National Foundation for Educational Research

A literature review of Assessment for Learning in science

Claire Hodgson and Katie Pyle

A literature review of Assessment for Learning in science

This literature review stems from a project funded by the National Foundation for Educational Research (NFER) which researched Assessment for Learning (AfL) in science with a focus on the primary age phase.

1 Key findings

- Assessment for learning has many generic features but there are some features of AfL which can be specifically honed for science teaching and learning.

- Classroom climate is particularly important. It is crucial that a co-constructivist, non-threatening environment is established in order for pupils to feel able to express their ideas and allow the teacher to establish what the pupils know, what they don't know and what they partly know – their misconceptions – and to develop teaching that will move their understanding on.

- The importance of talk, questioning, feedback, self- and peer-assessment are key in this area of research. The use of summative tests for formative purposes and the provision of resources and particular tools to support this area of teaching and learning are also discussed.

- There are some resource materials which have been written which support particular features of AfL in science, namely to identify pupil misconceptions and to encourage peer discussion. The work of Keogh and Naylor (Keogh and Naylor, 1997, 1998, 2004 and 2007) appears to be particularly well regarded in this area.

2 Background

Assessment for Learning has been defined as:

> 'the process of seeking and interpreting evidence for use by learners and their teachers to decide where learners are in their learning, where they need to go and how best to get there.' (Assessment Reform Group, 2002)

Research (e.g. Black and Wiliam, 1998) has indicated that AfL can benefit all children but is 'not something added to teaching, but is integral to it' (Harlen, 2006b p. 176).

A key starting point for this piece of work was *Science Inside the Black Box* (Black and Harrison, 2004) which was written to 'offer advice to teachers on how to interact more effectively with students, on a day-to-day basis, promoting their learning' (p. 1). Although aimed at secondary school teachers, Black and Harrison recognise that much of what they say is equally applicable to primary science teaching. They usefully point out that 'formative assessment has both *generic* features, i.e. features which will apply to learning across all stages and all school subjects, and features which are *specific*, for example to primary teachers and to individual subjects' (p2) and it is the focus on the 'specific' which has driven this literature review. Black and Harrison also helpfully offer an explanation for how science and formative assessment dovetail so well:

> 'Science provides the means by which learners can interact with the world around them and develop ideas about the phenomena they experience. So, when they attempt activities such as germinating seeds, their studies in science equip them with ways to observe and question what is happening. Through experiments they can begin to work out and predict what might happen if conditions for the growing plant change. To be able to learn science in this way, student needs help in developing process skills to investigate, and communication skills to question and discuss findings. Formative assessment fits well into this learning scenario, since its purpose is for teachers to sift the rich data that arise in classroom discussion and activity, so that professional judgements can be made about the next steps in learning.' (p. 3)

Against this background, the literature review sought to address the following research questions:

- What does AfL in science involve?
    - What AfL techniques are most commonly used in science?
    - Does science make use of different AfL techniques to other subjects?
    - How commonly is AfL used in science?

3 Literature Review

3.1 Literature searches

As the project was concerned with the interaction between assessment for learning and science, a formal systematic literature search was carried out by the NFER library to determine the extent of literature existing which included references to both of these areas.

The searches were defined by the following parameters:

- Material published between 1997 and 2008
- Focusing on key stage 2, but including other key stages where appropriate
- Research carried out in the UK primarily, but also extending to Australia, New Zealand, the Asian Pacific Rim and the USA

A range of different educational databases were searched and search strategies were developed using terms from the relevant thesauri (where available) and/or free-text searching. The same search strategies were adhered to as far as possible for all the databases. The full search strategy is detailed in Appendix 1.

A number of references were also taken from a bibliography published by the Institute of Education as part of their ‘Assessment for Learning’ module.

A bibliography of the reviewed literature is provided in section 5.

3.2 Findings

In reviewing the literature in this area, the aim has been to look at the interaction between AfL and science whilst attempting to focus on AfL principles and strategies which are most pertinent to the primary science classroom.

A total of 87 articles believed to relate to Assessment for Learning in science were found. A period of initial selection took place during which article abstracts were reviewed to see the extent to which they referenced AfL and science and were located within a primary context. Articles which were considered to refer to at least one of these areas (i.e. articles relating to AfL and primary science or to science more generally) were taken to the next stage of review. Following initial selection, 59 articles were reviewed in full of which 39 contributed directly to this review. Articles were reviewed following a standard format. In carrying out the reviews, it became increasingly obvious that the diverse features of what is known as Assessment for Learning were very often inter-related, with the literature often touching on a wide range of features within one article. For example, an article about pupil misconceptions may incorporate elements of classroom climate, talk and self-assessment. Once all reviews had been carried out a number of overarching themes emerged and the following literature review has been organised according to these themes. It should be noted, though, that due to the integrated nature of AfL it was not always possible to allocate a text to a discrete section. Several articles were, therefore, classified under a number of different themes and reference may be made to them in more than one area.

