

RISING TO THE CHALLENGE



***A contribution to the development of a holistic value-added
Assessment and Evaluation Framework
for Northern Ireland***

Contents

• Acknowledgement

This paper draws extensively on research-informed perspectives within

- the OECD NI report (Dec 2013)
- the OECD '*Synergies for Better Learning: An International Perspective on Evaluation and Assessment*' Report (March 2013), and
- '*Data-driven Improvement and Accountability*' (Hargreaves A & Braun H Oct 2013).

1: Introduction

- 1.1 Context
- 1.2 Aim
- 1.3 Objectives

2: Agreeing Principles

- 2.1: Shared Agreement
- 2.2 Fundamental underpinning principles

3: Addressing Challenges

- 3.1 The use of teacher assessment
- 3.2 Levels of Progression
- 3.3 Moderation
- 3.4 Contextual value-added
- 3.5 Pupil value-added
- 3.6 Revision of government targets

4: Proposed Assessment Tools and Processes

- 4.1 Teacher 'Assessment for Learning'
- 4.2 Assessment tools
- 4.3 Base-line assessment of productive language on entry to school:
- 4.4 Cognitive abilities analysis
- 4.5 Attitudes to learning analysis
- 4.6 Occasional standardised testing
- 4.7 Using outcomes to calculate value-added
- 4.8 Transfer of data

5: Moving forward

- 6.1 Testing for fitness for purpose, utility and manageability
- 6.2 Full economic appraisal
- 6.3 Interim arrangements

6: Capacity Building

- 6.1 Teacher skills in the use of formative assessment for learning
- 6.2 Senior management skills in managing data and pedagogical leadership

7: Summary of Proposals

Rising to the Challenge

A contribution to the development of a holistic value-added Assessment and Evaluation Framework for Northern Ireland

1 Introduction

1.1 Context

This paper has been developed in response to the OECD's recent report on assessment and evaluation processes in Northern Ireland which highlights that consensus is a pre-requisite for the successful implementation of policy reform and that building consensus is an iterative process of proposals and feedback to build ownership, trust, respect and transparency.

1.2 Aim

The paper focuses on the pupil assessment component of the framework and primarily on the Key Stage 1, 2 and 3 elements which have proved contentious to date. While acknowledging that CCEA is the lead agency in this area, this paper responds to the OECD's challenge that *'there is much to be gained from cross-fertilisation of distinct perspectives into compromises than from antagonism and the imposition of particular views over other stakeholder groups'* (OECD Dec 2013: 44).

The aim of the paper is to 'rise to the challenge' to contribute proactively to the iterative process of developing a more holistic, value-added assessment (and evaluation) framework for Northern Ireland, with clear synergies between the key components of pupil assessment; teacher and leadership appraisal and school and system evaluation. While each of these components is already a facet of official policy, the OECD has highlighted the need for developments and refinements within each to avoid *'duplication of procedures and prevent inconsistencies of objectives.'* (OECD 2013) and establishing clear *'synergies'* between the components *'for better learning'* (OECD March 2013).

1.3 Objectives

The objectives of the paper are to stimulate discussion and seek consensus on:

- The fundamental principles that should underpin Northern Ireland's approach to assessment (and evaluation) to ensure consistent objectives;
- manageable value-added solutions to addressing remaining concerns; and
- the key components and principles that should inform a holistic and integrated assessment framework (to be aligned with a refined evaluation framework¹).

The aspiration is to 'future proof' Northern Ireland's assessment and evaluation processes to make us a world leader in value-added assessment and accountability.

¹ *Refinements of the evaluation framework should build on OECD recommendations and the outcomes of the NI Assembly Inquiry.*

2 Agreeing Principles

2.1: Shared Agreement

There is shared agreement about the key components of official assessment policy (OECD P 57-63) including the focus on:

- formative assessment and teachers' professional judgement;
- strengthening assessment literacy among teachers and promoting student engagement in self- and peer-assessment;
- moderation to build teacher assessment literacy/capacity and to increase trust in teacher professional judgements;
- strong communication with parents and reporting on student progress;
- providing central diagnostic tools; and
- effective use of data and information systems to track progress in student learning.

