LEARNING
AND
ASSESSMENT FOR LEARNING

IDEAS, THINKING AND DIALOGUE FOR GROWTH AND ATTAINMENT

SCOTTISH EXPERIENCE

Paper for Conference: "Evaluating Schools, Assessing Pupils"
Tartu, December 2 -3, 2010

Louise Hayward and Ernest Spencer
University of Glasgow

Introduction

The major role of assessment in supporting - or in interfering with - learning has been recognised in Scotland for more than thirty years. This paper and our presentation to the conference draws on the story of assessment in Scotland over that period, including the original thinking that informed developments in the 1990s and how our understanding about what matters in assessment to enhance learning has developed. Lest there be any misunderstanding, Scotland does not have all the answers and we continue to face challenges currently as the new Curriculum for Excellence is implemented and teachers begin to address the assessment professional development needs that it entails. However, we and many others in Scotland believe that we will only make progress towards the world as we would wish it to be by opening up the challenges and engaging *all* teachers in addressing them. We hope that today is the beginning of a longer conversation in which, as small countries in Europe, we can learn from one another and work together to build a better future for all of our young people.

1. ASSESSMENT FOR LEARNING: WHY IT SHOULD WORK

Learning: Becoming a Person

An important question is: "Assessment for learning what?"
What kinds of learning are we hoping to use assessment to promote?

One useful approach to exploring current thinking, debate and policy on education and pedagogy might be to consider where we stand in relation to three central ideas about the purposes of education. (This categorisation can be found in Spencer, 2006, and, with detailed references to the research/theory underpinning each position, in Spencer, 1983.)

Education is concerned with initiation into knowledge, modes of thought, social mores and skills which are inherently valuable, valued by society or useful for future employment. Logic, order, cognitive development, "knowing that something is the case" are important. (Most of those holding this "initiation" view would also accept that motivation and active involvement on the learner's part are also important).

Education is for developing reflective thinking, hypothesising, problem solving, personal meaning, in the context of the learner's broad culture. These abilities, and indeed, all skills, develop in collaborative activity with other people - teachers and other learners. They become internalised, personal, through imitation, "apprenticeship" and purposeful practice of them.

Education is the process of "becoming a person", which continues lifelong and is not confined to experiences in educational institutions but includes all experience of the society in which people live. It involves conscious pursuit of the goal of becoming the kind of person one wants to be, with the kind of knowledge and skills one wishes to have. Curiosity and self-confidence, important components of the skill of learning how to learn, lead to a sense of continuous change in oneself. One has a continuous potential to engage with new ideas, contexts and groups of people and to become a

confident and fully accepted member of one's community or of several communities - and thus meet the emotional and social development need for a sense of belonging.

In most educational systems, aspects of all three of these positions can be discerned in policy and practice. The extent to which each receives particular emphasis depends on judgements about the differing values they reflect. Our own personal stance is quite widely shared in Scottish education and is reflected in much of the national guidance for the new Scottish Curriculum for Excellence, which aims to enable all young people to become successful learners, confident individuals, responsible citizens and effective contributors. It is that enabling young people to "Become Persons" should be the overriding aim of education, but that to achieve this they need intellectual sustenance and skills in thinking and collaborative working.

A key source of the ideas underpinning the idea of Becoming a Person is the work of the psychotherapist, Carl Rogers (1969). Similar ideas have more recently been advocated and developed by Carol Dweck (1999). Self-determination is a key factor in learning of the highest quality and in relation to emotional and social influences. This kind of education would develop belief that you can, if you desire it, escape from, or transform into what you want, family and social influences and constraining emotional reactions, such as anxiety, shyness and lack of self-confidence. An important implication of Rogers' and Dweck's theme is that one role of the teacher is that of counsellor in the sense that therapists use. The job, well done, involves helping pupils to perceive what they can do and be. There are various ways of achieving this, including offering them and sometimes even pushing them into potentially extending and valuable experiences. The job also involves acceptance of the pupil on his or her terms in a relationship based on equality as persons exchanging ideas.

Intellectual sustenance means experience of significant ideas, grappling with real intellectual challenges which matter in the subject areas studied, presented for analysis, critique, comparison and incorporation in each pupil's own fashion in her or his own set of values and constructs. Expectation of ability to understand and think about significant ideas should be high, though there will often be need for the teacher to present them in language or other representation which is accessible to the pupils. Nobody, at any level of abilities, should be asked to think about or work towards trivialities.