Many of the articles reported on small scale case studies, often carried out by teachers. These studies were mostly qualitative in their approach but there were some limited examples of quantitative data collection. These articles tended to explore the virtues of a particular AfL strategy and provided exemplification of how these strategies worked in class. A number of articles also reported case studies of teachers who monitored their understanding of AfL and how this changed over a period of time, for example as a result of an intervention or from doing particular activities. Another key feature of the articles included discussion of the need to develop a classroom climate which fosters a constructivist approach and encourages AfL practices. This was seen as particularly important in a science classroom where the need for questioning and discussion is so vital. Much of the literature also referred to the effect of AfL on the pupils who were involved. Discussions centred around achievement, performance, motivation, enjoyment and self-esteem which often related closely to the development of an appropriate classroom climate. These more generalised areas of the literature have mostly been incorporated into the thematically grouped reviews below.

3.2.1 Peer- and Self-assessment

In order to make learners more autonomous and able to identify their own learning needs and develop their own next steps, the skills of peer- and self-assessment are considered to be very important and several articles focus on the need to develop the skills though the science curriculum. Within this area, articles focused on a number of things including: the way that peer-assessment can feed into the development of self-assessment skills; the use of traffic lighting for peer-assessment and to indicate a self-evaluation by identifying own learning and understanding; the links between feedback from peers and from teachers and how the development of an AfL classroom climate can help to promote peer- and self-assessment activities.

The value of peer- and self-assessment is emphasised by Harlen in several articles and particularly in her 2007 paper in which Harlen identifies its importance as being in ‘helping children to take responsibility for their learning, an essential outcome of education’ (p. 30), thereby allowing pupils to direct their own learning, which will benefit them and their society in current and future situations inside and outside school. Following research about primary teachers who had implemented self-assessment by children, Harrison and Harlen (2006) reported a number of advantages and identified four aspects of self-assessment: self-monitoring and checking progress, diagnosis and recognition of learning needs, promoting good learning practices and linking learning practices. They identified that self-assessment is an essential component of AfL because it can help children direct their activities towards their learning goals. Within this, Harrison and Harlen identify the importance of a joint understanding of what ‘good work’ looks like which in turn relates to the need to develop appropriate success criteria. In discussion about peer-assessment, Harrison and Harlen describe one of the benefits as being that assessment takes place ‘without the pressure that comes from the unequal relationship between the child (novice) and the teacher (expert)' (p. 189). Harrison and Harlen further explain that peer-assessment builds on the AfL notion of learning as a co-constructivist activity whereby learning occurs as a result of social interaction. In this way, AfL contributes to learning that is in accord with current research into effective learning (e.g. Watkins *et al*., 2001; Wells, 2008)

Self- and peer-assessment is also the subject of some small scale studies in which the implementation of these strategies are discussed in relation to their use in the classroom. One such article, by Lindsay and Clarke (2001), describes the ways in which formative assessment builds a 'constructivist' classroom in which children are involved in 'creating and reflecting on their learning' (p. 15). Lindsay and Clarke reflect on a number of different methods of self- and peer-assessment which were used to support work done in science lessons in which practical investigations were carried out and include such activities as: self-marking, paired marking, plenary self-evaluation and self-assessment journals. In each of these activities the control over the learning process belongs to the children themselves. During self-marking children are invited to identify three areas where they believe they achieved the learning intention most effectively and one area which they felt needed addressing. Following on from this, the children worked on paired marking in which the partner was encouraged to mark the work and offer suggestions for 'closing the gap' and help one another to make an improvement. Lindsay and Clarke point out that peer-assessment offers valuable insight into the child's own understanding as they comment on the work of their partner, noting that 'children ... frequently demonstrate their level of understanding through their assessment comments' (p. 17). Through peer-assessment, teachers learn more about pupils and are therefore able to help them more effectively. Self-assessment journals are only touched on very briefly but do demonstrate children's ability to reflect on their work and make suggestions for the way forward, noting what they need to do next time to take account of what they have learnt from a particular experience. Lindsay and Clarke conclude that a large number of advantages have been recognised as a result of using the self- and paired assessment strategies including: children considering themselves as partners in the teaching-learning process which in turn raises self-esteem; teachers gaining greater insight into the child's understanding as a result of seeing them reflect on another child's assessment; children becoming more self-critical and pro-active learners as they focus on their next learning goal because it is set by themselves rather than being externally imposed. Lindsay and Clarke also identify advantages to these strategies in terms of science learning specifically and note that 'children become more scientific in their enquiries, as self-assessment encourages children to be constantly involved in the scientific process and their role within it' (p. 18). They also observed that self-assessment causes children to raise questions which 'constantly reinforces understanding of the skills and knowledge they are acquiring' (p. 18) and that as a result of paired-marking that children are writing and reporting for a purpose as an audience is defined for them.

Several areas of research into self- and peer-assessment explore the use of tools and techniques which can be used to support these areas. Concept mapping is one such way that is described as an effective way of self-assessing science in the primary classroom. Stow (1997) reported on concept mapping becoming ‘an increasingly common teaching and assessment strategy’ (p. 12) and indicated that it has been shown to fulfil formative assessment functions. Concept maps offer pupils a way to demonstrate their understanding of a particular concept by recording all the words and ideas they can think of related to the given word/topic/theme and showing the links between these. In this piece of research the pupils were asked to complete a concept map pre- and post-teaching of a given topic. The pupils were interviewed about their concept maps and they reported that they found the process useful and that it had helped some of them to identify the benefits of the mapping to their own thinking. It also helped some pupils to identify specific targets for future learning. In this way, concept mapping is seen as a useful self-assessment tool as it allows pupils to identify for themselves the areas about which they feel confident and the areas of knowledge which they need to develop. Furthermore, by completing the concept maps before and after a period of teaching pupils were able to monitor their own learning and progress. Stow concluded that concept mapping has two main positive features – that it is a motivational tool and that it offers metacognitive benefits.

