accessibility and clarity about who should, and should not, take part in national assessments should improve the quality of the assessment itself. Although such apparent exclusion of some groups seems at odds with the inclusive assessment approach espoused above (and identified as a concern in the US, e.g. Maxwell and Shah, 2011), this may be pragmatic for sample-based approaches to national assessment. The development of alternative assessment material for the different population of students with more complex learning disabilities may offer a useful insight into their progress, although we did not identify examples of this taking place in the literature for sample-based national assessments.

Inclusivity of access – accommodated and alternative assessments

All three countries have accommodation mechanisms in place to include students with SEND in some national assessments (e.g. QCA, 2007) [Table 3: 7; Table 4: 7; Table 5: 5]. This was most developed in the US (e.g. Lazarus et al, 2009), and least developed in Ireland (although Ireland is engaged in few national assessments, the emerging literacy and numeracy national assessment strategy appears to pay no attention to accommodation, see DES, 2011).

The alternative assessments offer a particularly interesting point of discussion here because the three countries differ to such a large extent. Ireland carries out few national assessments, but alternative assessments are not used in those it does [Table 4: 6-8].

The performance or P scales were introduced in 1998 to enable schools in England to measure attainment and progress of children whose attainment levels could not be recorded through English national curriculum scales (Ndaji and Tymms, 2009) [Table 3: 8]. The use of P scales is statutory when reporting attainment for pupils with SEND who are working below level 1 (QCDA, 2011) in English, mathematics and science. Ndaji and Tymms (2010) offer an evaluation of P scales use in England. They note that although the scales data collection and analysis have been found to be successful 'in the sense that many schools participate in it each year and their comments indicate approval' (p199), there have been questions concerning the data quality. Their analysis of data was based on information collected for 22,506 pupils aged five to 16 classified as having one or a combination of special

educational need and for whom schools had submitted data to the project (in English, maths and science). While the analysis identified poor assessment discrimination between curriculum areas, they concluded that 'the P scales are working' (p208).

However, it is the US that has invested the most effort into the development of alternative assessments (the term 'alternate assessments' is used in the US) and is worthy of further discussion here [Table 5: 6]. Of course, the standards-based education agenda is inextricably linked to the development of these assessments; Lowrey et al (2007) report that NCLB is 'the driving force behind alternate assessment' in the US (p245).

Cameto et al (2009) reported on the progress each state had made in implementing these assessments (the progress report was required by law). They found that a range of assessment approaches had been developed and implemented across the country. Unsurprisingly, a growing corpus of literature exists that examines the validity of selected state alternative assessment approaches both in psychometric and philosophical terms. Certainly, the efforts to assess all students' progress is intended to be inclusive, and given that progress in relation to the NCLB standards (reading, mathematics and science) are assessed for the vast majority of students in the US, it has been successful. Nevertheless, the development and implementation of such approaches is expensive (e.g. Elliot et al, 2008, p151). Perhaps more fundamentally, Lowrey et al (2007) highlight a concern about the use of alternative assessments for students with severe learning disabilities at the expense of other 'meaningful targets that will improve a student's quality of life after leaving the public school system' (p251). To this extent Lowrey et al question whether the alternative assessments miss the point, and do not attend to more important outcomes and progress which are of greater relevance to some young people's lives.

Breadth and relevance of assessment

Much of the discussion up to this point in the article has highlighted that a focus on educational assessment, particularly national assessments, and administrative records are associated with the measurement of attainment-related outcomes and progress – in other words, related to traditional curriculum areas, especially literacy and mathematics. We consider next the outcomes associated with the wider curriculum, linked to the third feature of our inclusive assessment framework.

These outcomes were labelled as happiness- and independence-related outcomes in our international literature and policy review (Douglas et al, 2012). They include more specific outcomes such as resilience, self-esteem, well-being, relationship building, optimism, employment, independent living skills and successful transition after school. This broader analysis of outcomes also included areas of the curriculum and additional curriculum that may be relevant to pupils with SEND.

The analysis shows that such an approach to data collection is relatively rare in the three countries: where national systems of data collection are in place in England and the US they tend to focus on attainment-related measures and both countries rarely collect broader curriculum data for monitoring purposes.

However, the analysis identified some links between system-based data collection and measurement of broader measures of progress and outcomes. For example, the US is unusual in that the policies of No Child Left Behind Act (NCLB) and Individuals with Disabilities Education Act (IDEA) require some data on employment outcomes and disability to be gathered and presented at state level [Table 5: 2, 3, 8]. Of relevance also is the use of a national pupil database in England. The National Pupil Database is central to England's standards-based reform agenda because it is the mechanism for tracking students as they progress through the educational system, enabling attainment-related outcomes data to be monitored and reported. The database contains information on student SEN, therefore disaggregation is possible [Table 3: 5]. However, such a database is not limited to a standards-based reform agenda with its associated reporting strategy (as typified by school league tables in England). Such national pupil databases are powerful because, if implemented appropriately, data gathered from a wide range of sources can be connected together. For example, in the evaluation work undertaken by Humphrey and Squires (2011) in England, the authors drew on attainment and attendance data already collected as part of England's national assessment programme and recorded within the National Pupil Database and combined this with other broader measures of pupil progress collected as part of the evaluation (e.g. developing positive relationships with others; increasing participation in extended services provision, including extra-curricular activities) [Table 3, 10, 11]. Such an approach is efficient because the same data (e.g. SEND status, gender and ethnicity) does not have to be collected many times.