The challenge is to ensure that each of these components (assessment for learning, moderation, effective use of data and reporting on progress) is implemented in manageable ways to support the primary purpose of improving teaching and learning.

2:2 Fundamental underpinning principles

The impact of data-driven accountability is profoundly influenced by the breadth and depth of data used; the care with which it is analysed; the use to which the analysis is put; the consequences that flow from it; and how these consequences affect different groups of teachers, students and schools. The OECD team highlights that:

- Effective use of data can help teachers and schools to evaluate pupil progress and their own teaching and to make appropriate data and research-informed classroom, school and system interventions in pursuit of continuous improvement and to inform accountability.
- Conversely, inappropriate and narrow use of data can lead to the distortion of teaching and learning and distraction from the broader purposes of schooling, with the danger of a deterioration of services, morale and commitment. For example an over-emphasis on narrow measures may well achieve improvement in numerical data in priority policy areas but is no guarantee of real improvement in overall education standards, due to an artificial emphasis on meeting designated targets while other equally important areas are neglected.

To ensure that a focus on data-driven accountability generates more positive and fewer negative outcomes.....:

It is proposed that a clear consensus and commitment is established around the following 5 fundamental underpinning principles: viz. broader purposes; broader data; sensitive analysis of data; value-added; and supportive accountability

- 1 Broader purposes:** as opposed to excessive concentration on meeting targets on just one or two policy areas such as 'Count Read Succeed';
- 2 Broader data:** that serves all system level goals so that the focus of teaching and learning is not distorted and no one indicator carries disproportionate weight. The OECD team's advice is that such data should take account of measures such as the development of critical thinking and personal capabilities, dispositions to learn and overall well-being.
- 3 Sensitive analysis of data:** taking account of *contextual factors* to enable comparisons that are fair to schools and pupils. To ensure sophisticated analysis of genuine trends, as opposed to 'bull-whip' responses to what may be short-term and unrepresentative blips, data analysis must take account of :
 - *the fallibility of data* drawn from different forms of evidence involving non-scientific scoring systems and human judgements with significant margins for error.
 - *The volatility of data* such as that derived from small class, key stage or school sample sizes which can make schools' value-added estimates vary inexplicable from year to year if the fluctuations are erroneously interpreted.
 - *Timescales factors* such as over- dependence on most recent evaluations which can contain potentially volatile swings in results from one year to the next, that are not representative of broader trends.
 - *Changes in 'standards'* such as when curriculum and/or assessment instruments are themselves changed making it impossible to draw conclusions about whether standards have improved.
- 4 Value-added:** taking account of *school and individual pupil contextual factors* to enable comparisons that are fair to schools and pupils (rather than reliance on "raw" results which may more accurately measure the school's intake, rather than the value it has added to student outcomes).
- 5 Supportive accountability:** not attaching external rewards or punitive consequences to the extremes of performance but rather operating on the assumption that poor performance is largely due to insufficient capacity and/or resources rather than to lack of effort or deliberate intransigence.

3. Addressing Challenges

The OECD team has helped to clarify the key challenges that need to be addressed in order to achieve consensus. The implications of each of the research-informed position is summarised below and workable proposals offered.

3.1 The use of teacher assessment

It has been clearly communicated in the most recent proposals relating to statutory assessment that teacher assessment against Levels of Progression is primarily designed for diagnostic and formative purposes. The OECD team cites three research-informed views in relation to the use of this data:

- that assessment designed for diagnostic and formative purposes should not be used for summative or accountability purposes as this would likely compromise its primary purpose (Linn, 2000);
- that the more purposes an assessment is intended to serve, the more each purpose will be undermined by compromises made during the design process (Pellegrino et al,2001); and
- that it is possible for an assessment to have multiple purposes as long as they are not logically incompatible (Newton 2007).

There is overwhelming evidence since teacher assessment against levels was first introduced in Northern Ireland two decades ago that the primary diagnostic and formative purpose of teacher assessment has been severely compromised by its dual use for accountability purposes. The clear and consistent message is that:

- the qualitative professional purpose of teacher assessment against progression criteria (for the improvement of teaching and learning and for informed feedback to pupils and parents) is considered by teachers to be extremely valuable; but
- the current quantitative measures as framed (and their use for accountability purposes) are considered of little or no utility to pupils, schools, parents, policy makers or politicians (*GTCNI survey findings 2013*).
- the evidence of distortion associated with their secondary use for accountability purposes is logical incompatible with their primary purpose and therefore educationally unacceptable.