Other techniques to support self- and peer-assessment are outlined in the research of Black and Harrison (2001b) and include colouring squares for goal statements, jigsawing and traffic lighting, question setting and comment-only marking. The research carried out as part of the KMOFAP project, involved secondary aged pupils and was based on science and maths teaching and learning. One method of self-assessment involved pupils in colouring squares to indicate their confidence in achieving a given goal. Black and Harrison reported that pupils were found to be honest in their completion of this task because it was carried out as a private activity, the result of which was only seen by the pupil and the teacher. This method allowed pupils to alert the teacher as to who needed help whilst actively involving pupils in their own work and developing skills of metacognition. A similar method is the use of traffic lighting. In this instance pupils were encouraged to colour code/traffic light (red, amber, green) particular words to indicate their level of understanding of that word with the target being to move up on the traffic light scale in their understanding as part of their homework or target setting. Traffic lights were also described in terms of a peer-assessment activity. Having revised a topic, pupils were asked to present a particular aspect to their peers. Peer-assessment was given via the means of traffic lights in which a green light indicates that the presentation was better than they could have given themselves; an amber light indicates the presentation was as good as they could have given themselves and a red light indicates that the presentation was not well explained. The immediate visual representation offers immediate feedback and allows for the presenter to determine their level of success. Comments from peers about what was and what wasn’t so successful also helps pupils to focus on their strengths and areas for development. By carrying out peer-assessment, the peers are also able to identify and assess their own understanding of the given topic. Black and Harrison also report on the use of peer-marking. Although pupils may initially focus on surface features, such as legibility, it can lead to pupils looking critically at content. The feedback from teachers involved in the research indicated that this method of marking led to ‘misconceptions com[ing] to the fore’ (p. 47) and had a greater impact on pupils willingness to present work clearly.

Harrison and Harlen (2006) conclude that self- and peer-assessment both 'engage children in being reflective, both about the task in hand and more broadly about the way they learn [and therefore] encourage a deep rather than a surface approach to learning' (p. 190) and this is reflected in all of the other research in this area whether or not use is made of specific tools and techniques.

Self- and peer-assessment are seen as the cornerstone of good AfL practice and have clear relevance for use in science teaching and learning because, for example, it allows for the identification of misconceptions and encourages pupils to direct their own learning and to become more involved in the scientific process.

3.2.2 Misconceptions and classroom climate

The importance of eliciting pupils' ideas about science through questioning and discussion is evident in much of the literature and finding out pupil misconceptions is considered a very valuable AfL benefit. Research (e.g. Keogh and Naylor, 1998, 2007) shows that children have a range of ideas about a wide variety of science topics and that some of these ideas are incorrect – and, perhaps more importantly, that it can be very difficult to encourage pupils to adopt the correct view. By finding out about a pupil's understanding at the start of a topic and establishing if there are any misconceptions (for example, believing that the moon only comes out at night) it is possible to plan teaching to challenge this idea and to shape future learning and understanding. Closely related to this area is the notion of developing a classroom climate in which pupils are willing to discuss their ideas and are not afraid of being wrong. Research shows that the willingness to discuss ideas in a non-threatening environment is crucial to finding out and addressing pupil misconceptions.

Much of the area of pupil misconception is discussed in relation to 'concept cartoons'. A concept cartoon is the presentation of a scientific concept which a group of 'cartoon children' are discussing. Pupils are encouraged to discuss a variety of different ideas around the given theme, allowing teachers to find out about pupil misconceptions in relation to their skills, understanding and knowledge. However, a number of articles also considered other ways of finding out about pupil misconceptions and the wider implications of this.

Keogh and Naylor (1998, 2004) are the creators of 'concept cartoons' with particular application in science and within that, investigative science. Concept cartoons were developed as a result of research, such as that of Newton *et al.* (1999), which suggested a lack of suitable stimuli for generating talk in science. They wished to develop a method or strategy that would look at everyday contexts from different viewpoints. In one article, Keogh and Naylor (2007) report that one way of doing this is to use puppets to represent different viewpoints, taking the ownership of the ideas and conceptions away from individual pupils so judgements are on the puppets. They found that when pupils were able to speak of an idea as a puppet's, rather than their own, that more of pupils' talk was justified in an evidence based manner. This is an essential quality for science learning, particularly investigative science.

Keogh and Naylor (2007) highlight that worksheets can inhibit conversation as tasks become individual and not shared. However, they further highlight that self-esteem is an issue amongst pupils when teachers are trying to establish pupil knowledge verbally due to reluctance by pupils to talk, fearing they may be wrong. It is important for teachers to suspend judgment in discussions in the class to avert damage to self-esteem and to allow for thinking and discussions to ensue. Appropriate open ended questioning can be fitting in such discussions.