Thinking skills are as important as the ideas. A view exists that the only criterion required for judging the quality of teaching is "Does it make them think?" Really effective teaching stimulates an excitement of mind, a desire to get a grip, work things out for yourself, find out more, experience more, try to solve problems. The cause of this excitement is complex, but it probably includes the nature of the ideas and tasks presented and the teacher's relationship with the pupils, in particular his or her expectation of independent thought on their part and his or her skill in supporting them in the process. There are two major types of intellectual activity under the heading 'Thinking', both of which would be happening continuously in really good school work. One is the understanding of new ideas and information in relation to what one already knows and in one's "own words" and reflection on what one is learning. A key principle is that the kind of thinking one does to clarify understanding is the same kind of thinking which is necessary to speak or write clearly about a topic. Requirement to write or talk is therefore often a valuable means of stimulating

thinking for understanding. However, clarification and communication of understanding can be achieved in other media, too: art, drama, music, video production…The principle is the same: production of an ordered statement requires hard thinking about the topic, as well as skill in and thinking about the means of communicating it. The other crucial type of thinking is the application of doubt, the questioning of assumptions and assertions, others' and one's own. The uncertainty and permanent testability of knowledge and value positions should be a key concept in young people's minds. There should be, as a matter of course, frequent encouragement in discussion and in the general work of any course, to question assumptions and test assertions and hypotheses, one's own and those discerned in, e.g., reading, radio and television programmes, political, social and moral arguments. This principle also implies that pupils should frequently be dealing with problems (of various types, according to the subject area) which require them to make decisions for which the correct basis is not obvious – decisions which involve grappling with and resolving one's own uncertainty about, e.g., a practical or social problem, a commitment to a view or a policy, an interpretation of a text…

Assessment for learning and Becoming a Person

In Developing the Theory of Formative Assessment, Black and Wiliam (2009) identify general characteristics of formative assessment – assessment for learning - which were in fact features of the Scottish Assessment is for Learning Programme, developed from 2000 in education authorities and schools.

They define formative assessment, drawing on earlier definitions (Black and Wiliam, 1998b, and the Assessment Reform Group (ARG), 2002), as the process in which teachers, learners or their peers elicit, interpret and use evidence about pupil achievement to make decisions about the next steps in learning "that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited" (p9). They list the five main types of activity developed with and by teachers in normal classroom work in early work on formative assessment:

Sharing success criteria with learners
Classroom questioning
Comment-only marking
Peer- and self-assessment
Formative use of summative tests.

These activities were promoted because of evidence (from Wiliam 2000; Black et al. 2003; Wiliam 2007b) of their potential effective contributions to three key processes in learning and teaching (identified by Wiliam and Thompson 2007, following Ramaprasad 1983)

Establishing where the learners are in their learning
Establishing where they are going
Establishing what needs to be done to get them there.

Black and Wiliam go on to conceptualise formative assessment as consisting of five key strategies (broader than the original five kinds of activity, which they describe as ways of initiating action, but which do not adequately reflect all aspects of formative assessment):

1. Clarifying and sharing learning intentions and criteria for success;

2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. Providing feedback that moves learners forward;
4. Activating students as instructional resources for one another; and
5. Activating students as the owners of their own learning.

The paper locates formative interactions within more comprehensive theories of pedagogy, which Perrenoud (1998) had argued "constitute the real systems of thought and action, in which feedback is only one element" (p. 86). The argument here clarifies the contribution of formative assessment or assessment for learning to the rich educational process we described above as Becoming a Person. If we are committed to a constructivist learning approach derived from the work of Vygotsky, we should think of formative assessment as "assessment for development" rather than just "for learning". Vygotsky distinguished between learning and development, describing the latter as involving changes in psychological functions available to the learner, not just acquisition of new knowledge or new mental capabilities. "The zone of proximal development (ZPD) is not, therefore, just a way of describing what a student can do with support, which might be simply learning, it is a description of the maturing psychological functions rather than those that already exist. A focus in instruction on the maturing psychological functions is most likely to produce a transition to the next developmental level and "good learning" is that which supports the acquisition of new psychological functions" (Black and Wiliam 2009, p. 19). Formative assessment contributes significantly to such development, partly because the quality of interactive feedback and reflection are critical features in learning activity. It develops the orientations, abilities and confidence characteristic of the independent and collaborative learning which are entailed in Becoming a Person.

Following Vygotsky's (1978) principle that ideas appear first in the external "social" plane, then become internalised by the individual, dialogue with others is a key means by which pupils learn.

While recognising (as do Black and Wiliam) that observation and interaction can provide only indirect evidence of a learner's internal affective state and cognitive processes, we can be clear that dialogue (pupil-teacher and pupil-pupil) is also a crucial means of challenging learners to reflect on their own thinking and to make unconscious learning processes overt, so that they can be considered, discussed and improved. The whole set of assessment for learning/formative assessment processes essentially comprise a sequence of three recurring activities: stimulating learners to think about the topic; finding out, often through dialogue, what and how they are thinking; on the basis of this evidence, identifying with them next steps for more effective thinking and fuller, more certain grasp of what is being learned. Particular "assessment for learning techniques", such as "traffic lights", "wait time", "fat (ie, open) questions" or WALT (We Are Learning Today …), may be means of engaging pupils in aspects of the necessary thinking, but they are not of themselves assessment for learning. Indeed they can be counter-productive if they are perceived as just teaching "tips" or as techniques that inevitably improve learning. What is needed is really effective teaching of "intellectual sustenance" incorporating any kinds of activity that cause thinking and reflection. Black and Wiliam (2009) show that successful learning programmes, such as Cognitive Acceleration (Shayer and Adey 2002; Adey 2005) and Dynamic Assessment (Poehner and Lantolf 2005) encourage

cognitive growth by challenging thinking, creating cognitive conflict. rather than giving answers, use dialogue to serve the social construction of knowledge and promote learners' reflection on their own learning. Assessment for learning practices are thus an essential feature of these programmes, which extend and develop the reasoning resources and the confidence and commitment that a learner can bring to any future task. Black and Wiliam (2009) argue that any teacher using formative and interactive dialogue for normal subject teaching and feedback that encourages self-regulated learning (targeting one's own cognitions, affects and action, as proposed by Boekaerts et al. 2005, p.150) is engaged in a subject-specific form of thinking skills programme. She/he is helping young people to Become Persons.