Ireland has traditionally not had a national pupil database, although recent developments suggest moves in this direction: a new primary (POD) and post-

primary online database (PPOD) has recently been introduced which seeks to gather and maintain data in relation to all pupils (e.g. DES, 2014). Current versions of the database have fields for learning support according to categories of low incidence SEND, general allocation model for milder needs, psychological or medical assessment report, and special/mainstream class placement [Table 4, 5].

In the context of this specific discussion it means that outcome measures that are broader than attainment-related outcome measures can be efficiently incorporated into a country's monitoring process. However, this is rarely done, which means insight into this tends to rely on standalone studies rather than national assessments. Of relevance here are research projects of varying sizes and complexity – with different designs (e.g. survey, longitudinal, retrospective and intervention studies) and focusing on different populations (a range of SEND groups or a specific sub-group). These studies can contribute to the understanding of progress and outcomes of students with SEND and they are particularly linked to the third proposed feature of inclusive assessment. The strength of such approaches is that they often seek to measure progress and outcomes that go beyond relatively narrow attainment-related measures. Depending on the scale of the research, these research studies can also be relatively inexpensive. A current large-scale study in Ireland is the Growing Up in Ireland study which is tracking the progress of almost 20,000 children [Table 4: 10].

A particularly powerful and large-scale example of this is the US-based longitudinal study of people with disabilities (NLTS2) [Table 5: 10, 11].The study not only gathered valuable data on attainment and a variety of procedural and experiential topics, but also gathered data on employment and lifestyle outcomes along with well-being. The study provides a wealth of broader outcome data including information on disability-specific areas (e.g. mobility and students with a visual impairment).

Discussion

The inclusive assessment framework we have used in the analysis provided a vocabulary which helped with summarising and comparing the national assessment practice in the three countries; in turn this helped to identify communality and differences. To this extent the framework was a useful descriptive tool. Nevertheless, the framework also helped us reflect upon broader tensions and dilemmas faced by those experiencing and designing national assessments in different countries. We present these reflections in this final discussion.

In the US, the National Council on Disability (NCD) reviewed the progress of US policies NCLB and IDEA and offered a 'progress report' (NCD, 2008). Their analysis was broadly positive, and a particularly strong theme in the report was that the inclusion of students with SEND in national assessment meant that schools were now being held to account for all student progress, including those with SEND:

> According to our analyses, one of the most important results of NCLB and IDEA appears to be that students with disabilities are no longer ignored. To that end, NCLB and IDEA have had a significant, positive impact. (p1).

To this extent, the national assessments in operation in the US and England are demonstrating features of the inclusive assessment framework: all students in those countries are included in assessment (feature 1 of the inclusive assessment framework), and the assessments have been developed to be accessible and appropriate (feature 2 of the inclusive assessment framework). Ireland has far fewer national assessments than England and the US, but the national assessments that exist in Ireland appear to be less inclusive in these regards.

Nevertheless, NCD's positive overall analysis assumes the quality of the national assessments and the associated data produced are adequate. We consider this issue first. Further, the analysis in the previous sections highlighted that national assessments tend to focus upon a narrow part of the curriculum. This is potentially at odds with our assertion that assessments should measure and report areas of relevance (feature 3 of the inclusive assessment framework). We consider if there are any potential consequences of this second issue.

1. Quality of assessment and data produced

As noted, the investment in national assessment design and implementation in England and US has been enormous, and the development of accommodated and alternative versions of these assessments (especially in the US) has been a significant technical challenge and achievement. Even so, concerns exist about the quality of some of the resulting data and the assessments themselves. For example, DfE (2010) reporting on attainment data for SEND sub-groups in England note that a (small) number of young people with 'profound and multiple learning difficulty' had achieved two Advanced Level qualifications or equivalent in 2008 and 2009 (p106). Given the nature of this disability group, it seems impossible that these data are accurate. It is unclear where the error occurred.

Our review presented here draws upon data from 2012 or earlier. Nevertheless, more recently the DfE in England queried the quality of national assessment sub-

levels more generally, arguing that sub-levels of assessment were poorly designed, used subjectively by teachers and schools, and misleading for parents and other schools (at primary-secondary transfer). This led to the abolition of the requirement for sub-level assessment to be carried out by schools in 2013. At time of writing a new commission is in place to review how assessment without sub-levels should take place.