Using concept cartoons is one method, developed from this thinking, for fostering a non-threatening method for eliciting misconceptions in pupils. A concept cartoon will generally pose a question which may gain mixed responses from pupils (cognitive conflict). A range of answers, already provided, will include a correct response but also will include responses which contain common misconceptions about the topic in question. Pupils can then select their answers from this range of options or suggest which ones might be true or false, allowing the teacher to find out the misconceptions held by the pupils. Naylor and Keogh (1998, 2004, 2007) report that pupils feel comfortable responding in this way rather than answering from a blank slate where they believe there may be stigma attached to their individually created response. Such questions and their response options can, for example, be used as the basis and motivation for investigative work in science so pupils can discover for themselves if options are true or false by conducting an investigation to find out. This is more likely to lead to a shift in thinking if pupils are able to demonstrate to themselves that their original thinking was incorrect.

In 1998, Keogh and Naylor conducted a review and update of their work using concept cartoons as a teaching and learning activity, following extensive feedback from teachers. They found that concept cartoons work most effectively when using everyday situations, minimal amounts of text and using positive statements rather than negative ones. The teachers in the study reported that children who are reluctant to put forward ideas normally are more relaxed in doing so when using concept cartoons. They also found that concept cartoons are effective in the promotion of discussion of ideas and also encourage shifts of ideas and thoughts. They were also found to be motivating and helped to keep pupils ‘on-task’. As a result of the nature of concept cartoons, limited text and cartoon-like appearance, teachers found them to be especially helpful for children with learning difficulties.

Macro and McFall (2004) highlight another method for encouraging responses from pupils, which in turn leads to teachers being better able to tease out misconceptions. They outline a method where pupils shut their eyes when they have been asked a question and then put up their hands when they have an answer. This encourages pupils to think for themselves without being intimidated by other pupils who instantly raise their hands. Macro and McFall found that this eye shutting method actually results in more pupils attempting an answer. They highlight that having more time to think also encourages more pupils to respond. For younger pupils, they highlight that drama or puppet use are effective mechanisms for drawing out responses as de-personalising issues in this way encourages more focussed and applied thinking.

Sato *et al.* (2005) conducted a study looking at how teachers modify class practice to take account of new ideas in assessment, and in particular ideas about assessment for learning. One teacher in the study employed the use of ‘question of the day’ where responses were recorded by pupils in ‘learning journals’. This allowed for the shift in ideas to be recorded by pupils. Marking of these journals reflected pupil awareness of the changing of their thoughts and ideas and the use of evidence to support this. This system makes it acceptable for a pupil to have a misconception about a situation and rewards pupils for shifting their belief about the situation in an evidence-based way.

Assessments can also help to identify misconceptions. McNair (2004) discusses three types of assessment; pre-assessment, formative assessment and post-assessment. These assessments have differing purposes throughout the learning process, assessing what is known before a particular topic, what is being learnt during the topic and what has been learnt. The modes of assessment McNair outlines, can range from drawings, journals and graphic organisers, to individual interviews. Graphic organisers, for example, allow pupils to explore their ideas and indulge in reflective understanding through the medium of flow charts and Venn diagrams. Drawings might reveal how a pupil perceives an object or an idea and how perceptions change from pre-assessment through to formative and post-assessment. Hence, these modes are useful for identifying misconceptions and shifts in thinking. Although McNair has not related these ideas specifically to science teaching and learning, the principles of using less written assessment and increasing different modes of communication, relate directly to Keogh and Naylor’s principles of more talk and emphasis on simple and visual stimuli. These principles are directly applicable to the science classroom where misconceptions are rife.

Millar and Hames (2002) highlight that it is a risk to assume that a pupil understands a model or a concept on the basis of their answer to one question. They provided science teachers with a bank of diagnostic questions which are designed to help provide information about what pupils understand and shed light on misconceptions. They found that only 54 per cent of pupils used the same scientific model to answer two similar questions. Millar and Hames highlight that many concepts in science are not understood or that there is only a limited amount of understanding. They found that diagnostic questioning leads to investigative science as pupils wish to find out answers. Anecdotal evidence suggests pupils are more accurate in their investigative science when it is presented in this way as they are very keen to make sure experiments are fair so they can find out the answers to the questions set. This mirrors Keogh and Naylor’s assertions about the use of concept cartoons in investigative science leading to increased motivation to conduct investigations in order to find answers and, ultimately, to shifts towards a more accurate understanding of scientific principles as a result.

Different methods for identifying and dealing with misconceptions in science have also been outlined and include verbal and non-verbal methodologies. All of these methods of teasing out misconceptions require, and highlight the need for, an appropriate climate in classrooms that allows pupils to be comfortable in sharing their ideas and misconceptions without ridicule or embarrassment. Furthermore these methods allow for a shift in thought pattern, guiding pupils through finding out the right answers to questions in a non-threatening and productive way. The literature highlights that this appropriate classroom climate is essential for the effective derivation of pupil misconceptions to allow for good teaching and learning to occur.

Keogh and Naylor (2004) suggest that concept cartoons allow children to feel that their ideas are valued, encouraging them to be shared and discussed. When using concept cartoons, all ideas have equal status and help to encourage an environment where mistakes and misconceptions are part of the learning process.

Qualter and Taylor (1999) emphasise the importance of developing a classroom climate where pupils feel comfortable to take risks. They document one teacher's approach to fostering this climate in science. They emphasise the importance of legitimising pupils' ideas by repeating them and writing them down rather than dismissing any that might be wrong. This gives ideas importance. Although upon reflection the teacher felt her enthusiasm for ideas was false, the pupils had not felt this way and it generated an open environment. This, Qualter and Taylor note, is in line with the theory of scaffold learning where an open and warm reception from the more knowledgeable adult, is key. Although it is naturally not the only exclusive condition for learning, it does suggest a way in which pupils will share ideas and hence, misconceptions can be identified.