2. ASSESSMENT FOR LEARNING: SOME SCOTTISH EXAMPLES

In 2002 the Scottish Government initiated a new attempt to improve the quality of assessment in Scotland, the Assessment is for Learning programme (AifL). The programme was concerned to create a coherent assessment system, including assessment for formative and summative purposes and for purposes of wider public accountability. Schools across Scotland were invited to take part in one of 10 projects relating to these themes. One of the projects was concerned specifically with formative assessment. The background to AifL is explained in some detail in Hutchinson and Hayward, 2005 and in Hayward and Spencer, 2010. This latter article is drawn on in the final section of this paper to identify important factors to be considered in developing and growing a programme to develop assessment for learning. We offer here some examples of assessment for learning in action in Scottish schools, mostly developed as part of the government programme. In their different contexts and with different emphases, they demonstrate how in practice some teachers provided intellectual challenge; created tasks/activities that stimulated independent and collaborative thinking; used dialogue as a means of learning and a means of finding out about learners' thinking and feelings; and interacted with pupils as persons, finding ways of building confidence and motivation and extending reasoning abilities.

2.1. Lynne (aged 7) (Adapted from Hayward and Spencer 1998.)

Lynne's teacher (and her teachers in the school she had just arrived from) and the learning support coordinator in the school were very concerned about lack of progress with reading. She was working on a different reading scheme from most of her classmates, since a miscue analysis had shown her reading book to be too difficult.

As they observed her work in class both teachers noted Lynne's enthusiasm for tasks which were well within her capabilities: 'She was keen to show us what she had done. She needed constant praise. Her work was very neat and tidy and she took a lot of time over it. Lynne was taking a real interest in her reading books.' However, discussion with Lynne in which the teachers aimed to find out what she really thought about reading provided what they considered "probably the most revealing of all the evidence gathered so far". Lynne's concept of reading was very interesting. She enjoyed reading books from reading schemes and did not value library books. They were not as important as "reading books": "Mrs Smith has the reading books. The other teacher gives us library books. They're not reading books!" The appearance of the book mattered: 'I like books with nice pictures. I don't like books that've been

scribbled on.' Lynne did not find reading easy and was very much aware that she was not as good at it as some other children in her class. For this reason she preferred to read quietly to herself so that no-one could hear her mistakes: 'I like to read it into myself and not out loud because people will say — "Ah, you can't read".' Even though she was having difficulties, Lynne thought that it was important that she learned to read. She had a very stereotypical vision of her future. She did realise that you would have to be able to read if you wanted a job, for example, as a teacher. However, her main aim was to become a mother: 'If you have a new born baby you can read to it and not make mistakes. I want to be a mother when I grow up.' She did not really appreciate that reading could be enjoyable or that you could learn from it beyond 'You wouldn't know what was on the telly if you couldn't read'.

As they considered the evidence before them, Lynne's teachers recognized that she was coping well with her present reading scheme books but there were a number of possible areas for development, if Lynne's reading abilities were to be enhanced. Initially they decided to focus on attitude and motivation, concentrating on strategies to build confidence as a reader and to encourage a wider enjoyment of reading. The steps they took included creating opportunities for Lynne to read to younger children, asking her to list all the things to do with reading at which she is "good" as a basis for a reading record of achievement, introducing books beyond the reading scheme and encouraging her enjoyment of them by asking her to recommend books for younger children.

Lynne's teacher also commented: "I realised that reading in my classroom had been too scheme-based. Reading for Lynne meant books from the reading scheme. There are things about my teaching programme that I will look at again."

2.2.History with 12-13 year olds (Adapted from Sliwka and Spencer, 2005)

This First Year secondary work featured self-and peer-evaluation of group presentations on researched topics and later also of individual written essays on the topics. The research activities, using carefully identified library and internet resources, and the preparation for presentations constituted about 50% of the classwork; pupils received direct teaching of subject content and skills in the other 50%, in the classroom. The staff justified this approach on the grounds that the process of learning was as valuable as the subject content and that practical application of history skills in the research/presentations deepened understanding of the ideas and evidence and developed not only interest in history but personal confidence and collaborative learning skills.

