PEARSON

Integrated Assessment Systems for Improved Learning

Jon S. Twing, Ph.D.
Executive Vice President
Pearson Assessment & Information, Iowa City

Professor Bill Boyle
Director, Centre for Formative Assessment Studies (CFAS)
University of Manchester

Marie Charles
Assessment Researcher
Pearson Learning Solutions, London.

Abstract

The use of assessments globally are changing from current accountability or outcomes models like those used by OECD in PISA, the Emirates in their "external outcome measures" EMSA programme and in the United States under NCLB to more formative, interim and mixed mode assessments or systems that support learning. This move is presumably due to the need to enhance learning and instruction in a more systematic manner, focusing on aspects of teacher improvement and support, continuous progress monitoring or individual student growth modeling in order to prepare students for success in university or prepare them for careers internationally.

This paper focuses on the policy and practice implications in designing a learning system that incorporates the best practices of instruction, assessment and professional development while fulfilling the needs of so many different and diverse stakeholders. Specifically the paper will focus on the use of technology in providing real time "authentic" and highlights how to fulfill these goals quickly and economically.

Paper presented at the 36th Annual Conference of the International Association of Educational Assessment (IAEA), Bangkok, Thailand, 22-27 August, 2010.

Introduction

This paper has been organized into three parts. The first part sets the stage for what is described and defined as an "integrated assessment system". This system, if effective, will marry classroom instruction (and hopefully improved learning) by exploiting data collection, analysis and management from a variety of measures and fully integrated into the instructional stream. This system integrates assessment for learning (A4L/AfL) with instruction and summary programme evaluation or outcome measures (typically used for accountability purposes). It is the belief of the authors that only by explicitly linking measures with instruction and thereby empowering teachers to improve instructional delivery, can real education reform and real learning improvement take place. The second section is a paper authored by Professor Boyle and educational researcher Marie Charles. In this section, the authors review survey statistics showing the value and impact of integrated A4L/AfL and outline how current instructional practices, despite the perceived value of integrated assessments for learning, need radical changes to become fully integrated and fully formative. Finally, this paper concludes with a review and outline how aspects of evolving assessment design may indeed drive instructional practices toward integrated learning systems. These changes would exploit technology, be implemented within a learning system that encompasses both formative and summative assessment and would drive the measurement of important aspects of 21st Century skills like problem solving and collaboration.

The Organisation for Economic Co-Operation and Development (OECD) provides evidence that education is key to economic growth and to people's ability to earn a living and that education is important for societies as they respond to increasing cultural and ethnic diversity, equality, and the needs of disadvantaged people. The reality is that education is not just a local or regional concern, or even just a national concern. In an increasingly interdependent global economy, it makes sense to think of education as an interdependent concern as well. Arguably then, a significant investment in education should yield sizeable returns. It is odd, therefore, that most education reform efforts recently have focused almost solely on assessment with most fiscal investment being applied to assessments. In the United States for example, the No Child Left Behind act was criticized for over emphasis and overuse of multiple-choice achievement testing. Similarly, in the United Kingdom there is seldom an assessment season where the rigor of the current marking schemes are not called into question and accusations of slipping standards abound. In 1989, the focus in Australia was on reaching common learning goals (literacy, numeracy, problem solving to name a few) which were to be reported annually via the Annual National Report. National reporting scales were established such that each state could compare their progress on meeting the learning goals across the nation. This process was quite successful for a certain period. However, the various states and territories started to compare their performance on the basis of the results (even making comparisons to the second decimal point)! This then led to the Federal Government becoming more involved (through significant funding initiatives) in pushing a national assessment agenda and to establishing a National Assessment Program Literacy and Numeracy (NAPLAN) which is now a common assessment in literacy and numeracy. It seems that even "down under" the desire to improve learning defaulted to increased assessment. Finally, the Abu Dhabi Education Council recently implemented an external assessment to evaluate school service providers in the Emirates. The External

Measure of Student Achievement is essentially a high stakes outcome measure and represents the cornerstone of education reform in the region.

While the need and use of assessment to make educational decisions is not the cause of the problem, it seems odd that when the goal is improved learning—that teachers, students and instruction are only tangentially linked to the goal. Hence, it is likely these educational reform efforts will not help improve learning until they can become integrated with instruction and learning. A system that integrates assessment with instruction for the purpose of learning is more likely to succeed in improving student performance than either assessments of learning which are independent of key programme and policy decisions or summary high-stakes outcome measures used in isolation.

The next section of this paper reviews one investigation into assessments for learning, teacher perceptions of such assessments and the need to use such formative interactions between teachers and students at the forefront of the teach and learn cycle.

Leading Learning through Assessment for Learning?[1]

Although the context of this paper is the education system in England, the paper highlights issues of accountability, measurement of standards and forms of assessment which have major relevance to international teaching and learning. Since a National Curriculum was first introduced into England in 1988 the teaching of that curriculum has been assessed by a series of external (i.e. government produced) statutory summative tests at the end of key stages of a child's schooling. Because of the English government's insistence on measuring school performance through test outcomes, assessment has become equated with 'testing' and is seen by teachers as a means of producing 'scores' to measure standards rather than as a means to supply continuous information for teaching and to support learning (Hall et al, 2004). The paper begins by looking at the origins of assessment for learning (AfL) within the National Curriculum Assessment context in England. The authors detail AfL's inclusion within national policy (DfES, 2003) and then describe the survey data and school visit reports which provide their evidence base for examining qualitatively how AfL policy has been understood and implemented in the six years since its national introduction.

Assessment for Learning (AfL) was formalised in 2003 through its inclusion within The Primary Strategy published in the government's Excellence and Enjoyment primary policy (DfES, 2003). AfL had initially entered into British classroom practice in the late 1980s under its synonym of 'formative assessment' through the study and interpretation of the work of theoreticians such as Scriven (1967), Crooks (1988), Sadler (1989), Perrenoud (1991), Ramprasad (1983) by UK researchers such as Christie and Boyle (1990) at the University of Manchester's Centre for Formative Assessment Studies (CFAS), Harlen (1997) and Russell at the University of Liverpool, Black and Wiliam (1998)at King's College, London. AfL has

[1] Professor Bill Boyle, Director, Centre for Formative Assessment Studies (CFAS), University of Manchester & Marie Charles, Assessment Researcher, Pearson Learning Solutions, London.

been defined as 'the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.' (Assessment Reform Group, 2002). Its core principles are to promote the use of assessment to support learning (rather than simply measure it) through an active involvement of the pupil in the learning process and encouraging pupils through self-evaluation to take responsibility for their own learning (Klenowski, 1995). Research in the UK by Black & Wiliam (1998) demonstrated that 'substantial learning gains are possible when teachers introduce formative assessment (AfL) into their classroom practice' (Wiliam et al, 2004, p.49).

Formative assessment was legitimised and became part of the education policymakers' and teaching fraternity's lexicon through the seminal Task Group on Assessment and Testing report (DES,1988) which developed the assessment system for the National Curriculum encompassed by the 1988 Education Reform Act (DES, 1988). However, with the commencement of paper and pencil testing of the National Curriculum (the 'sats') in 1991, soon the only form of 'assessment' which mattered was summative and this was embodied in the end of key stage tests. These quickly became a 'high stakes' priority for schools who felt pressured by both Ofsted (Office for Standards in Education) and the government who used the test results as the principal (often, it appeared to teachers, the sole) measure of national standards and each school's success or failure. This was a very one dimensional 'standards agenda' as its sole focus was on a school's test scores based on sub-domains of English and mathematics measured against arbitrarily set national percentage targets.

Officially summative Teacher Assessment (TA) has 'parity' (Dearing, 1994) with the test outcomes – but the school performance 'league' tables use only the test data . The (non-formative) purpose of TA was designed to be the holistic award of a teacher judgment 'level' for each child at the end of the school year. This attainment judgment was based on the child's progression through an eight level scale, the judgments to be made as a 'best fit' of the child's 'performance' against a prose paragraph describing performance at each level (Boyle, 2008; Hall & Harding, 2002). This task required standardisation of definitions of quality (at school, regional and national levels) for any judgments to be transferable as reliable and valid. 'Unless teachers come to this understanding and learn how to abstract the qualities that run across cases with different surface features but which are judged equivalent they can hardly be said to appreciate the concept of quality,' (Sadler,1989, p.128). This necessitated dialogue, communication and collaboration by teachers with their colleagues within and essentially across schools and as this strategy was financially unsupported by central government it was soon 'dismissed' by teachers. Their reasons included 'workload', difficulties of communication, administration and logistics of meetings to share understandings and meanings of pupil work. Significantly, the 'sats' scores were conveniently received by schools before the date for national returns of TA enabling schools to avoid disagreement between test and TA and reduce workload by returning as near a match as possible across the two scores (Reeves, Boyle & Christie, 2001). The test and TA reported levels were in accord so there appeared to be no need to further investigate a school's performance. The TA process has become even further obfuscated with the introduction of (currently piloting) Assessing Pupil Performance (APP), a government strategy which stresses the making of judgments at sub-sub-levels e.g. high 2c,secure 2c, low 2c (assessment@qca.org.uk).

Assessment for Learning was then publicised by the Assessment Reform Group (ARG) around the millennium with the production of a set of 10 principles (ARG 2002) and this work was formalised within the government's Primary Strategy: Excellence and Enjoyment (DfES 2003). The message contained within the government's juxtaposition of excellence and enjoyment was that there was a 'second way' of achieving measurable excellence in teaching and learning outcomes. The current first way i.e. tests and summative TA consisted of overt assessment OF learning measures with the attendant issues of teaching to the test and time spent on test preparation rather than on teaching for learning (Boyle & Bragg, 2006). The introduction into primary education policy of assessment FOR learning (i.e. assessment information being used in support of learning rather than for providing summative measurement of pupils' performance) was similarly focused on achieving the government's desired result of 'raising standards; hence excellence and enjoyment. The Primary Strategy emphasised the principles of AfL and encouraged teachers to import these principles, in brief these were: Assessment for Learning is part of effective planning, focuses on how pupils learn, is central to classroom practice, is a key professional skill, is sensitive and constructive, fosters motivation, promotes understanding of goals and criteria, helps learners know how to improve, develops the capacity for self- and peer assessment and recognises all educational achievement (ARG, 2002). These principles unfortunately were reduced by government 'strategy consultants' to a 'shopping list' of things to do which teachers could be trained to operationalise, i.e.: sharing learning objectives with the pupil; using written comments to 'feedback' to pupils rather than supplying marks or grades; using 'open' questioning rather than 'closed'; involving pupils more in their own learning process and introducing peer and self-assessment strategies.