It is proposed that teacher assessment should be used for diagnostic and formative purposes only to inform summative reporting to pupils and parents.

3.2 Levels / Indicators of Progression

It is acknowledged that it is not an easy task to develop criteria that are clear, [sufficiently detailed and fine grained] and widely agreed upon (Looney, 2011b; Nusche et al.,2011) but that this is crucial to inform subsequent teaching and learning and to develop a shared understanding of what may constitutes a specific performance at the different stages of learning progression.

Progression in learning is subtle and complex. There is no single linear developmental pathway which is neatly age-related with an expected end-point. Nor is the demands of 'a level' equivalent between key stages due to variation in context.

It is proposed that, when revising the Levels/ Indicators of Progression:

- **appropriate account is taken of progression in conceptual knowledge and understanding and associated thinking skills form an explicit and integral part of refined criteria; and**
- **numeric levels are replaced by progress indicators** *for example pre foundation, foundation, emergent, developing, competent, consolidated, advanced*

Framing progression indicators in this way will facilitate the effective use of the criteria by all teachers in all subject teachers and phases to promote a common understanding of standards within and across the curriculum. *A generic model and an exemplification of how the model can be used at classroom level in all key stages (including key stage 4 and 5) can be offered as a basis for discussion and development.* It may still be possible to translate this data into quantitative equivalence for aspects of system level performance analysis.

3.3 Moderation

We concur with the view that the involvement of teachers in moderation should develop their assessment capacity and improve the reliability of teacher assessed summative outcomes. In line with the proposal that teacher assessment should be used for diagnostic and formative purposes and to inform summative judgements, but not for accountability purposes, the continued emphasis within the latest moderation proposals towards verifying teacher and school numerical level judgements for accountability purposes is considered inappropriate.

It is proposed that the purposes of moderation support is to quality assure school's internal assessment processes and to enhance teacher capacity:

- **to use 'assessment for learning' pedagogy**
- **to devise appropriately challenging assessments**
- **to make valid assessments against knowledge and skills-based criteria across the Northern Ireland Curriculum.**

3.4 Contextual value-added

One of the strongest predictors of academic achievement is the socio-economic background of pupils and parental education. Statistical models can be used to incorporate a range of factors relating to contextual background. For example, *'In Sweden a model is used to assess a school's expected performance by adjusting its actual results with regard to student characteristics including parental education. A comparison is made between the school's expected and actual results to provide a measure of value added.* (Perry C. NIAR Oct 2013)

It is proposed that, in addition to the Free School Meals (FSM) Index, other mechanisms are explored to inform the development of a statistical model to enable the stratification of schools by intake (for example the use of such as Super Output areas potentially refined by using Geographical Information System (GIS) analysis of individual pupils post-codes (as in New Zealand) or parental education (as in Sweden)

3.5 Pupil value-added

'Individual value added' aims to measure the progress made by a pupil between different stages of education. The advantage of value-added assessment measurement over criterion or norm-referenced assessment is that it focuses on how far a pupil has progressed at the end of a specific period (for example, at the end of the school year or key stage, compared to the start). This requires a pre-test (or the use of relevant data passed on by the from the previous teacher) and a post-test (end of year/cycle) to determine what pupils have learned during a particular course of study. This data provides results that can be compared across classrooms and years. The more information teachers can gain about a pupil's potential, learning dispositions and progress, the better able they will be to tailor the learning environment and ways of teaching and learning to enable pupils to maximise their potential.

It is proposed that: a range of research-informed assessment tools and approaches should used to identify individual strengths and areas for development and to predict outcomes which can be used to evaluate value-added (See section 4 for detail)

3.6 Revision of government targets

The proposal to use value-added as opposed to raw outcomes has major implications for the nature of government educational targets and the way in which these are monitored and reported on by the NI Audit Office.