In the US, additional concerns include the practice of grouping together different types of SEND rather than reporting on SEND sub-groups. An interviewee quoted by NCD (2008) sums it up:

> … There is so much lumping together of disabilities, and we need to really differentiate them. NCLB should have more varied testing and accountability standards for students with disabilities given the differences in disabilities. NCLB should be more sophisticated in its requirements for proficiency, not just one standard (National Council on Disability study, 2008, p66).

Under NCLB, states are not required to report separate disability categories, rather all students in special education can be collapsed into a single category. Indeed, in 2009 fewer than half of the states disaggregated results by disability category (primary disability) (Altman *et al* 2010). Further, while disaggregating into primary disability groups is seen as preferable to not doing so, the concept of primary disability can be problematic because many young people have a complex set of multiple conditions (e.g. AFB, 2009).

2. Consequences of national assessments

Literature on unintended consequences of national assessment was identified in relation to the US and England. An important part of the unintended consequences debate is the notion of high stakes assessment. Commonly the latter refers to the importance of the assessment to the given student. Typically assessments that lead to a qualification or award are high stakes in this way. In some countries that implement a standards-based reform agenda, the term can also refer to assessments that have high stakes implications for the teacher or service (often school). SATs (England) and NCLB assessments (US) are examples. The assessments have limited high stakes impact on the student (there is no associated qualification), but they can have an impact on the school (e.g. league table position in England, potential funding in the US) or the teacher (e.g. disciplinary action or promotion opportunities).

Perhaps unsurprisingly, concerns exist about the negative consequences of an increased emphasis on testing and monitoring on the educational experiences of

students generally, and students with SEND in particular. Much of this literature relates to the US where enormous investments have been made to carry out such assessments. Examples of concerns raised include a narrowing of the curriculum, accusations of teachers teaching to the test and an overemphasis on low-level skills (see Darling-Hammond, 2007).

Of particular concern is that the narrowing of the curriculum around topics assessed, most notably literacy and mathematics, may be at the expense of wider curriculum areas that have value for all students, but perhaps particular value for those with SEND where measuring a broader range of achievement may be especially valuable. Indeed evidence presented in the US case study suggested a consequence of including students with severe learning disabilities in national assessments that focus on narrow parts of the curriculum may neglect other meaningful targets linked to functional life skills (e.g. Lowrey et al, 2007).

There is also a concern raised in the literature that assessments can place undue pressure on students with SEND as they may be the students at greatest risk of performing and progressing least well. Related to this, concerns exist that students with SEND are made scapegoats and held responsible for a school's poor performance. Such a view may hinder students with SEND being included within some schools or lead to increased dropout (e.g. Cole, 2006). In England, Galton and Macbeath (2015) make similar observations in their recent review of SEND provision in England and their recommendations refer to the exclusive effects of written examinations and call for the reinstatement of broader curriculum provision (pp13-14).

Tensions and dilemmas

Norwich (2013) argues for the importance of understanding the tensions and dilemmas in inclusive education, and navigating the uncertainties and risks associated with them. Our analysis and use of an inclusive assessment framework has helped us to draw out some of these tensions and dilemmas in relation to national assessments. Arguably, the national assessment strategies in the US and England reflect many of the principles of inclusive assessment, most notably the systems are equitable in intent because all children, including those with SEND, are included. Nevertheless, the previous sections noted some concerns about both the quality of some of the data generated by these assessment strategies and the unintended consequences of assessing narrow aspects of the curriculum. In contrast, the national assessment approaches in Ireland may be less inclusive as far as including all students with SEND through accessible and appropriate assessment

(features 1 and 2 of the inclusive assessment framework), but may suffer less from some of the consequences experienced in England and US in relation to narrowing of the curriculum (feature 3 of the inclusive assessment framework).

In part, the apparently different balances observed in these three countries are a product of contrasting education systems: most notably the standards-based reform approaches in combination with particular reporting and accountability mechanisms in England and particularly the US, which differ from the approaches in Ireland. It follows then that the analysis raises the following dilemmas:

- Does accountability (as per England and US) inevitably lead to narrowing of the curriculum (and therefore threaten feature 3 of the inclusive assessment framework)?
- Does lack of accountability (as per Ireland) inevitably lead to the exclusion of students with SEND from assessments (and therefore threaten features 1 and 2 of the inclusive assessment framework)?

In considering these dilemmas it is important first recognise that the education systems in the three countries can, and do, include assessments beyond the national approaches described (and these may or may not be used for accountability and monitoring purposes). Firstly, a range of one-off or repeated studies that attend to broader areas of the curriculum have an important role to play in this regard and examples from England, Ireland and the US are identified in the article. In the case of England, the national pupil database can enable efficient and powerful analyses to take place. However, with the exception of one example (in the US, linked to the IDEA Part-B indicators), these studies tend not to be used for accountability purposes. Therefore, these studies have limited direct impact upon the day-to-day running of schools and the curriculum emphasis schools adopt.