The research topics were controversial and therefore stimulating: eg, "The Romans did not really create a civilised society in Britain. What is the evidence for and against this statement?"; "William Wallace (a great Scottish hero) deserved to be executed by the English? How far do you agree?" Each group was expected to include in the presentation an introduction and background information, evidence to support the case for the argument, evidence against it and a reasoned conclusion. Before undertaking the research work the class spent a good deal of time discussing and agreeing with the teacher the criteria for an effective presentation (and also for the essays they would write later). They agreed on three possible levels of success: a very successful argument with full evidential support; a capable but not complete argument, with

some appropriate evidential support; and an argument that needed boosting in various ways. As each group made its presentation, the rest of the class made individual evaluations of its argument and evidence and then took part in group discussion to reach consensus on the criteria which had been met.

The teacher managed the feedback session after each presentation very effectively. He began with an open-ended class discussion, asking the class to consider the strong and weak points of the presentation and emphasising the need to provide evidence for the evaluation. This strategy kept open the possibility that pupils might come up with insightful comment on their colleagues' work without the help of the relatively pre-determined 3-level criteria statements (which the teacher nevertheless considered important as "scaffolding" for pupils who were not yet used to making evaluative comments on one another's work). He encouraged the class to agree or take issue with statements made by individual pupils in this open discussion. After it the pupils individually and then in groups confirmed or revised the initial evaluations they had made as they listened to the presentation.

One significant aspect of the arrangements was designed to ensure that all the pupils benefited from working through the challenges presented by the research and presentation tasks. The groups were carefully selected as mixed-ability and there was a requirement for each individual member to undertake some of the research work and contribute in the presentation. Since the pupils knew their presentation was going to be judged by the whole class, much supportive teamwork developed in the groups, with pupils helping others to access and understand their particular aspects of the research topic and ensuring they could contribute effectively to the presentation.

In the light of a frequently heard concern about the "dangers" of giving time to this kind of co-operative learning and assessment and so failing to prepare pupils properly for the type of examinations they will eventually take for qualifications, it is noteworthy that the teacher who devised and ran this approach to history teaching was the SQA Principal Examiner in the subject at Higher Grade. He clearly did not take the view that pupils needed only narrow "exam preparation" to achieve good success in the examinations that he was setting. It is also noteworthy that their early secondary history experiences led many of his pupils to choose the subject for continued study as they moved into the senior stages of the school.

2.3. Higher Mathematics (17 year olds) (Adapted from Sliwka and Spencer, 2005).

Much of the work of the class was based on the teacher's recognition that individual pupils vary in the ways they use mathematical knowledge to solve problems. It was designed to challenge the pupils to make their own thinking explicit, to explain it to others and to gain from what other pupils told them about their approaches. In groups of four they were expected to exchange ideas and discuss how to tackle mathematical problems of the kind they would eventually meet in the examination. "We argue in our group about the right way to do things. We use different methods, we compare the way we did it. If someone gets it wrong and others get it right, they explain how they did it to that person." The teacher came into play in these activities only when a group did not work out for themselves how to move ahead or there was controversy about the solution to a problem. His approach was often to ask the pupils questions they had not considered themselves in order to stimulate further thought in the group, though

he explained the relevant mathematics where he realised that pupils had misconceptions. He regarded the group discussion approach to addressing mathematical problems as extremely valuable in exposing both pupils' often differing ways of understanding, from which others could gain, and their misconceptions, which enabled him to focus explanations to address needs.

2.4. Psychology (17-18 year olds) (Adapted from Sliwka and Spencer, 2005).

In this work a deliberate use of co-operative learning freed the teacher to spend time with and provide scaffolding/advice for groups and individual pupils with different learning needs, based on her assessment of strengths and needs. In a study of anorexia nervosa, the pupils used a newspaper article, a case study and printed information on psychological theories explaining abnormal behaviour. After an introductory orientation, the 20 pupils worked in groups of four on a clearly defined task with a deadline. They would present the outcome of their task to the rest of the class later. The teacher interacted with each group, checking understanding of the texts and inviting comment on the relevance of the theories to the issue. She listened with great attention to what each group and individual pupils said, encouraged them to think beyond what the texts told them and added detailed expert knowledge to enhance understanding of key points. The pupils visibly enjoyed the professional, even academic, atmosphere of the work and respected the teacher as an expert responding to their interests and ideas and helping them develop their own knowledge and expertise.

2.5. Preparation for examinations at 16 (Intermediate) and 17 (Higher) (Adapted from Hayward et al 2009).

During academic year 2007-08, teachers in several Highland Council secondary schools participated in a project supported jointly by the Council (as part of its "Highland Journey" towards more effective learning, teaching and assessment) and the Scottish Qualifications Authority (SQA) to explore the extent to which they could enable pupils to make formative use of the formal criteria for Intermediate or Standard Grade examinations at 16 and Higher examinations at 17. SQA publishes Grade-related criteria (GRC) for all subjects at all levels of the examination system. The following example for the Critical Essay task in English (Literature) at Intermediate 2 (age 16) shows the kind of material that is available.