Daws and Singh (1999) also draw attention to the importance of the classroom environment. They reported on three case studies of teachers using AfL strategies in secondary science classrooms, including the scrutiny of exams and mark schemes with pupils, the provision of learning objectives before the start of a unit and the recording of marks. The case studies emphasised that assessment for learning helps to reduce stress and anxiety and develop positive teamwork skills as well as helping pupils to articulate views and listen to others. Interestingly this suggests assessment for learning techniques actually help to foster the necessary classroom environment for effective learning. This results in a circular model where a good climate is needed for assessment for learning to work, while assessment for learning itself promotes and maintains a good classroom climate.

The importance of identifying misconceptions is implicit in good teaching and effective learning. Ascertaining misconceptions to use in an assessment for learning framework is also important for the stimulation of confidence. Good classroom climate where pupils are free and comfortable to present their ideas, promotes confidence which in turn stimulates good learning and, as Stark and Gray (2001) found, better performance in tests and summative assessments.

In summary, the inter-relationship between misconceptions, questioning and classroom climate is particularly strong. The identification of misconceptions in science is very important as it allows for misunderstandings to be addressed and for learning to be developed and advanced. However, the discovery of pupils' misconceptions is reliant on the development of a classroom climate of trust where questioning can probe pupils' understanding in a non-threatening environment.

3.2.3 Talk

This was a recurring theme in many articles and the importance of talk in a science classroom was often linked to the development of an AfL climate and as part of the discussions about peer- and self-assessment and pupil misconceptions. Talk is vital as part of the process of collaborative knowledge construction. Through talk a number of AfL opportunities arise, as described in Hargreaves (2007):

> ‘ideas are proffered, group members assess them in open dialogue and in doing so the person proffering receives immediate feedback that moves their thinking forward (they learn). More than this, in a cascade of similar learning events, new knowledge is constructed by the group…’ (p. 195)

If pupils are working silently, their ideas are not being assessed, challenged or fed back on, but through talk they can be. It is this idea which is explored further below.

Black and Harrison (2004) extol the virtues of discussion within a classroom where ‘students feel they can reveal current understanding and be helped to firmer understanding’ and suggest that this is ‘an essential ingredient to making formative assessment function in the classroom’ (p9). One of the advantages of encouraging peer discussion is that it allows pupils to discuss ideas with one another and to check ideas before revealing a group answer to the whole class. This can offer useful reassurance and it can be less threatening if an answer is wrong but is offered as a group response. Ways for stimulating this type of discussion are discussed further in the section above.

Asoko and Scott (2006) highlight the importance of language in science learning, particularly opportunities in the science classroom to learn how to talk about natural phenomenon in a scientific way. Harlen (2004) also highlights the use of speech for reflection and communication, with reflection occurring when pupils are sorting out their ideas. This involves listening and expressing ideas in a way that is understandable and allowing for the linking of old ideas with new ones to create bigger ideas. This thinking is imperative to science learning. Harlen suggests particular importance for dialogue in *primary* science while the brain is developing at this early stage. Evans (2001) echoes this when he writes that there should be more use of the oral response mode in assessment, rather than written or drawn responses, particularly for younger pupils whose oral skills are generally more advanced than their written ones.

Keogh and Naylor (2007) are also keen proponents of the notion that science lessons should be environments where pupils are willing to consider alternative ideas, justify their opinions and base their decisions on evidence and reasoning. Talk is therefore highly valued in science and it can play a vital part in the process of teasing out misconceptions.

Keogh and Naylor (2007) collected data from pupils and teachers in science lessons and teachers on INSET training and found that pupils enjoy focused talk in science and freely engage in this kind of talk when presented with scientific questions. Pupils can engage in this kind of talk without either them or their teachers being taught special rules. It is suggested that productive talk is absent as a result of the burden of the work in the National Curriculum. Talk is seen an add-on rather than a central feature of the scientific classroom.

Different kinds of talk have been identified by researchers with reference to the talk between teachers and pupils. Many of these are linked to the discussion on questioning. The different kinds of talk include:

- Triadic dialogue - this involves three stages; an initiation (normally by the teacher), pupil response and then a teacher evaluation. Chin (2006) suggests this method is ‘often perceived to have restrictive effects on pupil thinking as responses are brief and teacher framed’ (p. 1316). Due to the fact that this type of talk minimises co-construction of meaning, this type of teacher questioning has received some criticism although it is acknowledged that it does have some functionality if used to scaffold pupils’ knowledge.

- Authorative discourse - in this interaction between teacher and pupils, the teacher conveys information that involves instructional questions, factual statements and reviews. Pupils respond with single, detached words to the questions. It is described by Harlen (2004) as being ‘controlled by the teacher’ (p. 18) and Asoko and Scott (2006) explain that it ‘does not really explore and take account of children’s ideas as they arise’ (p. 160). Chin (2006) suggests that in authorative discourse, ideas or questions that do not link to the school science, tend to be re-shaped by the teacher or ignored.