Secondly, classroom assessment in relation to assessment of broader areas of the curriculum, including disability-specific areas, is crucial. The greater flexibility associated with classroom assessment means that it can be tailored to the specific requirements of individuals or groups. Nevertheless, it is this flexibility and lack of accountability which may mean it does not take place consistently or systematically. Given this article focus is on national assessments, we have not discussed the role of classroom assessment. Even so, classroom assessment warrants more analysis to gain an understanding of national expectations made of schools in terms of recording progress across the wider curriculum.

Thirdly, broader national initiatives may have value in relation to inclusive assessment. A significant example in England was Every Child Matters (ECM), a national initiative (2003) that focused on the well-being of all children from birth to 19. The aim was to ensure children were provided with opportunities to achieve in five broad areas (referred to as the ECM outcomes): Be healthy; Stay safe; Enjoy and achieve; Make a positive contribution; and Achieve economic well-being. At that time significant efforts were made to develop methods that interrogate existing data sources to establish measures against these outcomes for children with SEND (e.g. Morris et al, 2008; Kendall et al, 2008). This interest in measurement of broader outcomes for all children, including those with SEND, generated much interest in England, and arguably the approach attended directly to some of the concerns about narrowness of assessment identified above. Successive UK governments appear to have pulled away from ECM, but it remains an example of how broader national initiatives could (and arguably should) provide a more balanced and inclusive assessment framework.

Conclusion

Continued analysis of England, Ireland and the US will highlight the impact of the dilemmas identified and how each country navigates these over time. Ireland is particularly interesting here: in the short time since we undertook this review, there has been further rollout of national assessments of literacy and mathematics in Ireland, as well as more development of their national pupil database. Time will tell whether this will lead to narrowing of the curriculum in Ireland as observed in England and the US.

The use of the inclusive assessment framework to analyse practice in other countries that operate different national assessment policies would also be useful. For example, Scotland is an interesting case because its current educational reforms include a central role for the Scottish Framework for Assessment as a mechanism for raising standards. However, unlike England and the US where assessment data are collected centrally, Scotland's model involves improving the quality and consistency of classroom assessment and record keeping without system-based national data collection for accountability purposes.

The inclusive assessment framework has proved useful in our analysis and we believe it warrants further development. Perhaps of central importance to an analysis inclusive assessment is that assessment can only be as inclusive as the broader education system will allow. Some countries will put greater value on some aspects

of the curriculum than others (and, ironically, they may use national assessments and accountability mechanisms as a method of emphasising this). This is why analysis of inclusive education needs to consider curriculum issues, and therefore a meaningful analysis of inclusive assessment procedures must do the same.

Acknowledgements

This research draws upon a review which was funded by the National Council for Special Education (NCSE) in Ireland (see Douglas et al 2012). A member of the original research team, Dr Penny Lacey, sadly died in January 2015. We remember our friend and colleague Penny with great fondness and thank her for the huge contribution she has made to the field of education, in particular the education of children with profound and multiple learning disabilities.

References

Altman, J. R., Lazarus, S. S., Quenemoen, R. F., Kearns, J., Quenemoen, M., & Thurlow, M. L. (2010). *2009 survey of states: Accomplishments and new issues at the end of a decade of change*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from www.cehd.umn.edu/NCEO/OnlinePubs/2009StateSurvey.pdf

AfB (American Foundation of the Blind) (2009). *Letter to Sec. of Education: Estimates of Severely Visually Impaired Children, 5 May 2009*. http://www.afb.org/Section.asp?SectionID=3&TopicID=138&DocumentID=4756

Banks, J. and Mc Coy, S (2011). A *Study on the Prevalence of Special Educational Needs.* Trim: National Council for Special Educational Needs.

Cameto, R. & Nagle, K. (2007). *Facts from NLTS2: Orientation and Mobility Skills of Secondary School Students with Visual Impairments.* Menlo Park, CA: SRI International. Retrieved from www.nlts2.org/fact_sheets/nlts2_fact_sheet_2007_11.pdf

Cameto, R., Knokey, A.-M. et al. (2009). *State Profiles on Alternate Assessments Based on Alternate Achievement Standards: A Report from the National Study on Alternate Assessments, NCSER 2009-3013*. National Center for Special Education Research. Retrieved from http://ies.ed.gov/ncser/.

Cole, C. (2006). *Closing the Achievement Gap Series: Part III What is the Impact of NCLB on the Inclusion of Students with Disabilities?* Vol. 4, No. 11, Center for Evaluation and Education Policy. Retrieved from http://www.ceep.indiana.edu/projects/PDF/PB_V4N11_Fall_2006_NCLB_dis.pdf

Cosgrove, J., McKeown, C., Travers, J., Lysaght, Z., Ní Bhroin, O., & Archer, P. (2014). Educational experiences and outcomes for children with special educational needs: A secondary analysis of data from the Growing Up in Ireland Study (NCSE Research Report No. 17). Trim: National Council for Special Education.

Darling-Hammond, L. (2007). Race, inequality and educational accountability: the irony of “No Child Left Behind”. *Race, Ethnicity and Education*, 10(3): 245−60.