The main focus of this paper is to report, from our survey evidence, what has happened six years after the Primary Strategy. Have any changes taken place in teaching and learning practices? Despite the national Standards agenda of 'hitting' targets, raising percentage success rates at national benchmarked levels and surviving Ofsted inspections, has AfL changed teaching and learning habits in primary classrooms?

Methodology

Firstly we designed a questionnaire for a representative (based on a random 25% of the total of primary schools) national sample of 4,000 primary schools to collect evidence of those schools' level of prioritisation of formative assessment as a philosophy for teaching and learning and whether that reported level of prioritisation extended into school practice. We asked schools to state which key aspects of formative assessment they utilised in their planning for teaching and learning; how formative assessment supported learning outcomes in the school; what links the teachers saw between formative assessment and learning and, specifically to address one of the AfL strategy 'key issues', how teachers actively involved children in their own learning.

Our second methodological decision was to have a range of qualitative investigations: these included sampling Local Authorities and talking to their school support staff, observing teaching and having conversations with as many senior leaders and classroom teachers as possible in the schools visited. We acknowledge the relatively low sample although the response rate is more than the anticipated normal rate of 10% for mailed surveys (Fresch, 2007, Alreck & Settle 1995). We did not use a telephone survey to follow-up responses

because years of longitudinal research (monitoring the national curriculum 1997-2007, QCA) has proved to me that the hours invested in contacting the required person do not produce an equitable return rate.

We visited 43 of the schools to observe teaching and learning. We selected those visits from the responses which after content analysis near-matched our own construct for formative teaching (based on Perrenoud 1991, 1998; Allal 2005; Sadler 1989). Analysis of the observations and transcripts (systematic observation schedule based on Galton et al, 1980 used by Alexander in his 1997 survey of 60 schools) indicated that the observations evidenced a profile of rigid, non-formative teaching, 'the formalism of highly structured lessons, whole class plenaries' (Alexander 2005, p.21). In nearly every case we were handed a formalised lesson plan which was rigidly structured from introduction to plenary and from which the teacher did not deviate to accommodate emerging learning needs. The focus was on the production of summative outcomes for measurement purposes 'just one kind of teaching, traditional direct instruction' (Alexander 2004, p 10). The majority of the teaching time was focused on English and mathematics (Boyle & Bragg 2006), specifically on the types of questions and product which were required for national test success.

Findings from the questionnaire survey

394 (9.8%) schools responded to our survey instrument, the normal rate for mailed surveys (Fresch, 2007). On being asked what importance they gave to formative assessment in their planning, 67% (223 of the responding schools), over two thirds, responded that they gave it a very high priority (90% responded that they gave it a 'high' or 'very high' priority). However on being asked to elaborate on 'why' they had assigned such a high level of priority, the schools supplied a range of responses. Some of these did not have a strong relationship between assigning a priority and the supplementary question 'why' (see Figure 1).

The main classifications of response on this question emerged as follows: approximately 40% of the sample reported that they had given a very high importance to formative assessment because it 'informs next steps' or 'it informs the next teaching plan', both of these responses were considered and counted in the same category. The next categories most reported were: 12% of schools reported that formative assessment 'informs all our planning', 8% stated that they gave a very high priority to formative assessment because it 'helped them assess where children are'. We felt this was vague but in the context of an open-ended questionnaire without telephone interview follow-up, it was as good a category description as possible for this aspect of formative assessment. Eleven percent of the sample reported that formative assessment enabled 'personalised learning' and this justified the high priority they gave to formative assessment.

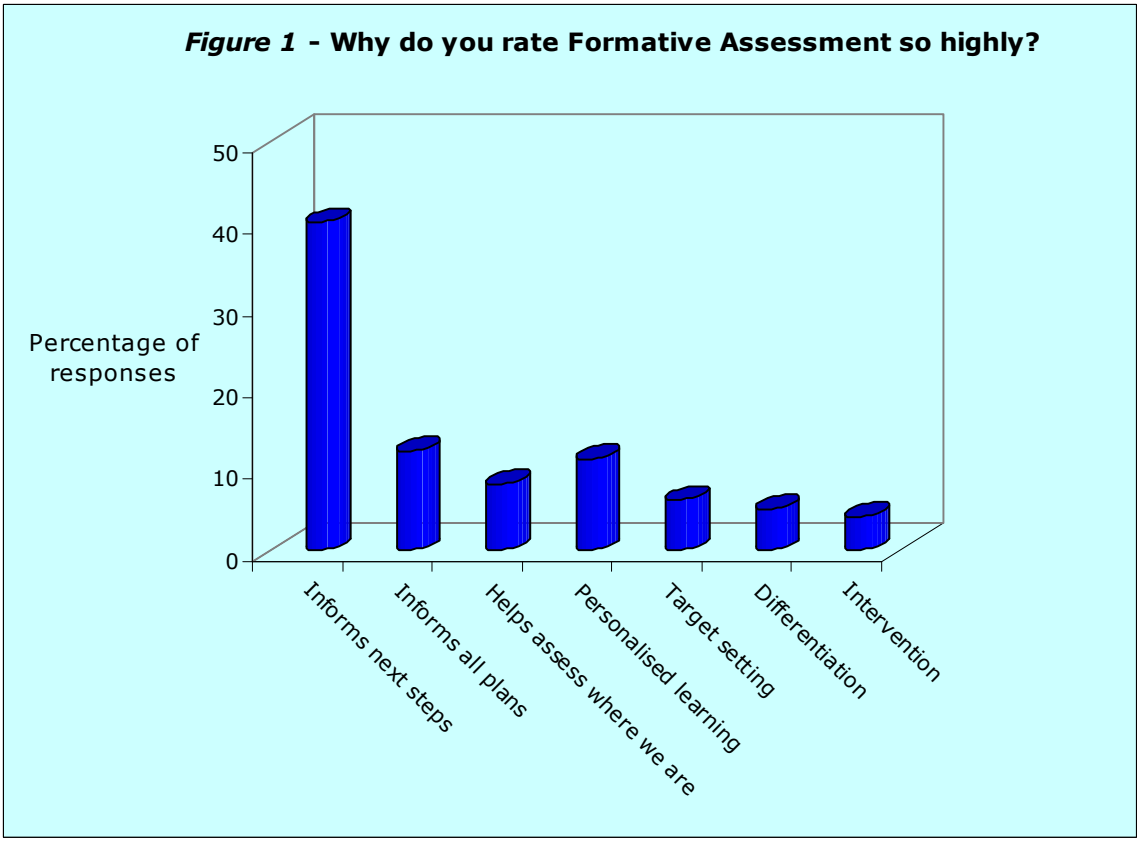
Six percent of respondents stated that formative assessment supplied 'an accurate way to set targets'. The only other significantly reported reason for the high priority given to formative assessment in planning was that 'it supports the identification of pupil needs enabling the setting of differentiated targets for lessons', this from 6% of the sample – a clear indicator that the notion of differentiated planning for teaching is not seen as a pre-requisite for formative assessment by the majority of teachers

There was then a wide range of low frequency responses across the schools, which we have tabulated in the figure as 'other'. In summary these included: 'child's personal next steps' ;

'informs pace and value added'; 'targeted activities'; 'effective comments for the child'; 'generates flexible teaching groups'; 'change planning to cater for pupils' needs'; 'match work to pupils' needs'; 'enjoyment'; 'accurate picture of what children are learning'; 'recommended by Ofsted' and 'a requirement says the SIP'. Equally low frequency but possibly more valid representations of what formative assessment means for the authors were supplied by 'teachers to be highly responsive to child's needs/adapt and adjust daily'; 'update plans on a daily basis for each child' and 'instant feedback to children' and more of the same.

As can be seen from the above despite the very high percentage reporting prioritisation of formative assessment, schools clearly have very different definitions of what it is and what is its purpose. The correlations between Q1a: what importance do you give to formative assessment and Q1b: what is the reason for that prioritisation, showed no significant relationship.

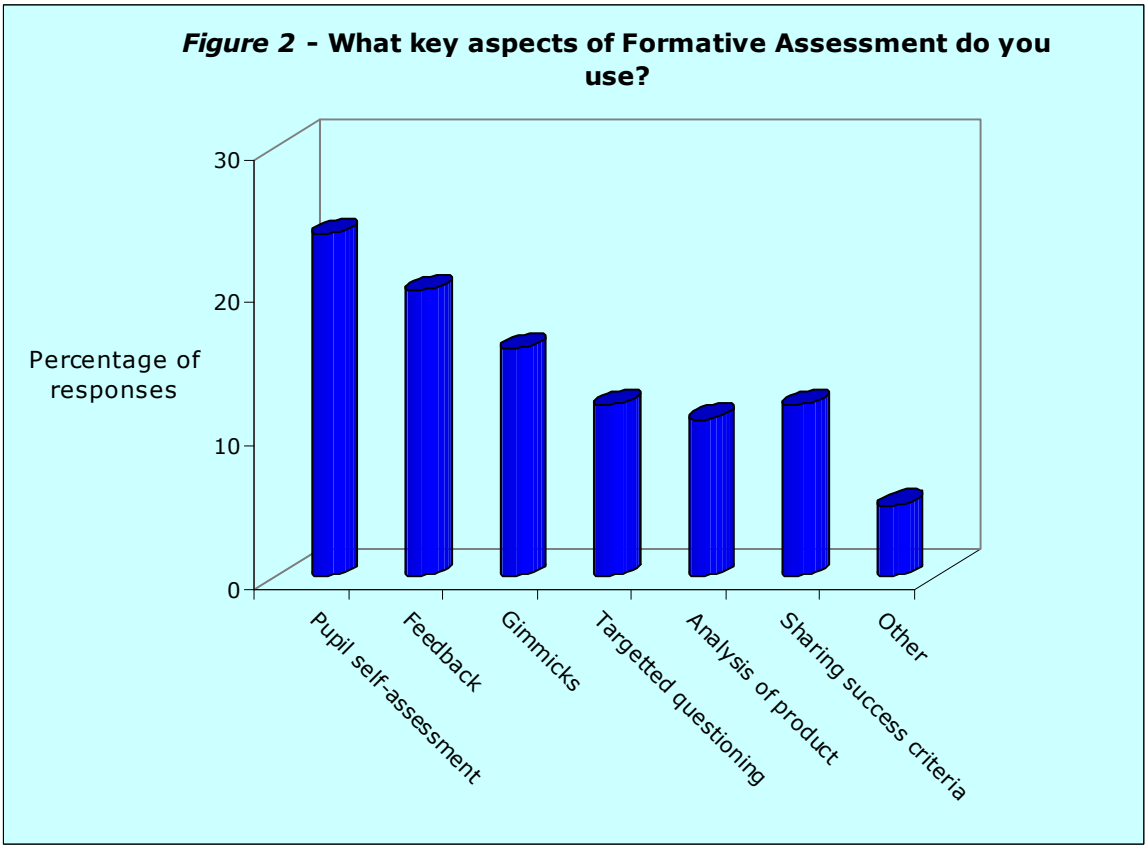
Figure 1: Why do you rate formative assessment so highly?