- Interactive/non-interactive: Interactive talk describes the interaction between the teacher and pupils. In interactive talk, the teacher explores pupils’ views and takes account of them, even if they are different from scientific ones. Chin (2006) outlines that the nature of questioning that runs alongside this kind of talk is constructivist-based or forms part of enquiry-oriented lessons. Teachers aim to elicit what pupils think about a topic and the questioning is used to diagnose what pupils think and extend pupils’ ideas using a scaffolding method. Pupils’ points of view and ideas are represented through talk. Responses from pupils are longer and require high order thinking. This links with what was outlined in the misconceptions sections, suggesting a climate that allows for the exploration of pupils’ views, even when these views may not be correct. Chin (2006) considers interactive talk as ‘dialogue’ but Asoko and Scott (2006) view dialogue in opposition to authorative discourse. They view it as a measure of how much the teacher takes into account different points of view and represents these through talk.

There can be a place for these kinds of talk depending on what is trying be achieved and Asoko and Scott (2006) suggest that ‘effective teaching involves all these approaches’ (p. 163) with many teachers being able to switch between the different kinds of talk as necessary.

Asoko and Scott (2006) also cite the research of Mercer *et al.* (2004) which discusses the dialogue between pupils and illuminates the need for exploratory talk between pupils. This is where pupils share information, all are invited to contribute, opinions are respected and considered and reasoning must be clear. Mercer *et al.* showed that pupils who had experienced how to do exploratory talk scored significantly better in assessments of science knowledge than pupils in control groups.

Strategies for promoting talk were also outlined in this literature. In particular, Keogh and Naylor (1998, 2004, 2007) discuss the use of concept cartoons to promote talk, especially to encourage interactive/dialogic talk - the kind of talk that allows different views to be explored and not dismissed. They suggest that using ideas such as concept cartoons to generate a purpose for talk makes the process more natural. It gives pupils the opportunity to look at different viewpoints and encourages debate. This encourages pupils to become self-motivated and consequently conversations sustain themselves.

Talk can be hampered if pupils feel a threat to their self-esteem. This links back to the necessity for a good climate in the science classroom where pupils feel that their opinions are respected. The use of concept cartoons is one example of fostering a climate of mutual respect of ideas which can generate talk which is more exploratory and interactive.

Science can be abstract and misconceptions can be rife. Talk and the stimulation of talk offer opportunities for pupils to develop ideas, to reflect and put ideas into context and to listen to other pupils' ideas. In turn, talk should allow pupils to develop better understanding of science through better learning. There is also a very clear link, and much overlap, between talk and questioning which should be borne in mind when considering the next section.

3.2.4 Questioning

Questioning is essential to both AfL and to science learning. It is one of the key AfL strategies discussed in the literature with the use of questioning referring to those questions asked by both teachers and pupils. Van Zee *et al.* (2001) identified that "questioning is a frequent component of science talk" (p. 160) and Harlen (2006c) states that "questioning is a key feature of scientific activity and of teaching science" (p. 167). The importance of questioning was evident within a number of articles and fell into two broad categories – articles which explained how to use questioning in the science classroom and articles which reported on research/case studies of questioning in use.

Articles concerned with questioning tended to discuss the use of:

- effective questions
- open questions
- questioning to find out about pupil misconceptions
- questioning to form part of feedback to prompt further learning

A number of articles also considered the development of a classroom climate to promote questioning.

Black and Harrison (2004) identify that teacher questioning can be used for a variety of purposes - including: to encourage comparison; categorising, grouping and recognising exceptions; predicting – and that they have a range of roles within the classroom. They state that 'to exploit formative opportunities it is necessary to move away from the routine of limited factual questions and to refocus attention on the quality and the different functions of classroom questions' (p. 6). Although teachers do sometimes need to ask closed questions in order to check pupils' knowledge, it is often the case that science questioning should ask the pupils to 'delve deep into their conceptual learning' (p. 6) and to ask 'rich' questions. Black and Harrison describe these as questions which 'cannot be answered immediately but rather requires the learner to work on a series of smaller questions and activities before they return to have a stab at answering it' (p. 7). Black and Harrison explain that through questioning teachers are able to collect evidence about pupils' understanding with the aim of finding out what they do know, what they don't know and what they partly know. This then provides the starting point for the teaching allowing for the pupils' knowledge and understanding to be moved on.

Several articles revealed that questioning and the resultant answers can be a useful tool for diagnosing misconceptions. Khwaja and Saxton (2001) report on the importance of asking suitably focused questions. They point out the need for clarity in the questions being asked in order to obtain the right level of detail from the respondee. They give the example of comparing responses given to the question/instruction to 'draw what is inside the body' compared to 'draw the bones inside the body'. One version may give rise to a child showing a good understanding of, for example, the ribs but the other may give misleading information because the child omits the ribs from the diagram in order to show what is deeper inside the body. It would be easy for a child's understanding or misconceptions to be misinterpreted because of the actual question being asked. Thus, the importance of clear, direct questioning is made apparent.

Millar and Hames (2002) indicate that questioning can be used to support and improve teaching. In their research into the use of diagnostic assessment to improve science teaching, they report on how children could be encouraged to carry out investigations in order to find their own answers. Related to this is the use of 'big questions' which Black and Harrison (2001a) describe as being used to set the scene for the whole lesson with the subsequent development of smaller questions being introduced to help answer it. Harlen (2006c) also discusses the usefulness of displaying a question and then reviewing it over the course of the lesson(s), thus allowing for the teacher and the pupils to monitor their developing understanding. Macro and McFall (2004) note that whilst teachers can plan their initial questioning it is often the children's responses which determine what the next question should be thus indicating the flexible and responsive nature of questioning.