Department of Education and Skills (DES) (2011). *Literacy and numeracy for learning and life, The National Strategy to Improve Literacy and Numeracy among Children and Young People 2011–2020*. Dublin: DES.

Department of Education and Skills (DES) (2014). *Revised arrangements for the submission of returns to the Department of Education and Skills by post-primary schools for 2014/15 and future years (Circular Letter 38/2014).* Accessed from. http://www.education.ie/en/Circulars-and-Forms/Active-Circulars/cl0038_2014.pdf

Department of Education and Skills (2015). *Revised arrangements for the submission of returns to the Department of Education and Skills by post-primary schools for 2014/15 and future years.* Accessed from. http://www.education.ie/en/Circulars-and-Forms/Active-Circulars/cl0038_2014.pdf

DfE (2011a). *Children with Special Educational Needs 2011: an analysis.* London: DfE.

Douglas, G., Travers, J., McLinden, M., Robertson, C., Smith, E., Macnab, N., Powers, S., Guldberg, K., McGough, A., O’Donnell, M. and Lacey, P. (2012). *Measuring Educational Engagement, Progress and Outcomes for Children with Special Educational Needs: A Review.* National Council for Special Education (NCSE), Trim, Ireland. http://ncse.ie/research-reports

Eivers, E., Close, S., Sheil, G., Millar, D., Clerkin, A., Gilleece, L. & Kiniry, J. (2010). The 2009 National Assessments of Mathematics and English Reading. Dublin: The Stationery Office.

Elliot, S. N., Kettler, R. J. and Roach, A. T. (2008). Alternate Assessments of Modified Achievement Standards: More Accessible and Less Difficult Tests to Advance Assessment Practices? *Journal of Disability Policy Studies*, 19: 140–152.

European Agency for Development in Special Needs Education (EADSNE) (2010). *Special Needs Education Country Data 2010*. Odense, Denmark: European Agency for Development in Special Needs Education.

Florian, L., Rouse, M., Black-Hawkins, K. & Jull, S. (2004). What can national data sets tell us about inclusion and pupil achievement? *British Journal of Special Education,* 31(3): 115–121.

Galton, M. and Macbeath, J. (2015). *Inclusion: Statements of intent – a report to the National Union of Teachers on the current state of special educational needs and disability provision.* National Union of Teachers. Retrieved from: https://www.teachers.org.uk/files/inclusion-report.pdf

Gorard, S. (2010). Serious doubts about school effectiveness. *British Educational Research Journal,* 36(5): 745–766.

Greenen, K. & Ysseldyke, J. (1997). Educational standards and students with disabilities. *The Educational Forum,* 61(3): 220–229.

Hopper, M. F. (2001). *The implications of accommodations in testing disabled students*. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED463627

Humphrey, N. & Squires, G. (2010). *Achievement for All Evaluation: Interim Report*. London: DfE.

Humphrey, N. and Squires, G. (2011). *Achievement for All: National Evaluation. Final Report*. London: DfE.

Isaacs, T. (2010). Educational assessment in England. *Assessment in Education: Principles, Policy and Practice,* 17(3): 315–334.

Kendall, S., Straw, S., Jones, M., Springate, I. & Grayson, H. (2008). *A Review of the Research Evidence (Narrowing the Gap in Outcomes for Vulnerable Groups).* Slough: NFER.

Kyriazopoulou, M. & Weber, H. (2009). *Development of a set of indicators for inclusive education in Europe.* Odense: European Agency for the Development of Special Needs Education.

Lazarus, S. S., Thurlow, M. L., Lail, K. E. & Christensen, L. (2009). A Longitudinal Analysis of State Accommodations Policies: Twelve Years of Change, 1993–2005. *Journal of Special Education,* 43: 67–80.

Lowrey, K. A., Drasgow, E. et al. (2007). Impact of Alternate Assessment on Curricula for Students with Severe Disabilities: Purpose Driven or Process Driven? *Assessment for Effective Intervention* 32(4): 244–253.

Maxwell L.A., & Shah N. (2011). Often Excluded, More Special-Needs Students Taking NAEP. *Education Week,* 31(12).

McLaughlin, M. J. & Thurlow, M. (2003). Educational Accountability and Students with Disabilities: Issues and Challenges. *Educational Policy*,17: 43

Morris, M., Rutt, S., Kendall, L. & Mehta, P. (2008). *Overview and Analysis of Available Datasets on Vulnerable Groups and the Five ECM Outcomes (Narrowing the Gap in Outcomes for Vulnerable Groups)*. Slough: NFER.

National Council for Curriculum and Assessment (NCCA) (2011). Towards a framework for Junior Cycle. Dublin: NCCA

National Council on Disability (2008), The No Child Left Behind Act and the Individuals with Disabilities Education Act: A Progress Report, National Council

on Disability, accessed from http://www.educationalpolicy.org/publications/pubpdf/NCD_2008.pdf

National Dissemination Centre for Children with Disabilities (2011). IDEA Part B indicators. Retrieved from http://nichcy.org/laws/idea#indicators

NCES (2011b), Digest of Educational Statistics 2010, Chapter 2: Elementary and Secondary Education, NCES. Retrieved from http://nces.ed.gov/programs/digest/d10/ch_2.asp

Ndaji, F. & Tymms, P. (2009). *The P scales: Assessing the Progress of Children with Special Educational Needs.* London: John Wiley.