In question 2 schools were asked which key aspects of formative assessment they used. As there was no 'supplied' list this gave an opportunity to note and analyse what schools would determine as key aspects of formative assessment (see Figure 2). The most highly reported aspect, by almost one in four schools (24%), was 'pupil self-evaluation/self-assessment'. The definitions of this category varied e.g. 'self assessment – checking off against given success criteria'; 'self-evaluation (traffic light system)'; 'self assessment against targets'; 'self and

peer assessment is used to assess understanding' and 'identify individual pupil targets which are used by the pupils to assess their own performance.' The second highest reported key aspect of formative assessment was 'providing feedback to the learner' (20%) with the definitions of that feedback including 'regular marking and feedback'; 'feedback on completed work'; 'feedback during lessons'; 'formative feedback when marking books' and 'feedback on targets set.' Sixteen percent of the responses reported 'gimmicks' related to their key practice of formative assessment e.g. 'two stars and a wish'; 'WALT, WILF and TIB' and 'traffic lights/thumbs up'. Twelve percent of schools reported 'targeted questioning', 12% 'sharing success criteria' and one in 9 schools (11%) reported 'analysis of product' as key aspects of formative assessment which they used.

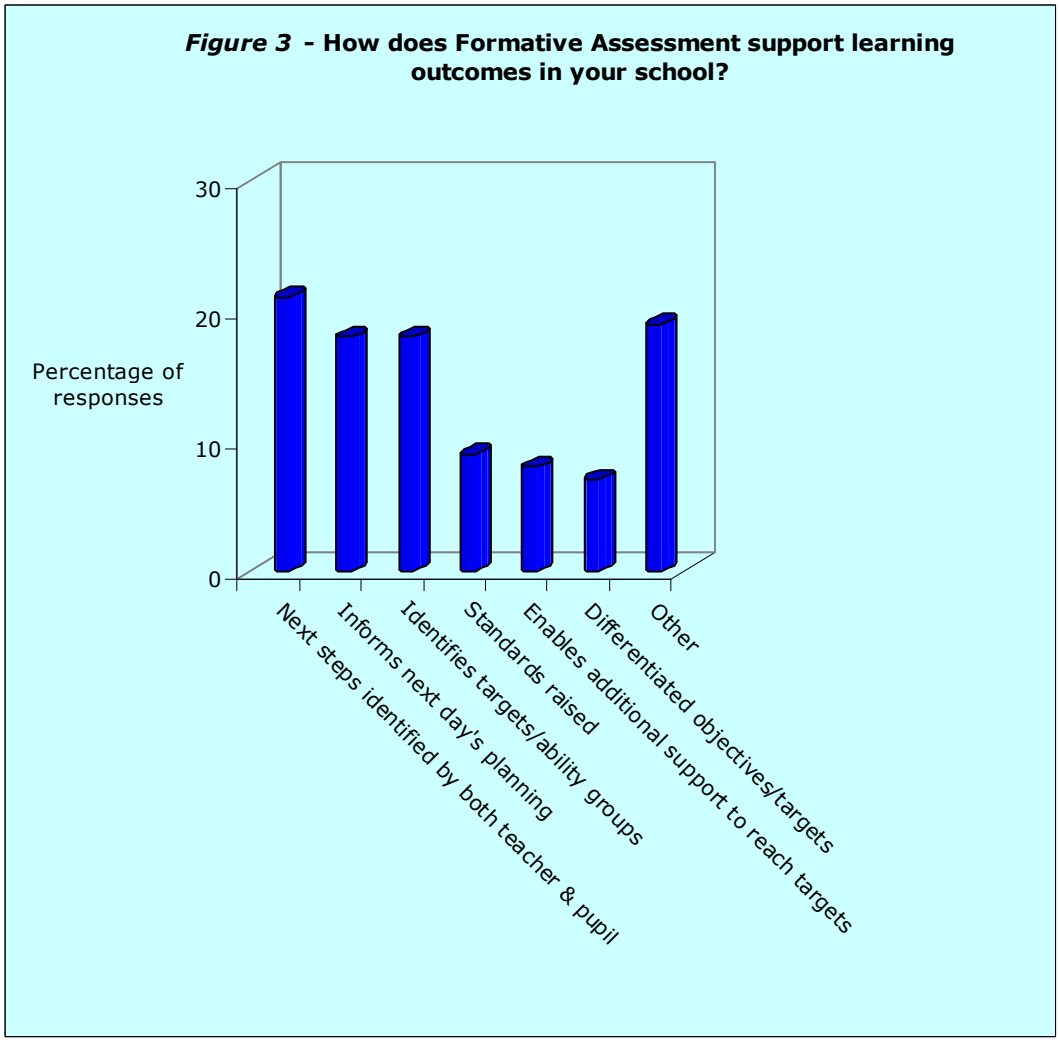
Figure 2: Which key aspects of formative assessment do you use?



One in twenty schools reported their key aspect as 'sharing learning objectives/success criteria with children'. These were the significant number responses identifying key aspects of formative assessment; there was also a scattering of wide interpretation but individual reports e.g. 'working alongside children', 'leveling/moderation of work', 'teacher and pupil setting targets together', 'checking children against targets', 'promoting children's learning', 'key questions: what do we know? what do we want to find out?' and 'APP/single level tests'. These responses led the authors to believe that formative assessment has no common understanding across teachers either in definition, components or aspects of practice.

Question 3 asked teachers to report how formative assessment supported learning outcomes in their schools. The responses (as with Questions 1 and 2) indicated a range of understandings not only of what formative assessment is but of what learning outcomes are now classified as ('achieving targets') and the link between assessment and learning ('enables additional support when not achieving targets'). The most supported response to the question how does formative assessment support learning in your school was 'next steps identified by both teacher and pupils' (21%) and that was regarded as both positive (teacher and pupil described as working together to identify next steps in learning) and formative (see Figure 3).

Figure 3: How does formative assessment support learning outcomes in your school?



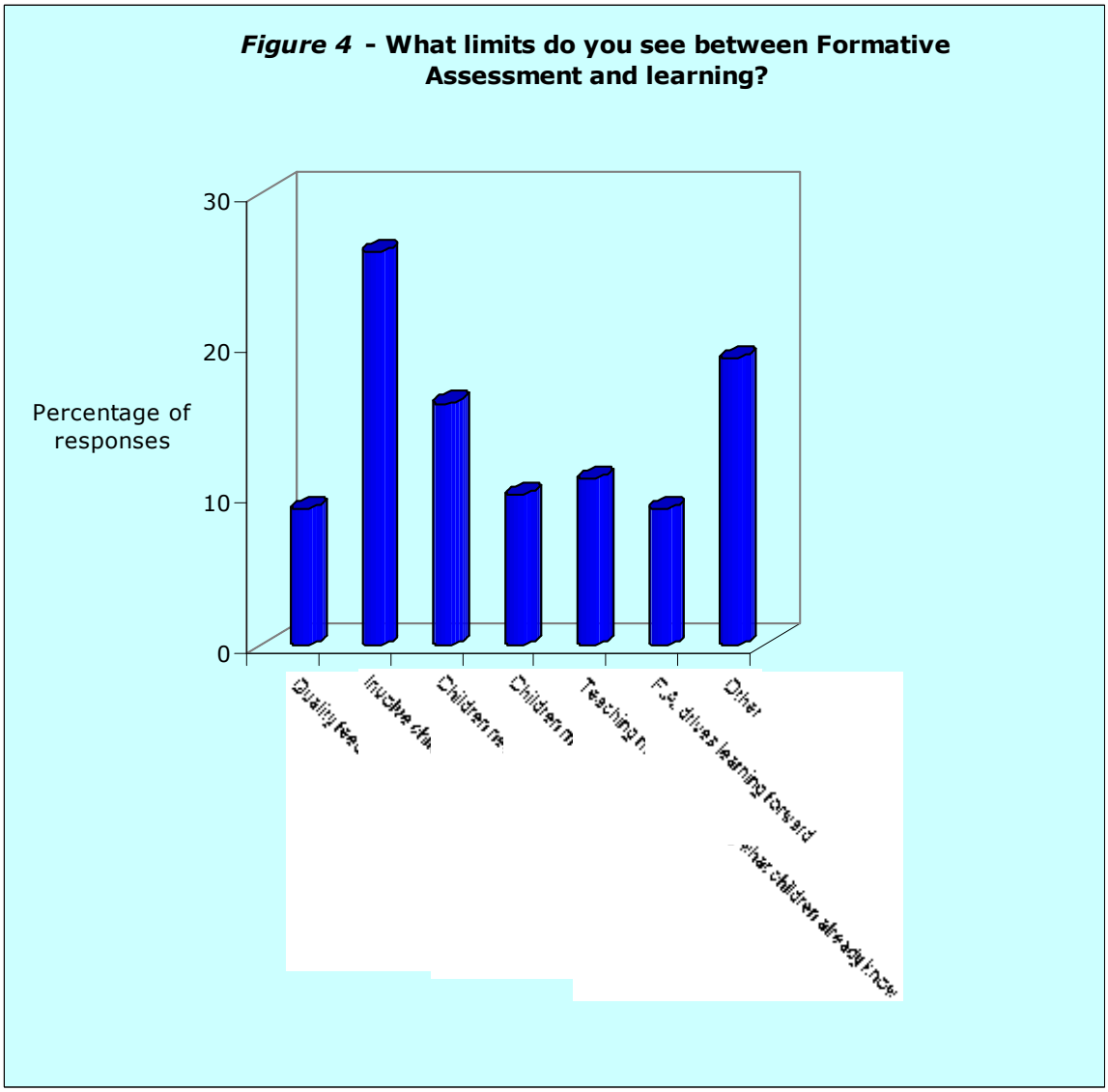
The other significantly (in numeric terms) reported responses were 'informs next day's planning' (18%), which was at least formative, 'planned to match differentiated objectives and targets', which hinted that it might or might not be formative and then 'identifies targets and ability groups' (18%), 'standards raising/achieving targets' (9%) and 'enables additional support/not achieving targets' (8%), all of which were not, in the authors' view, either

formative or supporting learning. There was a range of low frequency responses covering the possibly formative e.g., ‘individualise assessment for each pupil’ and ‘small steps which are reviewed and adjusted’ through the unspecific e.g. ‘central to learning process’, ‘pupils empowered’ ‘helps pace of lessons’ to the summative e.g. ‘enables teachers to make predictions’, ‘children’s individual half term targets’, ‘grouping of children relative to academic progress’ and ‘analysis of data allows appropriate targets to be set’.