As well as questioning referring to those questions asked by teachers, Harlen (2006c, 2006d) identifies the importance of pupils asking questions. Harlen (2006c) explains that children's questions show the 'cutting edge' of their understanding because they will ask questions about things that they do not know or understand. By asking their own questions, children are able to reveal their misconceptions as the question will show the limit of their understanding and the nature of their own ideas. When given the opportunity to investigate their own questions in order to find things out for themselves pupils gain satisfaction and motivation for learning.

Chin (2004) identifies that 'questioning is key to active and meaningful learning and is the cornerstone of scientific enquiry' (p. 107) and supports Harlen's stance that questions should come from both the teacher and the pupils. When posed by the teacher, Chin states that they can help pupils to 'explore and scaffold ideas, steer thinking ... advance students' understanding of scientific concepts and phenomena' (p. 107) whilst questions posed by pupils can help them to 'fill recognised knowledge gaps and solve problems' (p. 107). Chin comments that in a typical classroom pupils are much more likely to be asked than to ask and suggests that this pattern should be reversed. She suggests a number of ways that this can be achieved, such as asking for pupils to suggest questions at the start of a new topic, to demonstrate how a 'big question' can be broken down into smaller questions and encouraging pupils to pose questions before doing an activity in order to direct their own enquiry. Chin also proposes that teachers should model the formation of good questions and provide stimulus materials which will provoke such questions. She refers to a number of different question types and suggests that pupils should be made aware of them in order that they can think about answering them in different ways. She also references 'productive questions' (p. 110) which are designed to 'stimulate productive physical or mental activity and reasoning and take them forward in their thinking' (p. 110) and which are particularly useful in a science classroom context. Chin recognises that some teachers may feel overwhelmed by pupils who ask large numbers of questions, particularly if they may not know the answer, but suggests that this provides an opportunity for teachers to teach pupils about how and where, beyond the teacher, they might find answers and suggest that they offer answers and explanations to their own and each other's questions.

Chin (2006) reports on a study which looked at teacher questioning and feedback. Chin explains that questioning is constructivist based with the teacher's intention being to elicit what pupils think, to elaborate on previous answers and to construct conceptual knowledge. Questioning is seen as a way to diagnose and extend pupils' ideas and to scaffold their thinking. This study involved secondary aged pupils in Singapore and, as such, is outside the desired remit of this literature review, but it is interesting that Chin identifies that questioning is seen as a significant part of teaching and science talk. When teachers paraphrase a pupil's response, Chin states that this can allow pupils the opportunity to co-construct a response with their teacher and peers. Chin also points out the possibility of giving corrective feedback which can be overt or implicit. Implicit feedback may provide a constructive challenge if the teacher asks further challenging questions or recasts the questions. Questioning can also be used to help pupils with sequencing using a variety of cognitive processes such as comparing, hypothesising, predicting, explaining and interpreting and can also help to move students on with their thinking. Chin also found, much like Macro and McFall (2004), that it was important for teachers to be able to adjust their questioning to accommodate a range of pupil responses, to respond to pupils' thinking and to guide them through inquiry-based discussions.

Many articles discussed the importance of 'wait time' – allowing sufficient time for pupils to consider the question and their response, rather than rushing to answer immediately. Research (e.g. Budd Rowe, 1974) indicates that allowing a period of up to 8-9 seconds or more can encourage longer and more thoughtful responses which go beyond factual recall. Black and Harrison (2001a) identify that the practice of 'wait time' can be difficult to adopt initially and that waiting for several seconds can seem 'painful', but the benefits outweigh these difficulties as students become accustomed to having their ideas challenged. Alongside the 'wait time' technique, the strategy of 'no hands up' is also usefully employed as it implies that everyone is expected to have an answer and that thought rather than speed is valued. Harlen (2006d) further indicates that it is important to avoid rephrasing a question if it is not readily answered because it 'inevitably makes it more closed and less useful' (p. 65). This can lead children to expect that if they wait, the question will always be rephrased and made simpler. She also suggests that paired or small group discussion can encourage responses, a view supported by Macro and McFall (2004) who found that children are more confident at answering a question when working in a small group.

Van Zee *et al.* (2001) carried out a study to investigate student and teacher questioning during science talk. The study, based in the United States, involved students across a range of ages including primary, upper elementary and high school. Van Zee *et al.* discovered that students would ask questions if they were provided with opportunities to do so. Use of structures such as a KWHL chart, where a student identifies what they **Know** and **Wonder** about a topic and identifies **How** they can find out and decide what they have **Learned**, helps teachers to ask appropriate questions of their students. Brainstorming a topic and asking students questions during a discussion also helped elicit knowledge and prompt student questioning. Van Zee *et al.* discovered that students were able to ask questions that were grounded in their own observations and where they felt comfortable to discuss ideas with their peers in order to try and understand one another's thinking. Finally, van Zee *et al.* identified that student questions occurred where there were small groups of students who were collaborating with one another. This mirrors the findings of Harlen (2006d) and of Macro and McFall (2004). The research of van Zee *et al.* also looked at teacher questioning and they concluded that questioning could be used to develop conceptual understanding. Questioning could also be used to encourage pupils to elucidate their meanings and to explore a variety of points of view. The use of 'quietness' and reflective questioning was also investigated whereby 'wait time' and the provision of information on a 'need to know' basis was used to encourage student thinking. Van Zee *et al.* concluded that it was possible for students to formulate their own questions and will do so if given appropriate opportunities.