Ndaji, F. & Tymms, P. (2010). The P Scales: How well are they working? *British Journal of Special Education*, 37(4): 198–208.

Newman, L., Wagner, M., Cameto, R. & Knokey, A.-M. (2009). *The Post-High School Outcomes of Youth With Disabilities up to 4 Years After High School: A Report From the National Longitudinal Transition Study-2 (NLTS2)*. NCSER 2009-3017. National Center for Special Education Research. 400 Maryland Avenue SW, Washington, DC 20202. Tel: 800-437-0833; Fax: 202-401-0689. Retrieved from http://ies.ed.gov/ncser/.

Norwich, B. (2008). *Dilemmas of Difference, Inclusion and Disability.* London and New York: Routledge.

Norwich, B. (2013). *Addressing tensions and dilemmas in inclusive education: living with uncertainty.* London and New York: Routledge.

Qualifications and Curriculum Authority (QCA) (2007). *Assessment for disabled students: an international comparison: QCA briefing document.* London: QCA.

Qualifications and Curriculum Authority (QCA) (2009a). *Using the P scales: assessing, moderating and reporting pupil attainment at levels P1 to P8.* London: QCA.

Qualifications and Curriculum Authority (QCA) (2009b). *The P scales: Level descriptors P1 to P8.* London: QCA.

Qualifications and Curriculum Development Agency (QCDA) (2011). *Using the P scales to assess pupils' progress.* Coventry: QCDA.

Sanford, C., Newman, L., Wagner, M., Cameto, R., Knokey, A.-M. & Shaver, D. (2011). *The Post-High School Outcomes of Young Adults With Disabilities up to 6 Years After High School. Key Findings From the National Longitudinal Transition Study-2 (NLTS2)* (NCSER 2011-3004). Menlo Park, CA: SRI International. Retrieved from www.nlts2.org/reports/

Shiel, G., Kavanagh, L. & Millar, D. (2014). The 2014 National Assessments of Reading and Mathematics. Volume 1 Performance Report. Dublin: Educational Research Centre.

Sheil, G., Kellagham, T. & Moran, G. (2010). *Standardised testing in lower secondary education.* Dublin, Ireland: National Council for Curriculum and Assessment.

Wagner, M., Newman, L., Cameto, R., Levine, P. & Marder, C. (2007). *Perceptions and Expectations of Youth With Disabilities. A Special Topic Report of Findings from the National Longitudinal Transition Study-2 (NLTS2).* Menlo Park, CA: SRI International. Retrieved from www.nlts2.org/reports/

Watkins, A. (Ed) (2007). *Assessment in Inclusive Settings: Key Issues for Policy and Practice.* Odense, Denmark: European Agency for Development in Special Needs Education.

Williams, J., Greene, S., Doyle, E., Harris, E., Layte, R., McCoy, S., McCrory, C. Murray, A., Nixon, E., O'Dowd, T., O'Moore, M., Quail, A., Smyth, E., Swords, L. & Thornton, M. (2009). *Growing Up in Ireland National Longitudinal Study of Children: The lives of 9-year-olds*. Dublin: Office of the Minister for Children and Youth Affairs.

Table 1. Summary of SEN terms for England, Ireland and the US, % of school age population identified as having SEN, and % of school age population in special school / segregated provision

Country (and approximate school age population)	SEN term used / key policy document	% school age population identified as SEN	% school age population in special school / segregated provision	Notes / Sources
England (8.0 million)	Special educational need (SEN); e.g. Special Educational Needs and Disability Act (2001)*	20.6% (1) 2.8% (2)	1.3% (3)	(1) Based on all children with SEN (with and without a statement of SEN). (2) Based on all children with a statement of SEN (more severe). (3) Based on placement of children with SEN (with and without statement). Most of the 1.3% attend special schools, but also pupil referral units. An approximate additional 0.2% attend schools with a resource base in a mainstream school. *Source: DfE (2011a).*
Ireland (0.6 million)	Special educational need (SEN); Education for Persons with Special Educational Needs (EPSEN) Act (2004)	5.2% (1)	0.8% (2) 0.4% (3)	(1) Based on children with formal diagnosis of SEN. Does not include primary school pupils with high incidence SEN with no formal diagnosis but who may receive support under the General Allocation Model (GAM). Recent prevalence data suggest that up to 25% of young people may have SEN as defined by the EPSEN Act (2004) − Banks and McCoy (2011). (2) Special schools. (3) Special classes in mainstream schools. *Source: EADSNE (2010).*
US (49.1 million)	Disability; Individuals with Disabilities Education Act (IDEA), various revisions	13.2% (1)	0.4% (2) 7.7% (3) 2.9% (4) 1.9% (5)	(1) Based on all students with disabilities (2008). (2) Separate schools for pupils with disabilities (public and private). (3) Regular school (less than 21% of time outside general class). (4) Regular school (between 21% and 60% of time outside general class). (5) Regular school (more than 60% of time outside general class). *Source: NCES (2011b)*

* Note: Since our review, the Children and Families Act (2014) introduced significant reforms and led to the introduction of a new Special Educational Needs and Disability (SEND) code of practice (2014).