It is proposed that government educational targets are based on research-informed analysis of performance against a broad range of measures that align with system goals and are monitored in a way which avoids distortion of those goals

4: Proposed Assessment Tools and Processes

4.1 Assessment tools

The OECD team reported that both primary and post-primary schools have identified the need for diagnostic measures to monitor pupil and cohort progression against individual base-line starting points to enable comparisons that are fair to students and to schools and to facilitate the exchange of pupil information from primary to post-primary schools. Key considerations are:

- the diagnostic qualities and ease of use of the proposed mechanisms;
- who can access the data in what form for what purpose; and, crucially
- how it is used subsequently and reported for system accountability to ensure that it enhances and does not distort teaching and learning or overburden teachers.

Before describing the various tools, least what is suggested appears overly complex, it should be noted that all of the proposed tools:

- already exist in paper and digital format;
- are research informed and have been validated fit for purpose;
- can be customised to the specific context of NI;
- generate sophisticated statistical and narrative reports for teachers and senior management which provide valuable educational insights in relation to individual pupils, groups and cohorts;
- collectively address all assessment and value-added purposes; and
- should be affordable if procured as an integrated package at system level.

Many of these tool are already in use (independently paid for) by schools in Northern Ireland.

It is proposed that an existing range of diagnostic, predictive, performance monitoring and reporting tools is used as part of a holistic assessment framework to provide valuable data to enhance teaching and learning and enable value added reporting (as described below and illustrated in Figure 1 over)

4.2 Base-line assessment

Productive language on entry to school is a key indicator and determinant of ability to learn. A range of baseline tools exists to assess spoken language on entry to school, for example, *The Renfrew Bus Story (RBS)*, which is enjoyable for children, is a quick to administer short screening assessment which used 'narrative re-tell' or storytelling to assess receptive and expressive oral language for young children age 3 years to 6 years 11 months. The outcomes provide a quantitative and qualitative assessment of each child's oral language skills based on rich language data to identify children with language impairments, as well as to predict of later language and academic skill (Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998).

4.3 Cognitive abilities analysis

From the age of 7 it is possible to generate a comprehensive profile of individual pupil's dispositions to learn; and abilities to reason with, and manipulate, different types of material through a series of Verbal, Non-Verbal, Quantitative and Spatial Ability tasks. (Recent research has confirmed the importance of assessing pupils' spatial ability in order to develop and support skills that are important across the curriculum and particularly important for success in STEM subjects and careers). The analysis of outcomes provides teachers with a comprehensive profile of individual pupil's reasoning abilities, to identify strengths, weaknesses and learning preferences and to generate indicators of future attainment (for example at KS2, KS3, GCSE, AS/A Level). The data can be used, alongside attainment data (and other factors known to impact on learning, such as attendance and attitude), to set individual pupil targets; to plan focused teaching and learning (with interventions for different individuals or groups of pupils) and to monitor progress and track progress.

4.4 Learning dispositions analysis

Analysis of pupils' attitudes towards themselves as learners and their attitudes towards school on an individual basis can provide insights into motivation, and well-being to enable early identification and early intervention strategies to be provided for those at risk. Insights can help teachers and schools to set smarter monitoring and tracking targets to improve student well-being, behaviour and attendance and to reduce disaffection. The improvement of dispositions to learn and attitudinal measures such as improvement in liking, for example reading, can inform targets and interventions.

4.5 Occasional standardised assessment

While teachers' professional judgments are based on on-going day to day assessment, the use of occasional standardised tests can give teachers an informed snap-shot of how individuals and pupil cohorts compare against UK/NI standards on traditional literacies (i.e. communication and using maths). The occasional use of standardised tests (a few weeks after the beginning of the year and/or at the end of a year) can provide helpful in-depth information to establish a baseline and possible gaps in learning in order to plan and adapt teaching and against which to track monitor and report progress. Outcome scores can provide insights into bands of performance across a cohort and short-comings in progress in skill areas to inform the focus of future teaching and learning