Survey question 4 probed the schools’ views of the links between formative assessment and learning. The responses ranged in specificity from the generalisation of ‘they are inextricably linked’ to ‘children need to know how to continue to improve’ (16%). The most reported response was ‘involves children in measuring their own learning/increases confidence’ (26%) which the authors felt summed up two positive aspects of formative assessment and supplied a link between the assessment and learning. Other responses which had some numerical support were ‘children cannot move in learning unless AfL is in place’ (10%), ‘teaching must be driven by what children already know’ (11%) and ‘learning has to have formative assessment to move it forward’ (9%), all of which we thought were too general to detail any specific link between formative assessment and learning.

There was the usual range of low frequency responses which we have encapsulated in the ‘other’ classification column (19%). These ranged from the esoteric and unspecific e.g. ‘it’s a continued cycle, teacher challenges children and keeps them motivated’, ‘so we all have the same philosophy’ and ‘the greater the quality of the formative assessment the deeper the learning process’. Then we received the vaguely formative but at least linked to learning e.g. ‘if pupils do not understand a concept this must be returned to’, ‘assessment is seen as an integral part of teaching and learning’, ‘quality feedback to signpost areas of work they need to concentrate on’ and ‘involves children in actively monitoring what they have to do next’. The latter signaled both the active involvement of the child in the process (a crucial component of formative assessment) and gave the authors hope that the ‘next’ referred to was actually a specific micro-learning step rather than a generalised statement of intent. The hope was soon dampened by a school response which ‘formatively’ stated the link between formative assessment and learning as ‘only as a means of testing’.

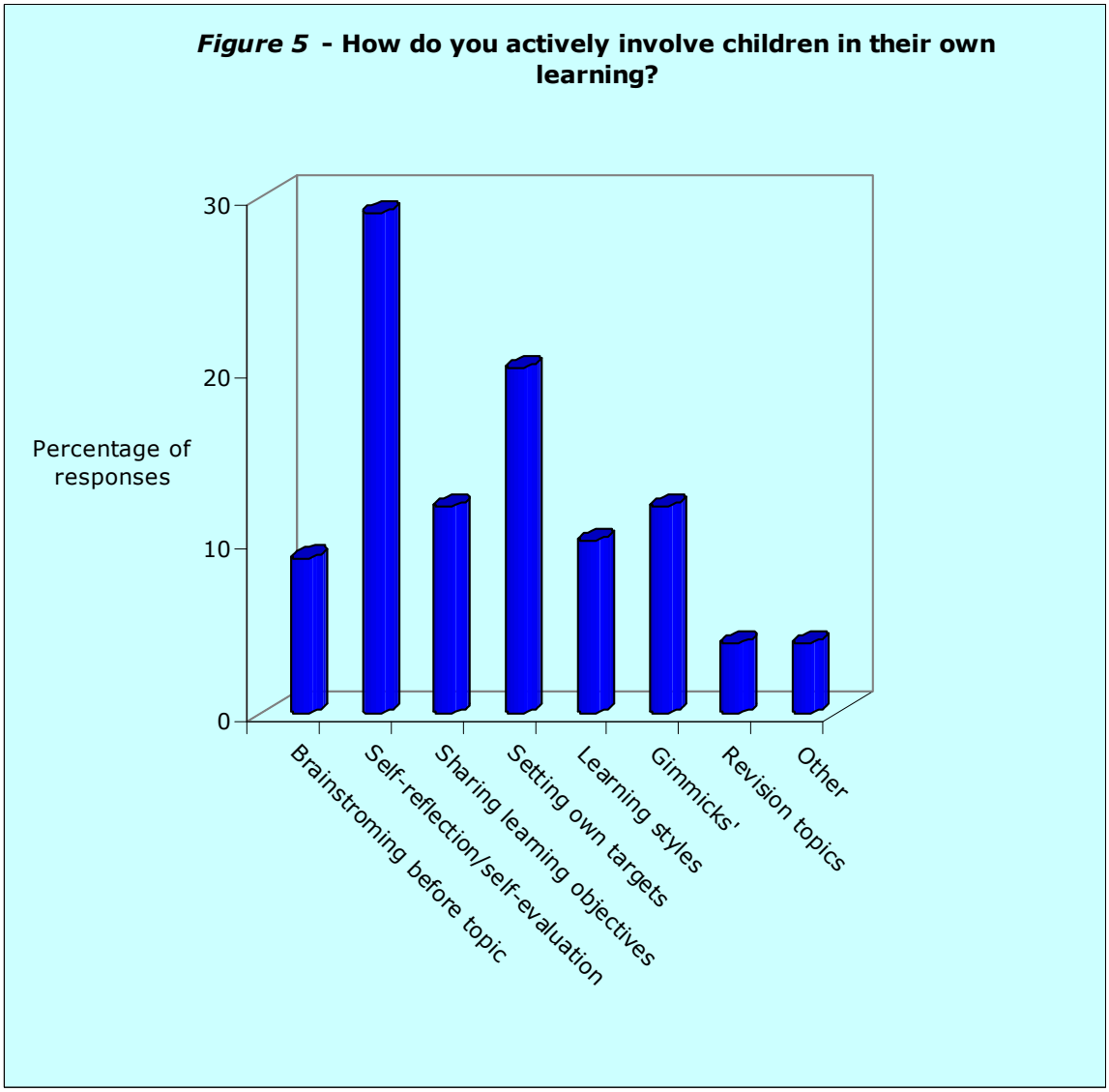
Figure 4: What links do you see between formative assessment and learning?



Survey question 5 specifically focused on the AfL principle of actively involving children in their own learning and asked how this was being done in practice. The highest supported response (29%) stated that children were involved in their own learning through 'self-reflection/self-evaluation' which seemed ambiguous to the authors as it was not clear (until we observed the teaching sessions) when, how or if this self-reflection took place or the results of the self-reflection transferred into active involvement in learning. One in five schools (20%) identified 'setting own targets/reaching own targets' as an active involvement while 10% of schools reported 'learning styles/what they like to learn and how'. Twelve percent of respondents reported 'gimmicks' as the route to actively involving children in their own learning e.g. 'thumbs up/down/sideways', 'WALT, WILF', 'star checkers', 'two stars and a wish', and of course the ubiquitous 'traffic lights'. The range of individual responses covered the bold but unspecific 'HOW COULD YOU NOT!' (sic) through

'enthusiasm', 'circle time' and good plenary sessions' to the summatively oriented 'revision topics'. None of the responses matched the authors' understanding of the involvement of children as co-constructors, self-regulated learners and negotiators of their own learning.

Figure 5: How do you actively involve children in their own learning?



Discussion and Implications

Following those varied definitions and understandings we have still to see a formative classroom or a teacher whose pedagogy is based on formative principles on any of our forty three observation visits to schools, despite these visits being based on the most 'formative' responses to our survey. We are however seeing a succession of teachers following a formula of planned predictability, controlling the content of the 'three part menu' which is being delivered to passive children. This rigidity has its pedagogical roots firmly in the National

Numeracy and Literacy Strategies with 'the formalism of highly structured lessons, whole class plenaries' (Alexander, 2005). Among the concessions to AfL, from our observations, are the 'solitary' learning objective on the interactive whiteboard (how can one learning objective be adequate for the range of learning needs present in every class?). Our interviews provided some informative responses on that theme, such as 'But I have the top set (for literacy)' indicating that Teacher C believed that she did not need to differentiate within a class which had been streamed as a top set for literacy. We observed and were treated to rhetoric about 'doing' self and peer assessment and the free scattering of terms such as 'open questions', 'feedback' etc.

Our anticipation was that after six years of a national AfL strategy we would not just see the isolated individual formatively teaching but there would be in a majority of schools a shared learning community of formative teachers working collaboratively with children at the centre of the whole school's teaching and learning ethos and culture (Allal, 2005). From our survey responses and our observations in the classroom, the following five issues have emerged. Teachers for whatever legacy or conceptual reasons clearly have problems with differentiation; 'differentiation implies the imposition of different curricula for different groups of pupils - or it means nothing,' (Simon 1985, p.6); the following of a formulaic 'lesson plan' seems to be the sole pedagogical model so there is no 'divergence'; understanding of formative assessment (or its synonym, AfL) in practical operation is poor so no clarity of definition; 'coverage' has precedence over depth and security in learning and the associated 'jargon' around the simple truth of formative teaching needs demythologising. These issues are itemised below with an introduction to each supplied by a theorist in the field.

Differentiation

> 'If formative assessment is carried out on a fairly regular basis, the result is pressure to differentiate' (Perrenoud 1991, p.89).

The evidence from our sample indicates the notion of differentiated planning for teaching is not seen as a pre-requisite for formative assessment by the majority of teachers (only 7% of our sample states that formative assessment enabled the setting of differentiated targets for lessons. This is a still-strong legacy from the National Numeracy Strategy with its discouragement of differentiated teaching 'we are concerned that children should not continue to work at many levels, with the teacher placing them in a wide range of differentiated groups (DfEE. 1998, p.54) and 'its format as a fixed curriculum to be taught to all pupils regardless of attainment indicates that very little curricular differentiation is recommended' (Brown et al, 1998). So strong that five years after Excellence and Enjoyment: The Primary Strategy we are observing a majority of lessons in which differentiation is totally absent. Why are we observing lessons with one static learning objective which embraces the whole extent of a class's learning? This signals two things. The pedagogical messages of the National Literacy and Numeracy Strategies in which differentiation was frowned upon and the absence of teacher understanding of the need for a differentiated menu to match the range of learning needs and the presence of a 'one size fits all' mentality are proving difficult to shift.