Table 2. The education year group labels for the England, Ireland and the US.

Age	Ireland	England	US
4+	Junior Infant	Reception (Foundation)	Pre-school
5+	Senior Infant	Year 1 − Key Stage 1 (Primary)	Kindergarten (Primary / elementary school)
6+	First Class	Year 2	1st Grade
7+	Second Class	Year 3 − Key Stage 2	2nd Grade
8+	Third Class	Year 4	3rd Grade
9+	Fourth Class	Year 5	4th Grade
10+	Fifth Class	Year 6	5th Grade
11+	Sixth Class	Y7 − Key Stage 3 (Secondary)	6th Grade (Secondary: Junior high school)
12+	First Year (Secondary School, Junior Cycle)	Y8	7th Grade
13+	Second Year	Y9	8th Grade
14+	Third Year	Y10 − Key Stage 4	9th Grade (Secondary: High school)
15+	Transition Year	Y11	10th Grade
16+	Fifth Year (Senior Cycle)	Y12 − Key Stage 5 (non-compulsory)	11th Grade
17+	Sixth Year	Y13	12th Grade

Table 3. England: Summary of inclusive assessment approaches (assessment approaches, accommodation / alternative, relevance) in relation to national assessments.

Inclusive assessment approaches	Summary details and example sources
Assessment approaches	1. A key feature of the approach taken in England is linked to standards-based education – the specification of a national curriculum and the assessment of most children's progress on national tests and assessment tasks (Isaacs, 2010). 2. Most pupils are included in assessment tasks and tests at the end of each key stage (i.e. age seven, 11, 14, 16). Currently, at the end of Key Stage 2 (age ten to 11) most children take national tests (or standard attainment tests, SATs) in English, maths and science. Similarly most children take a range of accredited national tests (General Certificate of Secondary Education, GCSE) in a range of subjects at the age of 15 to 16. 3. All local authority schools, including special schools, must deliver the national curriculum which should be sufficiently flexible to accommodate different paces and styles of learning. 4. Performance tables, compiled and published since 1997, contain statistics on student test and examination results for each school (Isaacs, 2010). All subjects are assessed through teacher assessments with progress reported to parents every year. 5. Data on student characteristics (including details of SEN), school characteristics, attainment and attendance are collated within the National Pupil Database (NPD) (e.g. Florian et al, 2004; Gorard, 2010). 6. Analysis of this data allows disaggregation of pupils with SEN and there is a legal requirement that the government presents data on the progress and outcomes of students with SEN each year (e.g. DfE, 2011).
Accommodation / alternative	7. Accommodated versions of national tests are provided to increase inclusion in the assessment. This is true for assessments leading to student accreditation (e.g. GCSEs at the age of 16) and for non-accredited national assessments (SATs) (e.g. QCA, 2007). 8. As an alternative version of national assessments, a series of preparatory levels and associated framework (known as performance or P scales ranging from P1 to P8) monitors progress at the foundation level of the national curriculum for some pupils with learning difficulties (e.g. QCA, 2009a; 2009b; Ndaji and Tymms, 2009; 2010). Foundation level is below Level 1 of the

	curriculum (expected to be achieved between age five and six).
Relevance	9. No national assessments or system-based progress and outcomes data is gathered beyond the national curriculum for children with SEN (i.e. in relation to an additional or wider curriculum). Schools would be expected to carry out intervention and assessment of progress in relation to a student's Individual Education Plan, but this is not collated. 10. The NPD provides a flexible mechanism for combining a range of datasets gathered from different sources (e.g. surveys and evaluations), and these could include measurement of progress and outcomes of children with SEN in relation to wider curriculum areas. 11. For example, the evaluation of Achievement for All (AfA) school improvement framework in which data on 7,770 pupils development in social and emotional aspects of learning (e.g. behaviours, friendships, participation in extra-curricular activities) was combined with data from the NPD (Humphrey and Squire, 2010; 2011).

Table 4. Ireland: Summary of inclusive assessment approaches (assessment approaches, accommodation / alternative, relevance) in relation to national assessments.