INTERMEDIATE 2 EXTERNAL ASSESSMENT (EXAMINATION)

Critical Essay Performance Criteria (Pass at Grade C)	**Indicators of Excellence (Qualities leading to Grade A)**
Understanding As appropriate to task, the response demonstrates understanding of key elements, central concerns and significant details of the text(s).	**Understanding** The response reveals some insight into key elements and central concerns of the text(s). Clear explanation is given of significant detail.
Analysis The response explains in some detail ways in which aspects of structure/style/ language contribute to	**Analysis** The response reveals some insight into the use of literary/linguistic technique

meaning/effect/impact.	
Evaluation The response reveals engagement with the text(s) or aspects of the text(s) and stated or implied evaluation of effectiveness, substantiated by some relevant evidence from the text(s).	**Evaluation** Evaluation is valid and appropriate and reveals clear engagement with the text(s). Critical stance is clearly established and fully supported by appropriate use of textual evidence.
Expression Structure, style and language, including use of some appropriate critical terminology, are deployed to communicate meaning clearly and develop a line of thought which is generally relevant to purpose; spelling, grammar and punctuation are sufficiently accurate.	**Expression** Expression, including use of critical terminology, is generally effective in establishing a clear and consistently relevant line of thought.

In addition, examiners' marking schemes for previous examination questions are publicly available.

Hayward et al, 2010, describe the ways in which teachers sought to use the criteria formatively in a range of subjects, including English, mathematics, sciences, social subjects and modern languages. Examples include:

- Getting pupils to think out collaboratively how they would approach particular past examination mathematics questions, calling on mathematics knowledge covered in the course, working out the nature of the problem in the question and using the GRC to identify key requirements of the answer if it is to achieve a high grade.
- Engaging pupils in "jigsaw" co-operative learning, working in a group to ensure they had a full grasp of topics they and the teacher agreed were weak for them, then reporting to the rest of the class.
- Enabling pupils to access privately their own and their peers' English speaking assessment tasks on the teacher's blog, so that they could themselves assess them using the GRC without public embarrassment.
- Operating a "reading market" in the English classroom: different groups of pupils were responsible for producing and explaining to others who came to their "stall" how to produce very good answers to particular literature questions, taking account of the GRC.
- Having pupil groups produce their own marking schemes for summative assessment tasks they had taken in class; this involved much learning discussion and led to an agreed class scheme which they then used to assess their own and others' work.
- Using brainstorming approaches and pupils using GRC and the examiners' published marking scheme in marking essays written by pupils in previous years to develop much discussion about the qualities of a very good discursive essay in French. (The teacher reported that 9 times out of ten the pupils became as accurate in marking others' essays as she was herself.)

3 DOES IT WORK?

3.1 Improving the quality of learning

By 2005, AifL in Scotland had developed a positive public profile. It was described by the then Education Minister as a quiet revolution in Scottish education. The original formative assessment project within the broader AifL Programme was evaluated by Hallam et al (2004). They reported evidence of positive impact on pupils' motivation, levels of engagement and confidence, particularly among pupils whose current levels of attainment were not high. They also reported significant developments in teachers' practices: "It would not be an overstatement to say that the project has focused teachers' attention on how their pupils learn and how they can support this process, rather than what or how much they have learned." (Hallam et al. 2004, 133). Hayward, Simpson and Spencer (2005) undertook a further study specifically aimed at clarifying the characteristics of the formative assessment part of the Programme that had motivated and engaged teachers and had led to the particular success of that aspect of the government development. (Hayward and Spencer, 2010, offers an analysis and commentary on the findings and implications of that study.)

Teachers and education authority staff who were interviewed were extremely positive about the benefits of the assessment for learning (formative assessment) developments in which they had taken part. The following quotation from a secondary teacher exemplifies the impact on teachers' ways of working:
"I don't think I'll ever go back to being the kind of teacher I was before. And it just becomes part of you. Part of your teaching manner. And it didn't happen overnight. It happened gradually, as you read material, as you saw what other people were doing, as you were listening to others, as you tried things. And I'd be prepared to say that some things worked and some things didn't work. And maybe there were reasons for that. But the whole ethos of the system, I still think it's of huge value and that's the way we should be teaching. So I don't want it to die. I mean, at my stage in life you could say, 'Right, okay, four or five years to go, too late to change'. But never. These pupils are only in your room once."

Another secondary teacher, close to retirement, said with chagrin that he had only just learned to teach properly!

Teachers highlighted the significance of moving from "teaching" to "learning". A primary head teacher summed up the development as "It's just the shift of emphasis from the teaching to the learning, you know." A secondary teacher said: "It made me think how the pupil was thinking and get my mind into their mind and think….right, where is this pupil at? Why are they not understanding, while to me it's crystal clear? How can I help them progress? How can I encourage them not to give up? How can I imagine I'm sitting in their seat listening to this guy?"

Impact on self-esteem, engagement, and attainment was frequently mentioned in the interviews, for example by this primary teacher: "Confidence grows. They would never go back. There's a wee boy working out there who had learning difficulties….he needed support in language. He's on the star pupil board. You wouldn't believe what he can write. And that's all with his new self-esteem. That's all starting out with just a few words, saying, 'That's brilliant.' "

3.2 Attainment

The value of the changes brought about by assessment for learning in pupils' motivation, attitudes, collaborative learning and independent thinking and in teachers' professional development as educators is clear. Those actually implementing the formative assessment activities tended to regard their benefits as self-evident. However, such benefits are sometimes regarded as "soft", not comparable to the "hard" evidence represented by test and examination results. This might be thought an odd position to take if the argument presented at the start of this paper about the benefits of assessment for learning to development as a thinker and learner are accepted – it was evidently not the view taken by the history examiner whose own teaching developed those "soft" skills so effectively, with a view to enabling his pupils to achieve high grades in later examinations, as well as developing their learning and thinking abilities and their confidence.