In conversation with teachers during our 43 school visits we raise the word 'differentiation' and the vagueness of the responses begins. We are told 'I set one task and then I differentiate

by what they produce' or 'I have an extension task ready for those who finish'. Our observations indicate that this is what AfL in practice has been reduced to. The responses demonstrate misconceptions of the basic principles of how children learn and the sacrifice of developmental learning on the altar of 'coverage', 'pace', 'moving the cohort on' and 'getting through the pre-planned package'. Clearly our initial teacher training programmes need to return to the conceptualisers and theorists of formative assessment to change technicians back to pedagogists.. Principal amongst these theorists is Philippe Perrenoud whose philosophy is based on 'to the extent that pupils do not have the same abilities nor the same needs nor the same way of working, an optimal situation for one pupil will not be optimal for another…one can write a simple equation: diversity in people + appropriate treatment for each = diversity in approach' (Perrenoud, 1998, p.93-4). In even simpler terms, 'good teaching forces differentiation' (Perrenoud, 1998) is called for. Linda Allal reinforces the point, 'differentiation of instruction is planned rather than just being added on after observing difficulties' (Allal 2005, p.246).

Divergence

> *'So in the face of pace, objectives, targets and tables that have become part of the dominant linguistic and conceptual discourse of education reform in England, we might wonder how confident good divergent teachers will be to stray from pre-set paths for better pastures. We might wonder what the absence of divergent thinking will mean, in the longer term, for children's motivation and interest in their learning experiences' (Dadds 2001, p.53)*

Dadds in 2001 described a scenario which has further deteriorated by 2009. In our classroom observations we keep looking for the first teacher who 'diverges' from the norm of the pre-packaged lesson. This is delivered (usually script perfect) around a sole common learning objective (or alternatively in some cases a 'whole class task') to the class who are then invited either (i) to talk to their partner, or (ii) to complete a common task, or (iii) indulge in an 'AfL professional development day' gimmick ('snowball', 'traffic lights', etc) or (iv) to wake from their lethargy induced by this format being repeated day after day, to recall some of the detail from the teacher's (lengthy) contextualising or introductory remarks. We see teachers 'covering' work at pace, we see teachers 'controlling' and imposing the narrowness of the learning agenda and forgetting that the learning is not in the tidiness of the schema but in the response, the involvement, the energy, the interest of the child as participant learner; learning is a 'messy' and complex process not a neat and tidy one. Therefore their misinterpretation of divergence as inevitably resulting in chaos and reduction in quality must be challenged. We need to see but are not, teachers having the confidence to relax that control, to allow children to be involved in the 'why?' am I doing this, 'what if?' and in the 'how?' can it best be done, and encouraging collaboration and conversation and children setting personal progressive targets.

In the classroom example below children were consciously or sub-consciously connecting prior learning to a present theme and they were re-drafting openly and orally their developing conceptualisation of counting in tens in a non-rigidly controlled classroom environment. This is an example, rare in our observations, of co-construction between the teacher and the child enabling the children's dialogue to expand by non-intervention from the teacher at the point of the first child's question thus enabling the children to 'drive' the learning direction.

Mathematics lesson Year 1 children

Learning focus: counting in tens (10 more/10 less)

Context: The previous week the children had explored the concept of odd and even numbers. In this lesson the whole class was on the carpet exploring counting. The teacher recorded the following dialogue which took place as the children worked on grouping as part of the process of understanding the concept.

Teacher (T); Let's count to 100 in tens.

Burhan: three sets of ten make 30 but it is an odd number.

Mohammed: is it an odd number?

Burhan: Yes, it is odd.

T: Well is it an odd number?

Burhan: If you had three people, one would get 10, one would get 10 and one would get 10.

T: What about two people?

Reem: One person would get 5, 5 and 5.

T: How many is that?

Reem: 15.

T: What is that doubled?

Reem: 30.

T: Burhan, you can share 30 as 15 and 15.

In this formative classroom situation the children were demonstrating the following: they were consciously or sub-consciously connecting prior learning to the present theme and they were re-drafting orally and collaboratively their developing conceptualisation of 'counting in tens' in an open classroom culture. This is a genuine example of co-construction between the teacher and the group of children (through enabling the children's dialogue to expand and by non- intervention at the first child's question) and of divergence from a planned format to enable the children to 'drive' the learning direction.

The misconception i.e. that the digit 3 makes 30 into an odd number is explored and rectified in group discussion. The teacher, by not closing the learning agenda by responding with an answer to the first child's question, has enabled the children to orally work through two concepts i.e. multiples of 10 and odds and evens.

In conversation the teacher reflectively observed 'I should have given Burhan, Reem and Mohammed a task outside the main group to explore their own numbers'.

However unlike the above example the norm is that far from the formative principles of involving children in their own learning, teachers are controlling the learning agenda even more firmly. 'Many schools give the impression of having implemented AfL when in reality the change in pedagogy that it requires has not taken place. This may happen when teachers

feel constrained by external tests over which they have no control. As a result they are unlikely to give pupils a greater role in directing their (own) learning.' (ARG, 2007, p.9)

Definition

> *'Formative assessment takes place day by day and allows the teacher and the student to adapt their respective actions to the teaching/learning situation in question. It is thus, for them, a privileged occasion for conscious reflection on their experience' Audibert, 1980, p.62.*

On our visits to schools we ask teachers how they define their teaching, how they conceptualise their role, what their philosophy of teaching is. 'In terms of her philosophy for teaching and learning – this was something teacher F had not given any thought to. Much of her practice she claimed was based on the modelling of others she felt were worth copying.' (Case study school 5). We ask teachers if they think they are 'formative' teachers. To which they have replied in the majority, 'what does that mean, I have never heard of that before?'. We are wondering if teachers in 2009 need a philosophy. From 207 responses to a survey question 'what is your teaching philosophy based on?' and from 13 case study visits the typical response was 'that's a really hard one - I've never been asked that before' (school X), they certainly don't think they do. They are 'reliant on prescriptive centrally disseminated materials from which 'politicians and bureaucrats are demanding greater conformity of education offerings which are transparent and superficially testable' (Patrick, Forde & McPhee, 2003, p.239). They have 'Strategies' for most of the important things i.e. numeracy, literacy, AfL, and they have 'Frameworks' to plan to and from and they have centrally supplied schemes of work to save the need for matching teaching material to developmental or interest levels, in short they have been reduced to technicians. If they follow these formulae they are 'safe' and 'secure' in the accountability and auditing processes conducted by their own Senior Leadership Team. So, in summary the central purpose for becoming a teacher has been lost. Our belief is that teachers need to understand and to embrace what formative teaching is. It is not disguised within a programme or strategy to improve 'level scores' and it should not respond to the summative bell but instead to the learning needs of the child, involving the child centrally in the origination and the development of his/her learning. One example from our observations centres on an experienced teacher who expounded widely on her formative practice, her current studying for a further degree and then spent the 40 minutes teaching time dictating language, content, control and materials in a closed format which did not enable the children to connect to or be involved in their own learning. In short, they did not see the relevance of the lesson (Vygotsky, 1986).

Depth

> *'Teachers bring skills in devising and constructing tasks to elicit revealing and pertinent responses from children.' (Sadler 1989, p.80)*

Depth of learning: this equates with the immersion of the teacher and the child in the teaching and learning process. Our search is to identify through our observations of teaching and in conversations with children, teachers and LA officers their priorities in planning for teaching and learning. How is an independent and lifelong learner developed? Is there a relationship between the intrinsic development of engagement, self-motivation, interest and research skills at an early stage of a child's education and current pedagogical practice? Is

the current observed paradigm of controlling teacher/passive recipient moving at pace through a prescribed programme going to develop a generation of 'deep and reflective thinkers' and lifelong learners? From our recent classroom observations the authors' response is that in the current summative framework the chances of developing reflective children involved in self-motivated research activities is negligible. A missing component is the acknowledgement of the child as learner within the affective domain, in short acknowledging that social development is equally important as and a primary factor in cognitive development – but the latter is the area on which curriculum and assessment and therefore pedagogy focuses in a summative Standards agenda. The importance of a nurturing pedagogy is recognised by Reeves (1993) who argues that 'if we are to take quality seriously we have to get closer to our learners, their needs, their learning styles and their motivation' (Dadds, 2001, p.53). The evidence of our observations across the 43 schools justifies the necessity of reminding the teacher that he/she is working with discrete individuals, all with emotional and learning needs, not just delivering a centrally devolved teacher-controlled subject diet. Wink suggests that pedagogy involves human interaction and joy, of playing with new ideas and '[to] challenge all educators to look beyond the complexities and familiarities of their own teaching' (Wink, 2005, in Graziano 2008, p.162). Within the current climate this position seems both irrelevant and unobtainable as the dominant discourse is one of controlling pedagogy and performativity.

Demythologising

> *'The search for theoretical frameworks could lead to an increasingly abstract vision of formative assessment cut off from the realities of classroom practice. This is why it is essential to articulate theoretical work with the study of how assessment is actually practised in the classroom' (Allal 2005, p.251)*

Already AfL has collected too much 'clutter' of terminology; it is dominated by gimmicks (WILFs, WALTs, TIBs and OLIs) rather than focusing on the specific understanding and practical application of formative assessment (FA), assessment for learning (AfL), continuous assessment (CA) and teacher assessment (TA). Just as the Education Reform Act in 1988 issued in a plethora of abbreviations i.e. SAT, AT, SoA, etc, similarly, as our previous sentence illustrates, assessment now has its own potential for confusion through abbreviation. This confusion over terminology derives from a scant understanding of the works of the original formative assessment theorists, misrepresented or 'popularised' by the travelling consultants who see money to be made from the centre-periphery training model for AfL. Is there a pack? we were asked by one teacher, misunderstanding both the purpose of our visit and the purpose of formative assessment. If there isn't a 'pack' do not expect it to be done because that demands experimentation with pedagogy (to be frowned upon by School Leadership Teams and School Improvement Partners), then inevitably deviation (to be frowned on by everybody!) which, contrary to the cynics, produces successful (and deep) formative teaching and learning.