Inclusive assessment approaches	Summary details and example sources
Assessment approaches	1. In contrast to England and the US, Ireland has traditionally not adopted a standards-based education approach, in which children's progress is measured through national tests (e.g. Sheil et al, 2010). 2. Nevertheless, the National Literacy and Numeracy Strategy (DES, 2011) includes proposals for more assessment at various ages. There are associated requirements for schools to report pupil results to parents and to report aggregated results to the Department of Education and Skills. The first implementation (May/June 2012) of the standardised tests did not include all children with SEN. 3. National award-bearing assessments are the Junior Certificate (usually at approximately 15 years) and Leaving Certificate (usually at approximately 17 to 18 years). Data in relation to SEN is not collected for these assessments (other than numbers and type of examination accommodation requests) so no analysis is possible (Douglas et al, 2012). 4. Ireland carries out periodic sample-based national assessments (e.g. in 2009 in mathematics and English reading in primary second and sixth classes, see Eivers et al, 2009). The sample for each age group was about 4,000 pupils, although this excluded some students with SEN (e.g. exemption figure is 1% for 2014, see Shiel et al, 2014). Pupil SEN is not recorded as part of the assessment. 5. Ireland does not have a national pupil database, although significant developments are ongoing in relation to the Primary (POD) and Post-primary online database (PPOD), e.g. DES (2014). The most recent version of the database contains fields in relation to SEN.
Accommodation / alternative	6. National Literacy and Numeracy Strategy draws upon standardised tests and at this stage do not offer accommodated or alternative versions. (DES, 2011) 7. National award-bearing assessments in the junior (age 14) and senior cycles (age 17) include options for accommodations. The junior cycle assessment has a limited range of levels of assessment criteria, but there are plans to increase this range to include assessment of learning at lower levels (NCCA, 2011). 8. The sample-based national assessments do not include accommodations or alternatives.

Relevance	9. No national assessments or system-based progress and outcomes data is gathered in relation to an additional or wider curriculum for children with SEN. Schools would be expected to carry out intervention and assessment of progress in relation to a student's Individual Education Plan, but this is not collated. 10. Another study in Ireland which has the potential to gather outcomes and progress data in relation to pupils with SEN is the use of data from 'Growing Up in Ireland' (Williams *et al*. 2009), e.g. Cosgrove et al (2014). Given Ireland does not currently have an operational national pupil database, these analyses are not able to take advantage of combining datasets.

Table 5. US: Summary of inclusive assessment approaches (assessment approaches, accommodation / alternative, relevance) in relation to national assessments.

Inclusive assessment approaches	Summary details and example sources
Assessment approaches	1. A key feature of the US approach is the link to standards-based education, where schools and school districts are held accountable for progress towards state-defined learning standards which are a key focus of the No Child Left Behind Act (NCLB). NCLB requires that states assess performance annually in grades 3 to 8 (8-13 years of age) in reading/language/arts, mathematics and science with additional tests for grades 10 to 12 (15-17 years of age). 2. The Individuals with Disabilities Education Act (IDEA) and NCLB require standards-based accountability monitoring for all students with SEN. States must assess student progress against these standards (which includes high school graduation and employment outcomes) and these data are published. 3. Analysis of these data allows disaggregation of pupils with SEN, although capacity for disaggregation by disability/SEN sub-groups varies from state to state (Altman *et al.* 2010). 4. The National Assessment of Educational Progress (NAEP) is a sample-based assessment which takes place annually for grades 4, 8 and 12 in relation to a range of curriculum areas. Students with disabilities are included in NAEP, although they can be excluded if teachers feel the assessment is inappropriate. Targets have been set to ensure 85% of the sample with disabilities are included (Maxwell and Shah, 2011).
Accommodation / alternative	5. All students must be included in state assessments, and therefore strict requirements exist for states to provide accommodated and alternative versions of assessments to include students with disabilities. Lazarus et al, (2009) provides an analysis of accommodation procedures across the US. 6. As an alternative version of state assessments, states are required to provide 'alternate assessments' for students working at achievement standards at a basic level. These alternate achievement standards are an expectation of performance that differs in complexity from grade-level achievement standards. Cameto et al, (2009) provides an analysis of approaches developed in different states (the analysis was required by IDEA). 7. NAEP sample-based assessment includes accommodations but not alternatives.

Relevance	8. Monitoring and assessing progress and participation is required in relation to the 20 IDEA Part-B indicators. This includes curriculum areas defined by NCLB, but also wider indicators: e.g. post school outcomes including employment; high school graduation and drop-out rates (National Dissemination Centre for Children with Disabilities, 2011). 9. On an individual level, pupils identified as having a disability have their development monitored and assessed through an individual education plan (IEP). By law, schools must include a statement of the child's present level of performance, annual goals and short-term objectives plus details of all special support to be provided (McLaughlin & Thurlow, 2003). 10. Another significant source of data is the National Longitudinal Transition Study-2 (NLTS2) which has a nationally representative sample of around 11,000 students receiving special education services. The survey gathers data on educational progress and outcomes in a range of areas, including attainment as well as broader outcome areas (e.g. Wagner et al, 2007; Sanford et al, 2011). 11. Additionally, the scale of the NLTS2 means that disability-specific data and outcomes are gathered (e.g. Cameto and Nagel, 2007).

Figure 1: Inclusive Assessment Framework