4.6 On-going Teacher Assessment

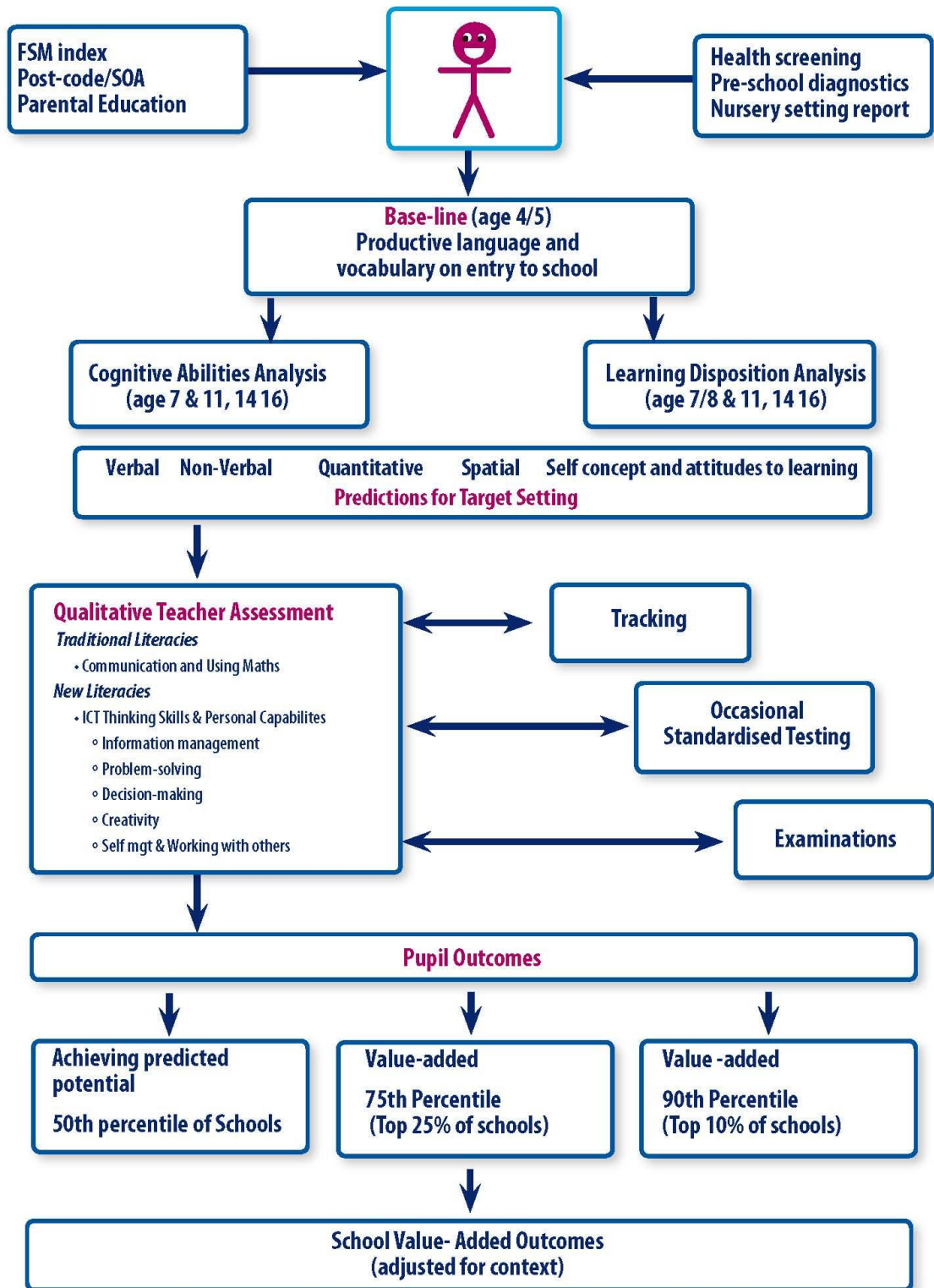
The information provided by these diagnostic, predictive and monitoring tools aim to inform on-going teacher assessment for learning, which is at the heart of raising standards, involving the development of:

- stimulating curriculum planning
- appropriately challenging assessment tasks
- shared learning intentions
- agreed success criteria,
- effective questioning,
- peer/self assessment and
- targeted feedback on next steps in learning.