Conclusion

After five years of a national AfL strategy with the attendant government funding and supply of 'consultants' to support the strategy, our anticipation was that there would be in a majority

of schools a shared community of formative teachers working collaboratively with children at the centre of a whole school philosophy of formative teaching and learning. From our forty three school observations this does not appear to be the case; the summative agenda is still firmly straitjacketing the teaching and learning ethos. As can be seen from the survey data despite the very high percentage reporting prioritisation of formative assessment, schools clearly have very different definitions of what it is and what is its purpose. There is a huge range of misconceptions, these cover the most blatant misunderstandings of formative practice e.g. 'testing', 'analysis of data' and 'levelling and moderation' through sweeping generalisations and rhetoric. The correlations between Q1a: what importance do you give to formative assessment and Q1b: what is the reason for that prioritisation, showed no significant relationship. The only other significantly reported reason for the high priority given to formative assessment in planning was that 'it supports the identification of pupil needs enabling the setting of differentiated targets for lessons' from 5% of the sample – a clear indicator that the notion of differentiated planning for teaching is not seen as a pre-requisite for formative assessment by the huge majority of teachers (95%). That latter conclusion is worrying in itself because clearly the legacy of the national strategies with their sustenance of the summative accountability agenda through didactic teaching has survived the introduction of both Excellence and Enjoyment and AfL.

Our research has evidenced the following five issues. Teachers for pedagogical or philosophical, legacy or conceptual reasons clearly have problems with understanding differentiation. Differentiation has been evidenced as equating with setting and labelling children in static inflexible groups that remain constant throughout the time that child remains in the year group (usually based on the sole evidence of test scores from the previous year). 'Differentiation implies the imposition of different curricula for different groups of pupils - or it means nothing,' (Simon 1985, p.6). The following of a formulaic whole class 'lesson plan' seems to be the sole pedagogical model so there is no 'divergence'. The understanding of formative assessment (or its synonym, AfL) and its practical operation is poor so there is no clarity of definition. Pedagogy is driven by 'coverage' and 'pace' which have precedence over depth and security in learning. 'Coverage and elicitation of facts rather than the creation and co-construction of interconnected learning' (Myhill, 2006, p34). The associated gimmicks camouflage what is the simple truth of formative teaching i.e. the child's learning needs at the centre of a teacher's planning, and therefore the concept of formative assessment needs demythologising.

References

Allal L & Lopez M (2005) Formative Assessment of learning: A Review of Publications in French (OECD p.241-263)

Alexander R (2005) Culture, Dialogue and Learning: Notes on an Emerging Pedagogy, paper presented at International Association for Cognitive Education and Psychology Conference, Durham, July 2005.

Assessment Reform Group (2002) Assessment for learning: 10 principles

Assessment Reform Group (2007) The role of teachers in the assessment of learning

Audibert S (1980) 'En d'autres mots…l'evaluation des apprentissages' Mesure et evaluation en education, Vol 3 pp 59-64

Black P & Wiliam D (1988) Inside the Black Box: Raising Standards through Classroom Assessment London: Kings College

Boyle B (2008) Holistic assessment: A model without status in the Assessment hierarchy in England. Paper delivered at an International Colloquium on holistic assessment, Kuching, Malaysia, 4-7 May, 2008

Boyle B & Bragg J (2006) A curriculum without foundation, British Educational Research Journal, Vol 32 (4) pp.569-582

Brown M, Askew M, Baker D, Denvor H & Millett A (1998) Is the National Numeracy Strategy Research Based, British Journal of Educational Studies, Vol 46, (4) pp. 362-385.

Christie T & Boyle B (1990) A Guide to Teacher Assessment Packs A, B & C, Schools Examination & Assessment Council, London: Heinemann

Crooks T.J. (1988 The Impact of Classroom Evaluation Practices on Students Review of Educational Research Vol 58 pp 438-81

Dadds M (2001) The Politics of pedagogy Teachers and Teaching: theory and practice Vol 7 (1)

Dearing R. (1994) The National Curriculum and its Assessment, School Curriculum and Assessment Authority

Department of Education and Science (1988) Education Reform Act, HMSO

Department of Education and Science (1988) Task Group on Assessment and Testing – National Curriculum: A Report

Department for Education and Employment (1998) The Implementation of the National Numeracy Strategy: The final report of the numeracy task force. London: DfEE

Department for Education and Skills (2003) Excellence and Enjoyment; A Strategy for Primary Schools

Fresch M J (2007) Teachers' Concerns About Spelling Instruction: A National Survey, Reading psychology, Vol 28 (4) pp.301-330

Hall K, Collins J, Benjamin S, Nind M & Sheehy K (2004) SATurated models of pupildom: assessment and inclusion/exclusion, British Educational Research Journal Vol 30 (6)

Hall K & Harding A (2002) Level descriptions and teacher assessment in England: towards a community of practice, Educational Research

Harlen W & James M (1997) Assessment and learning: differences and relationships between formative and summative assessment. Assessment in Education Vol 4 (3) pp. 365-379

Klenowski V (1995) Student Self-evaluation Processes in Student-centred Teaching and Learning Contexts of Australia and England Assessment in Education Vol 2, pp. 145-63

Perrenoud P (1991) Towards a pragmatic approach to formative evaluation. In Weston P, editor, Assessment of Pupils' Achievement: Motivation and School Success. Swets & Zeitlinger.

Perrenoud P (1998) From a Formative evaluation to a Controlled Regulation of Learning Processes towards a Wider Conceptual Field. Assessment in Education Vol 5 (1) pp 85-102

QCA Assessing Primary Progress, assessment@qca.org.uk

Ramprasad A (1983) On the definition of feedback, Behavioral Science, 28, pp.4-13

Reeves D, Boyle B & Christie T (2001) The relationship between teacher assessments and pupil attainments in standard tests/tasks at key stage 2, 1996-98. British Educational Research Journal.

Sadler D.R. (1989) Formative assessment and the design of instructional systems. International science, Kluwer Academic Press

Sadler D.R. (1998) Formative assessment: revising the territory. Assessment in Education Vol 5 (1), pp. 77-84

Scriven M (1967) The Methodology of Evaluation American Education Research Association Monograph Series on Evaluation 1, pp.39-83

Simon B (1985) What is differentiation? In Weston P, British Journal of Special Education Vol 19 (1)

Vygotsky L (1986) Thought and Language Cambridge

Wiliam D, Lee C, Harrison C & Black P (2004) Teachers developing assessment for learning: impact on student achievement, Assessment in Education Vol 11 (1), pp.49-65

Aligned Assessments for Improved Learning[2]

Advances in technology, coupled with innovative assessment task design and advanced psychometric and cognitive models, make it possible for us to obtain a richer, more intelligent picture of what students know and can do than ever before. While the historical opportunity to change the direction of education is real, so are the challenges inherent in any change in assessment paradigm. As such, some guiding principles to provide a framework of discussion is warranted.

1. Assessment must be used to improve learning

   While this would seem to be self evident, education reform is rife with examples of assessments for various purpose, some quite distant from improved learning. As such, for an assessment to be effective it must be aligned to instruction and evidence must be provided that it links directly to instruction and improved learning.

2. Assessment should be one component of an integrated learning system

   Assessments need support and can not be used in isolation. They must provide instructionally actionable information—information a teacher can use to improve

[2] Jon S. Twing, Ph.D., Executive Vice President, Pearson Assessment & Information, Iowa City

learning. Additionally however, various stakeholders will want to use the same information in different ways. As such, this same instructionally actionable information will have to feed accountability systems, and inform judgments about teacher development needs and evaluation, for example. While many people will claim this is asking too much of an assessment, the reality is that such multiple decisions are currently being made and are likely to continue to be made into the future. As such, assessment experts should delineate what types of assessment data are most appropriate for which decisions under what circumstances and stop arguing that multiple decisions are not useful.

3. Assessment must exploit technology if we want to measure meaningful learning attributes.

   We can no longer in large scale assessment rely on the measurement of content mastery only, but rather must incorporate skills-based measures such as problem solving, critical analysis and collaboration where appropriate. This will surly mean that performance assessments, scenario-based tasks, simulations, data collections, generation and editing of student information will all be required. If so, then only a technology delivered assessment (online testing) coupled with technology aided scoring (computerized performance scoring or machine scoring via artificial intelligence) can facilitate collecting such measures.

4. Assessment must account for a transition from where we are now to where we want to be in the future.

   Such a learning revolution (i.e., more measures of problem solving, critical analysis and collaboration) coupled with the logarithmic increase in technology infrastructure to support it, particularly in the international assessment area, can not happen over night. As such, there must be a transition plan that moves us from our current assessment paradigm to the new technology laden arena. Decisions regarding priorities will have to be made. For example, in the US, the summary outcome measures are being funded first. In Abu Dhabi, interim assessments and growth measures are now being considered after the implementation of accountability assessments.

5. Assessment must be practical and implementable and must not sacrifice what we want to measure due to limitations in psychometric, cognitive or technology models.

   While the psychometric properties of reliability and validity are important we will no longer be able to refine and constrain the construct to be measured in order to achieve acceptable levels for these parameters. We can not simply measure content mastery with multiple-choice achievement tests because we know how and can do so reliably. We have already seen degradation o f the curriculum and over contextualization of the measures to make them useful for meaningful learning improvement. Similarly, we can not wait for new psychometric or cognitive models to solve this dilemma for us as these could be years away. Rather, what we need to do is understand that a variety of information (some objective, some not; some directly observed, some not; some more defensible for one purpose than another) will coexist inside this learning

system and that our parameters of psychometric excellence must accommodate it. This can be accomplished with the expanded use of a validity framework that ask "…what evidence do you have to justify such inferences made from the data?" and to contextualize the data used in these inferences. For example, if teachers will administer performance assessment tasks that will then be combined with online delivery of objective items, what evidence has the test developer collected to compare the reliability and standardization of the teacher scoring as compared to the objective scoring. Only then can knowledgeable inferences be made from the combined scores.

6. Different stakeholders will want different and often competing outcomes and this will be a mater of fact for the assessment.