In any case, the argument that assessment for learning leads to more success in tests and examinations, where these are important summative assessment tools, does not depend solely on the argument that it *should* do so, given its nature. There is hard evidence that it does do so. Black and Wiliam (1998b) reviewed over 250 studies linking assessment and learning and obtained clear and incontrovertible evidence that initiatives to enhance effectiveness of the way assessment is used in the classroom to promote learning can raise pupil achievement. In England, the scale of the effect would be the equivalent for an individual of between one and two grades in a General Certificate of Secondary Education examination at age 16. Black and Wiliam estimated that attention to formative assessment throughout the country would have raised England's position in the Third International Mathematics and Science Study from the middle of the 41 countries involved to one of the top five. They also found evidence that the gain was likely to be even more substantial for lower-achieving pupils.

Education Authority co-ordinators involved in the Hayward and Spencer (2005, 2010) study of the success of formative assessment in the Scottish Government initiative also identified an impact on attainment as measured in national tests and other summative assessments. One said "Already attainment is improving. There has been clearly observable progress in Level A achievement in a P1 task after formative assessment had been tried out with them …this was also true of a secondary school geography class. Formative assessment encourages talk… relearning and improvement. It deepens learning considerably, achieves learning intentions and improves the quality of discourse."

Another's view was: "We have lots of evidence from videos of classroom work, observations, discussion. ….we have a wider range of children engaged and classwork has improved even in one of our high attaining schools. AifL and Building Bridges (a primary-secondary liaison project focusing on literacy) have led to improved 5 - 14 test results at Primary School X and Secondary school Y".

Interesting evidence emerged also from the study by Hayward et al (2009) of Highland Council schools' use of examination criteria in formative ways. The

changes in pedagogy varied slightly from teacher to teacher, but tended to be in a similar direction — towards more group work, pupil-pupil dialogue, peer -and self-evaluation and creative thinking. Teachers cited reflection, peer- and self-assessment, pupil autonomy and understanding as the main changes in pupils' learning. The key word was dialogue: pupil-pupil, pupil-teacher and teacher-teacher. The teachers were most animated when they spoke about this aspect of their pedagogy. Indeed, they seemed to value this more than examination success, although they never deviated from their duty to get their pupils through the exam successfully.

The strong tendency was for the teachers to focus on learning and pupil engagement. One argued: "My exam results have improved over the last few years … but more than the results, it's just being in the classroom and seeing the engagement of the pupils."
Another said: "So I would say that's the key; the engagement with both pupils and colleagues."

The final question put to all the teachers as the data gathering for the study came to an end in June 2008 invited them to speculate on how their pupils might perform in their examinations. They were all reticent and unwilling to be over-optimistic, though they emphasised the benefits they considered the pupils had gained as learners.

The teachers came together again with the researchers in September 2008, after the publication of examination results. The first question was, "How did the pupils do?" The response was almost unanimous — the results had been excellent.

However, what was also important were the comments from those teachers who said that the results had been more or less as expected. They felt that the pupils had been more self-aware and had come after the examination to discuss the paper and talk about how they had gone about their answers, whereas, in the past, they would have gone straight home. Others had come to see the teacher at the beginning of the current year academic year, as they began S6, talking about the strategies they would have to work on to do better next time.

Key issues identified in the Highland study were common to teachers from all areas of the curriculum. They argued that self-evaluation, critical and creative thinking and reflection were not perceived to be rewarded by the current Scottish examination system, but that they should be.

4. MAKING IT WORK

The 2002 Assessment is for Learning developmental strategy was designed in awareness of the historical reality of failed implementation during the 1990s of a national assessment policy based on essentially the same principles as AifL itself - a coherent system strongly promoting formative assessment ("assessment as part of teaching") and summative assessment based on teachers' professional judgement of a large body of classroom evidence, with national testing used simply as a means of confirming teachers' own assessments.

Hayward and Spencer 2010 refer to a substantial body of evidence that plans for curriculum or pedagogical innovation often lead to little change in practice

(Cuban1994; Swann and Brown 1997; Barnes et al. 2000; Olson 2002). Some initiatives have very little impact at all. Many are successful in their early stages but fail later, although reasons for failure appear rarely to have been analysed in any systematic way. Understanding the relationship between research, policy and practice in assessment in Scotland has proven to be a complex task (Hayward and Hedge 2005). Previous articles analysed why the earlier attempts to change assessment practice in Scotland had failed, even where efforts had been made to ensure that assessment policy was well informed by evidence from research (Hutchinson and Hayward 2005). In common with many well-intentioned educational initiatives assessment changes in Scotland in the 1990s had been what Gardner et al. (2010) describe as under-designed. Planning had focused on producing research-informed assessment *policy*. Although some attention was paid to engaging policy-makers and practitioners, the exercise was seen largely as a process of disseminating policy across Scotland in ways that would inform and enthuse teachers. The impact of the wider policy context of performativity was underestimated as was the effect of multiple, parallel local and national policy initiatives (Hayward 2007). In essence, the relationship between research, policy and practice was oversimplified. As a consequence a set of factors that get in the way of effective assessment for learning remained prominent -