(An approach to devising and assessing tasks using revised progression indicators involving both traditional and 'new' literacies can be provided for discussion)

Tools to enhance Teaching Learning and Assessment

(including informed target-setting and the assessment of value-added)



4.8 Reporting outcomes and calculating value-added

The combination of outcomes from cognitive abilities analysis, dispositional analysis, base-line and occasional standardised progress data will help teachers and schools to set informed aspirational (but achievable) targets for each individual pupil. The assessment of pupil achievement and value added should draw on the extent to:

- Evidence of meeting or surpassing predicted targets
- Improvements in dispositions to learn and sense of well being
- Internally moderated assessments of
 - ICT and
 - the development of thinking skills and personal capabilities

The outcomes can be adjusted using appropriate statistical models to give a measure of contextualised school value-added.

4.9 Transfer of data

The transfer of detailed pupil data from teacher to teacher and school to school is crucial to:

- to avoid gaps in information
- to assist future planning; and
- to prevent unnecessary and costly duplication of processes

<p>It is proposed that rich pupil data is transferred each year in an agreed format to assist future planning, teaching, learning and assessment</p>

5 Moving Forward

5.1 Pilot testing and phased implementation

Best practice recommends that any proposed model be thoroughly piloted and that feedback from the pilot be used to assess and amend the model as necessary before procurement and planned, phased roll-out on a systemic scale.

The advantages of the model proposed is that some of the core components are already in use (and paid for independently) by a large number of schools in Northern Ireland. However, it is not known how many schools make use of the full suite of components and their analytical reporting and value-added potential

It is proposed that a number of case-study schools be identified at both primary and post-primary level to explore the quality, educational utility and manageability of the proposed model and associated diagnostic and monitoring tools before considering customised procurement.

5.2 Full economic appraisal

The majority of schools in Northern Ireland already expend significant funds on standardised testing but many (possibly most) do not make use of the analytical predictive and advisory components that are available alongside these tools to enhance their educational utility. The proposed has the potential to achieve stakeholder buy-in because of: familiarity with, and trust in, the diagnostics already offered by elements of the model; the synergies between the various components; its ease of use and manageability and its potential to address all quality assessment and data analysis needs for the foreseeable future, freeing up teachers' time to focus on the core professional task of quality teaching and learning to meet pupils' needs and to improve their outcome.

It is proposed that a full economic appraisal is undertaken of the merits of procuring a completely integrated suite of tools that has the support of all stakeholders in terms of:

- **The use of quality information for educational and accountability purposes,**
- **manageability and teacher time saved for core professional purposes; and**
- **freeing up other agencies from a focus on accountability to a focus on providing much-needed capacity building support.**

5.3 Interim arrangements

While the merits, cost and potential procurement of the model are explored:

It is proposed that:

- **Teachers continue to assess and report to parents in qualitative terms as working at, above or below expected standards;**
- **Schools are invited to register to have their internal assessment processes quality assured by CCEA; and**
- **Best practice schools are enabled to act as centres of good practice for other schools in their catchment /area learning community.**

6. Capacity Building

The OECD Team recommends that to achieve effective implementation will require capacity building at all levels of the education system.

6.1 Teacher skills in the use of formative assessment for learning

Agreement on the fundamental principles that the sole purpose of teacher assessment and moderation is for the improvement of teaching and learning and the quality assurance of moderation purposes will free up CCEA capacity to focus on supporting assessment for learning as opposed to accountability. Scotland has supported a major 'assessment as, for and of learning' initiative. The NCCA in the Republic of Ireland is currently providing £100K of bursaries for practicing teachers to undertake PhD study in assessment for learning and ICT to develop deep capacity within the system.

It is proposed that over the next few years that assessment support resources should focus on developing teacher assessment for learning capacity; and

6.2 Senior management skills in managing data and pedagogical leadership

Agreement on the use of broader data for informed target-setting and the assessment of value-added will require Principals and senior management in schools to be able to understand, interpret and use data in increasingly sophisticated ways in pursuit of improved teaching and learning for improved outcomes. The latest digital developments in pupil assessment facilitated the immediate feedback to senior managers and teachers (as well as pupils and parents) of outcomes in the form of analytical and advisory graphical and narrative reports to support more focused teaching and focused interventions for individuals and groups.