   The learning assessment systems of the future will simply have to provide information for different stakeholders. Actionable information teachers might use to improve instruction, program evaluation and accountability information policy makers might use for system decisions, measures of teacher effectiveness, measures of college and work place readiness, international benchmark comparisons and other initiatives will want to use this data. Only a robust learning system that is planned with these and other multiple purposes in mind will survive such demands.

When digital natives become the majority, our current thinking about things like: Technology infrastructure (Computer labs for example), Grade-based or age-based enrollment instead of enrollment based on learning progressions, Transadaptations to other languages instead of building assessments in the native language by native language speakers, manual systems requiring multiple logins or paper history files or "permanent records" as well as traditional textbooks and homework assignments will all be things of the past. We need to get ready for this revolution now!

A Practical Implementation Example

Such a learning system led by the guiding principles previously outlined is not simply "pie in the sky" desire. In fact, we have many examples of components of such systems that are ready or nearly ready for implementation now. Online marking schemes, automated scoring systems using artificial intelligence, online tutorial systems, online interactive testing systems and voice analysis feedback are examples of technology enabled assessment enhancements that are currently being used. Pulling these together into one system that is aligned with instruction is the next step. The following provides one example of one such interim system that meets some of the goals of the guidelines review and may set the stage for how such a future system might look like.

Figure 6 shows a high level organizational scheme that outlines the attributes of an existing learning system as an example of one such system that can meet the needs of the educational challenges moving forward.

Figure 6. Limelight Example

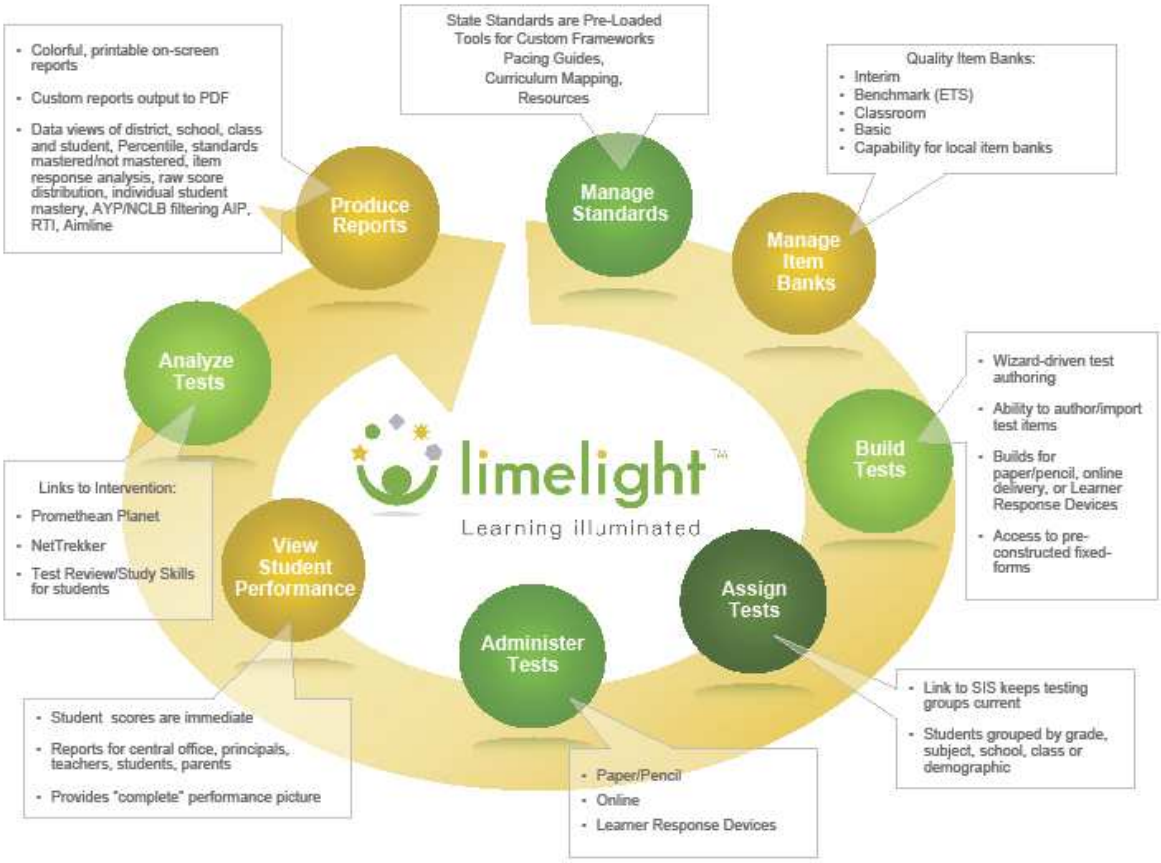
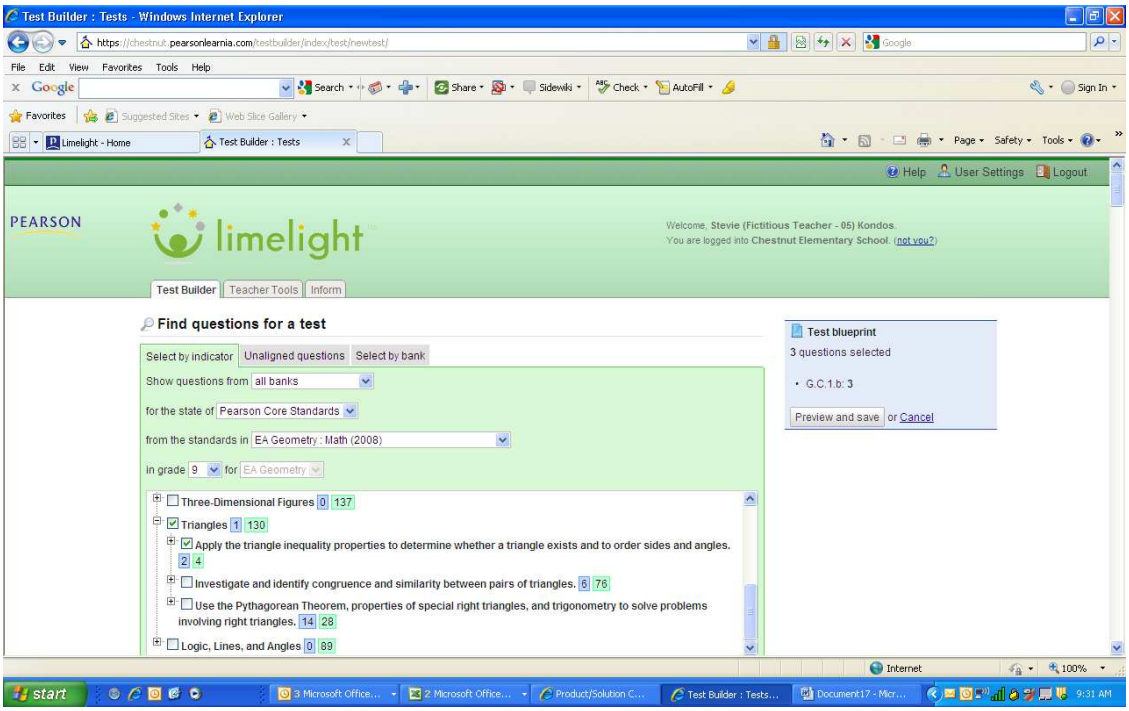


Figure 6 begins with the need to manage content standards and curricula. Arguably, all learning comes from standards and, as such, no small part of any learning system will require aspects of managing such information. Maintaining item pools and learning assets is another example of a component or attribute of a learning system that is required. If the learning system is an item bank, task or performance module or learning asset agnostic then it has the flexibility to be useful and helpful to educators of all areas. Similarly, flexibility in how educators combine assessment tasks and learning assets will facilitate the use of assessment for learning as oppose to using it as simple summary measures. Regardless of how the learning system will be used, assignment of tasks and managing administration will be a fundamental requirement of the learning system and will need to take into account a wide range of user facility with technology. Arguably, the information management aspect of the learning system is the most important. Being able to understand, present, analyze and take action is the empowering feature of a comprehensive learning system that may ultimately transform education.

Specifically, the content management aspect of the learning system needs to allow educators with easy to use and intuitive ways to review, select and manage learning and assessment assets is needed. Figure 7 displays one example of such a management system.

Figure 7. Limelight Assessment Task Management Example



Similarly, the ability to quickly and easily add learning assets to the learning management system is required. Figure 8 shows and example of how this might be done.

Figure 8. Limelight Assessment Construction Management Example.

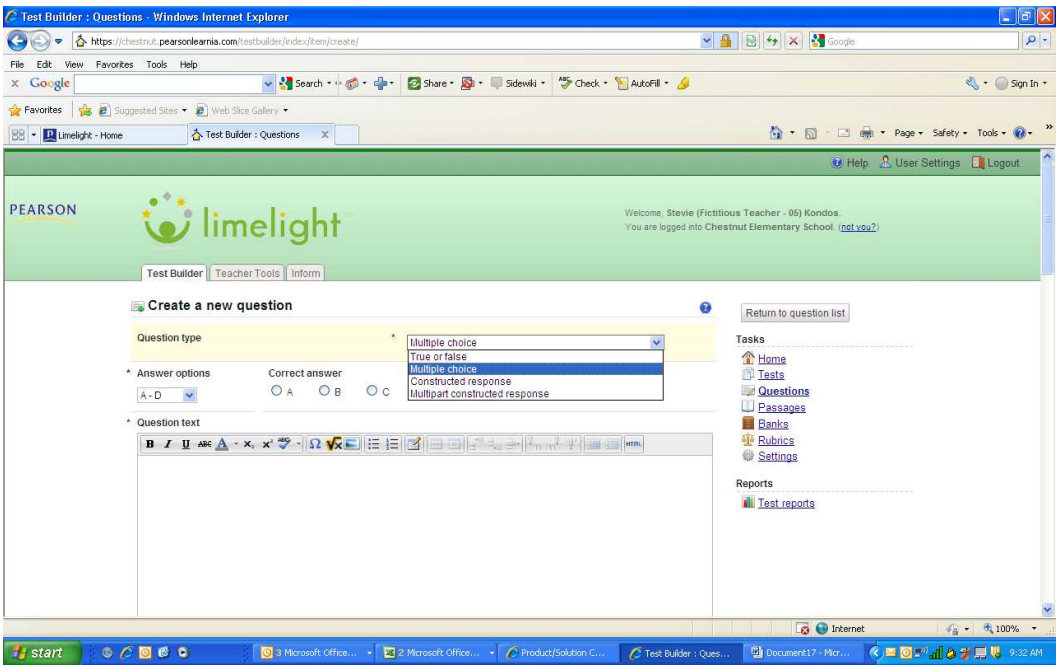


Figure 9 shows how one rendering of what an assessment asset looks like as presented from the learning management system. A robust and flexible system will accommodate all types of learning assets as well as a range of media types desired for use in assessment, instruction or learning as well as interventions.