- a tendency for teachers to assess quantity of work and presentation rather than the quality of learning;
- too much attention given to marking and grading, much of it tending to lower the self-esteem of pupils, rather than to providing advice for improvement – in reality the national tests became almost the sole means of assessment;
- a strong emphasis on comparing pupils with each other, which demoralises the less successful learners;
- teachers' feedback to pupils often served social and managerial purposes rather than helping them to learn more effectively;
- teachers did not know enough about their pupils' learning needs.

Accordingly, the 2002 AifL strategy aimed to tackle issues of impact and sustainability, drawing on research evidence not only about assessment but also about the processes of individual and collective change, in particular the work of Fullan (1993) and Senge and Scharmer (2001). The intended AifL model is described in Hayward et al. (2004, 399). Three key features were that:

- the initiative should focus on real issues important for the communities which would participate in it;
- the programme should be inclusive, involving all relevant communities in its development and thus seeking to address issues which might inhibit valuable change, such as competing policy demands;
- AifL should recognise the complexity of the change process and should not seek simplistic models that would be unlikely to achieve meaningful change, such as informing teachers of research findings and expecting practice to change as a consequence of that act.

The study of the successful implementation of the formative assessment aspect of the AifL Programme by Hayward et al. 2005 and the commentary on its findings by Hayward and Spencer 2010 confirm that these are indeed key characteristics of effective change. These publications identify three crucial factors emerging from the investigation of successful implementation.

- **Educational integrity**
  Those participating recognised that the initiative manifestly focused on what matters for good learning, including focus on learning and learners' ways of working, more consistent checking for understanding and developing pupils' independence as learners.

- **Personal and professional integrity**
  Teachers recognised the development as central to their own professional concerns: personal conviction was a key factor – a belief that what they were doing really mattered, that they were not simply responding to someone else's priorities. A second key aspect of professionalism for teachers was the sense of being listened to, of having a significant role in constructing the programme, in deciding how to use formative assessment in their own classrooms, rather than being the passive recipients of policy directives and advice from others. This participative role gave teachers a sense that their professionalism was respected and crucial to the development. A third important factor was having opportunities to work through ideas and challenges with other teachers, in their own school and elsewhere.

- **Systemic integrity**
  The development was clearly, explicitly and consistently supported and maintained by all key players in the education system - government, local authorities, inspectors, school managers, university researchers …

Hayward and Spencer 2010 argue that the process of effective implementation of a significant educational change like assessment for learning is complex and, like the Teaching and Learning Research Programme publication by James et al (2007), they suggest a need to understand more deeply the nature of complexity and to find ways of responding to it, rather than seeking to simplify, to make manageable. Consistent with the contention by Gardner et al. (2010) that innovation needs to be designed for sustainable development from its outset, they propose that the design process itself should be collaborative, in order to build in different perspectives from different communities.

It is salutary to list some of the key factors in the complexity of successful change in pedagogy that emerged from the 2005 study by Hayward et al and are discussed in Hayward and Spencer 2010:

- educational integrity, what matters for learning, which is itself a complex idea, and actual improvement of learning occurring;
- ensuring depth of understanding about what really matters – and implications for the roles of teaching, research and policy communities;
- personal conviction on the part of teachers, researchers and policymakers, and their full professional participation in deciding action to take the development forward;
- openness, equality, sharing of issues, problems, solutions and professional expertise across all the communities involved;
- effective interaction and sharing, networking in teacher groups in school and the wider peer group
- consistent policy and advice from policy and research;

- deep understanding of learning and teaching principles and of the nature of participative learning, and an awareness of different community priorities;
- recognition that the change process is ultimately personal;
- full attention to all these important interacting factors and avoidance of over-simplified strategies.

Achieving real change is not easy but it is important to remember that it was achieved in this first phase of the AifL formative assessment project. We need to make sure that future major developments create their own informed routes through their own set of complexities. In Scotland, previous models of change were often referred to as 'pilot' and 'roll out' or 'cascade', on the assumption that once ideas had been developed by a few people, others could simply be informed or instructed how to make the initiative work. However, the things that matter in the process of real change need to be fully worked through by *all* participants, whether an individual or a school is in the first phase of the development or the last.

REFERENCES

ARG.2002. *Assessment for Learning: 10 Principles*.
Available on the Assessment Reform Group website:
www.assessment-reform-group.org.uk.

Barnes, M., Clarke, D. and Stephens, M. 2000. Assessment: The engine of systemic curricular reform? *Journal of Curriculum Studies* 32, no. 5: 623–50.

Black, P. and Wiliam, D. 1998a. Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5(1), 7–73.