It is proposed that there should be a major emphasis over the next few years on: the development of

- **Principal's pedagogical leadership skills and**
- **Senior management skills in managing and interpreting data.**

6.3 System capacity for value-added assessment

In order to ensure the transfer of skills in the development of research-informed assessment tools opportunity should be taken as part of any procurement exercise to ensure the transfer of knowledge and skills.

- **In time consideration might be given to establishing a centre of excellence in diagnostic, predictive and standardised assessment and analysis in Northern Ireland.**

7: Summary of Proposals

Fundamental principles	It is proposed that.....
	A clear consensus and commitment is established in relation the following 5 fundamental underpinning principles: viz. broader purposes; broader data; sensitive analysis of data; value-added; and supportive accountability
The use of teacher assessment	Teacher assessment should be used for diagnostic and formative purposes only to inform summative reporting to pupils and parents.
Levels of Progression	It is proposed that, when revising the Levels/ Indicators of Progression: <ul style="list-style-type: none"> • progression in conceptual knowledge and understanding and associated thinking skills form an explicit and integral part of refined criteria; • numeric levels are replaced by progress indicators
Moderation	The focus of moderation is to quality assure school's internal assessment processes and to enhance teacher capacity <ul style="list-style-type: none"> • to use 'assessment for learning' pedagogy; • to devise appropriately challenging assessments; • to make valid assessments judgements.
Contextual value-added	In addition to FSM other mechanisms are explored to inform the development of a statistical model to enable the stratification of schools by intake (<i>for example the use of such as Super Output areas potentially refined by using Geographical Information System (GIS) analysis of individual pupils post-codes (as in New Zealand) or parental education (as in Sweden)</i>)
Pupil value-added	A range of research-informed assessment tools and approaches should used to identify individual strengths and areas for development and to predict outcomes which can be used to evaluate value-added
Revision of government targets	Government educational targets are based on research-informed analysis of performance against a broad range of measures that align with system goals and are monitored in a way which avoids distortion of those goals

Assessment tools and processes	It is proposed that :
	<p>A suite of diagnostic, predictive, performance monitoring and reporting tools (<i>illustrated in Figure 1 over</i>) is used as part of a holistic assessment framework to provide valuable data to enhance teaching and learning and enable value added reporting to include:</p> <ul style="list-style-type: none"> • Base-line assessment of oracy on entry to school • Cognitive abilities analysis • Learning dispositions analysis • Occasional standardised assessment of traditional literacies • Teacher assessment of 'Traditional' and 'New literacies'
Transfer of data	<p>Rich data is transferred from teacher to teacher and school to school</p> <ul style="list-style-type: none"> • to avoid gaps in information • to prevent unnecessary duplication of assessment and • to assist future planning.
Piloting the model	<p>A number of case-study schools be identified at both primary and post-primary level to explore the quality, educational utility and manageability of the proposed model and associated diagnostic and monitoring tools before considering customised procurement.</p>
Full economic appraisal	<p>A full economic appraisal is undertaken of the merits of procuring a completely integrated suite of tools that has the support of all stakeholders in terms of:</p> <ul style="list-style-type: none"> • The use of quality information for educational and accountability purposes, • manageability and teacher time saved for core professional purposes; and • freeing up other agencies from a focus on accountability to a focus on providing much-needed capacity building support.
Interim arrangements	<ul style="list-style-type: none"> • Teachers continue to assess and report to parents in qualitative terms as working at, above or below expected standards; • Schools are invited to register to have their internal assessment processes quality assured by CCEA; and • Best practice schools are enabled to act as centres of good practice for other schools in their catchment /area learning community.

Capacity Building	It is proposed that:
Teacher skills in the use of formative assessment for learning	Assessment support resources should focus on developing teacher assessment for learning capacity
Senior management skills in managing data and pedagogical leadership	<p>There should be a major emphasis over the next few years on the development of:</p> <ul style="list-style-type: none"> • Principal's pedagogical leadership skills and • Senior management skills in managing and interpreting data
System capacity for value-added assessment	In time consideration might be given to establishing a centre of excellence in diagnostic, predictive and standardised assessment and analysis in Northern Ireland.