Figure 9. Limelight Example Assessment Asset.

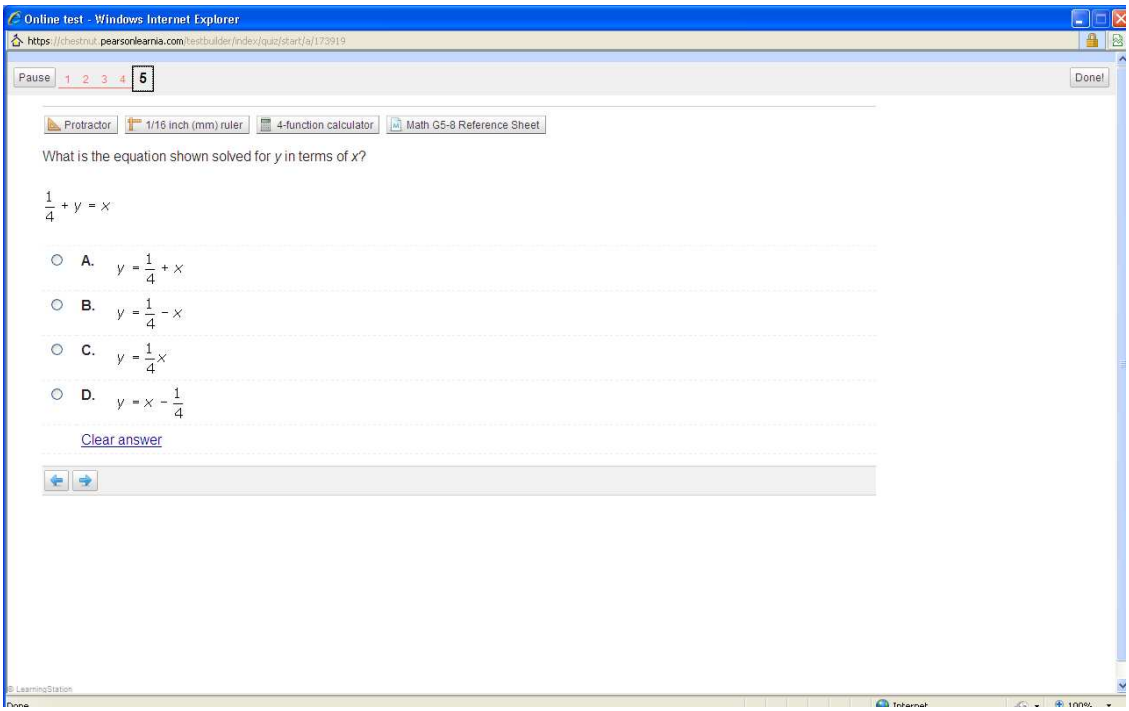
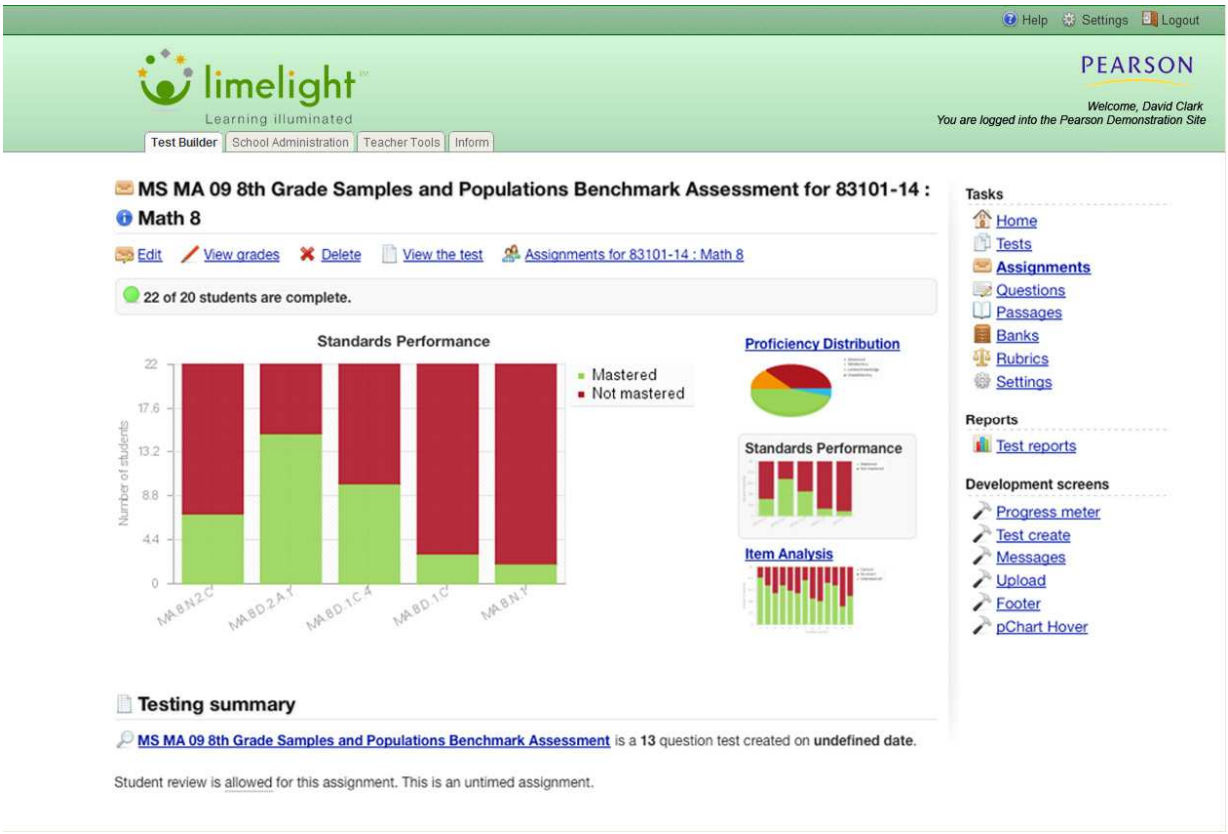


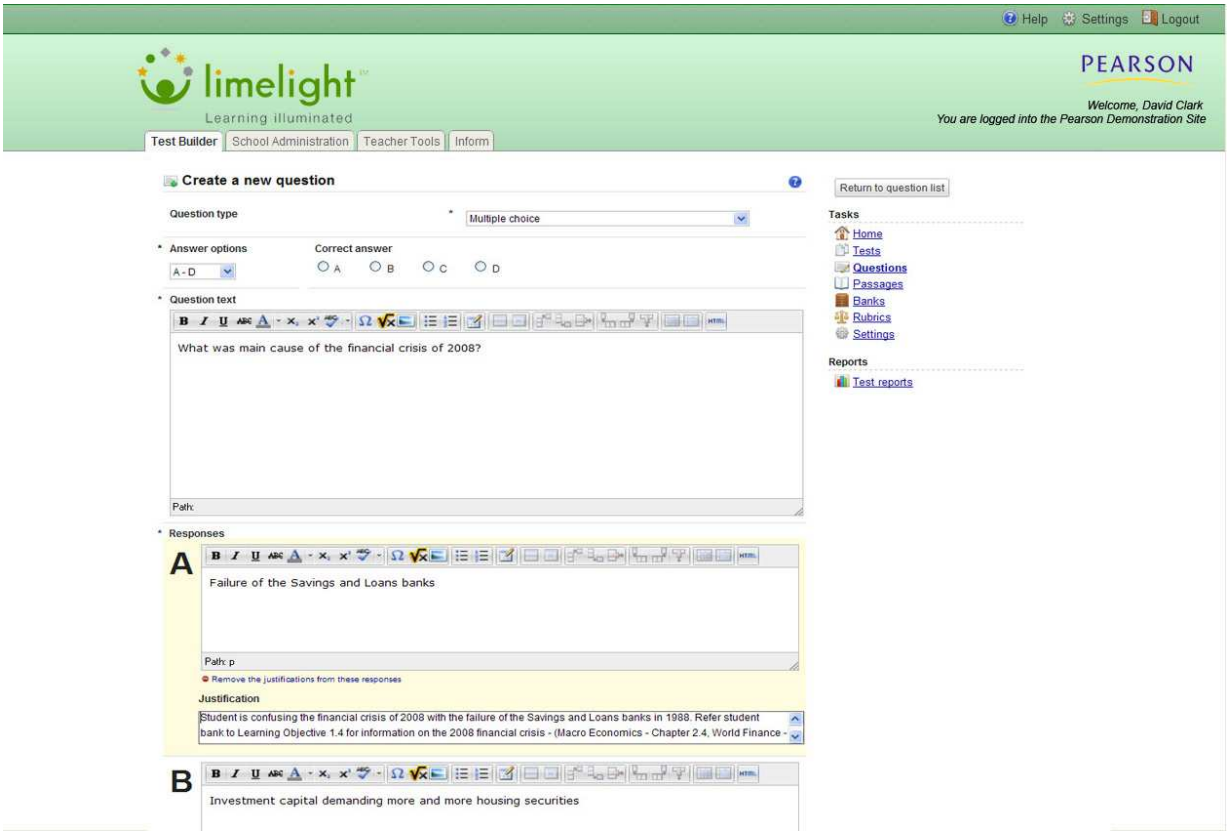
Figure 10 shows one aspect of a flexible series of reports summarizing learning information coming from the learning system

Figure 10. Limelight Example Proficiency Report



One additional and key aspect of the use of such a management system must be improved learning. As such, learning assets, instructional interventions and feedback to students and teachers is required. Figure 11 shows an example of one such type of feedback.

Figure 11. Limelight Example Question Rationale Report.



This section highlighted some guiding principles required for an integrated learning system and then provided examples of how one such system has visualized such a solution. Many such solutions exist and will continue to evolve. All follow a common theme, however, namely that only assessment integrated with learning will likely lead to a change in how we conceptualize the future of learning and with it the future of assessment.