Black, P. and Wiliam, D. 1998b. *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. 2002. *Working inside the black box: Assessment for learning in the classroom*. London: GL Assessment.

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. 2003. *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.

Black, P. and Wiliam, D. 2009. Developing the theory of formative assessment, *Educational Assessment, Evaluation and Accountability*, Volume 21, Number 1, 5-31.

Boekaerts, M., Maes, S. and Karoly, P. (2005). Self-regulation across domains of applied psychology: Is there an emerging consensus? *Applied Psychology*, 54(2), 149–154.

Cuban, L. 1994. *How teachers taught: Constancy and change in American classrooms 1890–1980*. New York: Teachers College Press.

Dweck, C. S. 1999. *Self-Theories: Their role in motivation, personality, and development.* Philadelphia, PA: The Psychology Press.

Fullan, M. 1993. *Change forces: Probing the depths of educational reform*. London: Falmer.

Gardner, J., Harlen, W., Hayward, L. and Stobart, G., with Montgomery, M. 2010. *Developing teacher assessment*. Maidenhead/New York: Open University Press/McGraw-Hill.

Hallam, S., Kirton, A., Pfeffers, J., Robertson, P. and Stobart, G. 2004. *Evaluation of Project 1 of the Assessment is for Learning development programme: Support for professional practice in formative assessment.* London: Institute of Education, University of London.

Hayward, L., Priestly, M., and Young, M.. 2004. Ruffling the calm of the ocean floor: Merging research, policy and practice in Scotland. *Oxford Review of Education* 30, no.3: 397–415.

Hayward, L. and Hedge, N. 2005. Travelling towards change in assessment. Assessment in Education: Principles, Policy & Practice, 12, no. 1: 55–76.

Hayward, L. 2007. Curriculum, pedagogies and assessment in Scotland: The quest for social justice. 'Ah kent yir faither'. *Assessment in Education: Principles, Policy & Practice*, 14, no. 2: 251–68.

Hayward, L., Simpson, M. and Spencer, E. 2005. *Assessment is for Learning: Exploring Programme Success*. Edinburgh: Scottish Executive Education Department. Available online at:
http://www.ltscotland.org.uk/resources/a/genericresource_tcm4354152.asp?strReferringChannel¼assess

Hayward, L., Boyd, B., MacBride, G. and Spencer, E., 2009. *Just making them think: a tension between teaching and assessment in the high stakes stages*. Glasgow. Scottish Qualifications Authority (Research Report 13.).
Available:
http://www.sqa.org.uk/sqa/files_ccc/Report%2013_Make_Them_Think.pdf

Hayward, L. and Spencer, E. 2010. The complexities of change: formative assessment in Scotland, *Curriculum Journal*, 21: 2, 161 — 177

Hutchinson, C., and Hayward, L. 2005. Assessment in Scotland: The journey so far. *The Curriculum Journal* 16, no. 2: 225–48.

Olson, J. 2002. Systemic change/teacher tradition: Legends of reform continue. *Journal of Curriculum Studies* 34, no. 2: 129–37.

Perrenoud, P. 1998. From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in Education: Principles, Policy and Practice,* 5(1), 85–102.

Poehner, M. E. and Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9(3), 233–265.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4–13.

Senge, P. and Scharmer, O. 2001. Community action research: Learning as a community of practitioners. In *Handbook of action research: Participative inquiry and practice*, Ed. Reason, P. and Bradbury, H. 238–49. London: Sage.

Shayer, M., & Adey, P. (Eds.) (2002). *Learning intelligence: cognitive acceleration across the curriculum 5 to 15 years.* Milton Keynes: Open University Press.

Sliwka, A. and Spencer, E. 2005. Scotland: Developing a Coherent Assessment System in *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris, OECD.
(Also available in French as *L'Evaluation Formative: Pour un Meilleur Apprentissage dans les Classes Secondaires*).

Spencer, E. 2006. *Features of Excellent Schools: Some Research Background* (paper to support *How Good Is Our School: The Journey to Excellence*. Her Majesty's Inspectors of Education (HMIE)).
Available at www.hmie.gov.uk/documents/publication/hgiosjte.pdf

Spencer, E., 1983, *Writing Matters Across the Curriculum*. Edinburgh. Hodder and Stoughton.

Swann, J. and Brown, S. 1997. The implementation of a National Curriculum and teachers' classroom thinking. *Research Papers in Education* 12, no. 1: 91–114.

Vygotsky, L. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Wiliam, D. (2007a). Keeping learning on track: classroom assessment and the regulation of learning. In *Second handbook of mathematics teaching and learning.* Ed. Lester Jr., F. K. (pp. 1053–1098). Greenwich, CT: Information Age.

Wiliam, D. (2007b). Content then process: Teacher learning communities in the service of formative assessment. In *Ahead of the curve: The power of assessment to transform teaching and learning.* Ed. Reeves, D. (pp. 182–204). Bloomington, IN: Solution Tree.

Wiliam, D., and Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In *The future of assessment: Shaping teaching and learning.* Ed. Dwyer, C. A. (pp. 53–82). Mahwah, NJ: Erlbaum.