As this area of research shows, questioning is a crucial part of science learning and, used appropriately, is an extremely useful AfL tool as it allows opportunities to find out what pupils know, to identify pupils' misconceptions and to plan the next steps in the pupils' learning.

3.2.5 Feedback

In terms of providing feedback to pupils about their learning a number of methods are identified which relate specifically to feedback in science. Articles refer to comment only marking, the use of feedback to inform next steps; the role of feedback in peer- and self-assessment and the relationship between feedback and learning objectives and success criteria. The development of success criteria is a key AfL strategy. The success criteria (which may also be the assessment criteria for a task) are shared with the pupils with the aim of raising awareness of what is expected of them during their learning. Articles concerning success/assessment criteria discuss the value of pupil involvement in devising them and creating awareness of their importance. Success criteria are an integral part of many of the discussions on feedback and, as such, have been integrated into this section.

Black and Harrison (2004) identify that there are two types of feedback which are 'essential to formative assessment: the first is from student to teacher, the second from teacher to student. Learning is effected by alternation between these, in which each contribution responds to the other' (p. 3). They also explain that effective feedback arises from 'learning experiences that provide rich evidence so that judgements about the next step in learning can be made' (p. 5) and this provides the link to other areas of AfL such as the provision of challenging activities to promote thinking and discussion, the use of rich questions, provision of strategies to support learners in revealing their ideas and opportunities for peer-discussion and larger group discussion to encourage open dialogue.

Harlen (2006b) also identifies that feedback is a two-way process. She comments that it is important for pupils to give feedback to teachers in terms of allowing the teacher to see where the children are and what would be the appropriate next steps. Similarly, the children obtain feedback directly from self-assessment or from the teacher or other children. Harlen usefully points out some of the key features of feedback if it is to be useful, such as the need for non-judgemental comments and to show where improvements can be made. She highlights the need to allow time for pupils to react to and act upon the feedback they have been given in order to 'convey the message that responding to the comments is part of their learning' (p. 178). She further suggests that in valuing the comments in this way, pupils can incorporate them when giving their own feedback as part of peer-assessment.

Harrison *et al.* (2001) make the link between learning objectives and feedback and demonstrate feedback as a two-way process in their research into formative assessment. They refer to teachers developing self-assessment strategies which not only helped pupils but formed an 'important feedback mechanism from the pupil to the teacher regarding the pupil's confidence in their current work' (p. 19). This was achieved by asking pupils to rate their understanding of a learning objective on a scale of 1-3. This visual response, with fingers held up to indicate where they fell on the scale, allowed the teacher to see who needed help with particular concepts thus providing immediate feedback on what further teaching was required. The teacher involved also identified that the sharing of learning objectives provided a useful reference point for any written feedback as it has 'an immediate term of reference' (p. 20). The research indicated that because children knew what the learning objective was that they knew what the marking criteria would be based on. It was also felt that the task was 'intrinsically more worthwhile' (p. 21).

Related to the giving of feedback is the notion that learning objectives must be clear, afterall, assessment should be to see if the learning objectives and projected outcomes have been achieved. Leakey (2001) reported on her own experiences of sharing learning objectives with children of a range of ages and abilities and found it to improve her teaching mainly because it gives children ‘ownership of their own learning’ (p. 68). Leakey suggests that it is only through having an understanding of the learning objectives, and thus what will be assessed, that children are able to make sense of any feedback that they receive. Leakey describes the need to provide constructive feedback during discussions and practical tasks as well as on written exercises. These can be given in the form of prompts aimed at moving the learning forward. In terms of written work she suggests that questions to prompt more accurate and appropriate answers are required but also suggests that simply writing comments is not sufficient – dialogue with the pupils based on the written feedback is required in order to encourage the pupils to take the next step forwards.

Whilst it is clear that it is important to provide feedback, what is vital is that the feedback given is effective in moving the learners on. Black and Harrison (2004) explore the features of effective feedback and these can be summarised as follows:

Effective feedback is that which:

- should initiate thinking enabling the learner ‘to discuss his or her thoughts with the teacher or a peer’ (p. 12) in order to instigate improvement
- prompts immediate action
- ‘relates back to the success criteria’ (p. 13)
- allows learners to match their own judgement of quality against that of the teacher or peer
- may direct learners ‘where to go for help and what they can do to improve’ (p. 13) their work.

In order to set some of the findings in this area in context, it is worth summarising some of the influential research of Butler. Although the actual research of Butler (1987, 1988) falls outside the remit/timeframe of this review, it has influenced much of the research which is discussed in the area of feedback. Butler compared the effects of providing different types of marking where feedback was given in different ways. Pupils in one group were given only marks or grades, pupils in another group received grades and comments and a third group of pupils received only comments. Butler concluded that comment-only marking tended to result in the greatest learning gain when compared with children who received just marks or marks and comments. References to comment-only marking are relatively frequent in the literature relating to feedback.

Many of the references to feedback in the articles in this area concern feedback to written work produced in the secondary school context. For example, Black and Harrison (2001a) report on research carried out to look at feedback given by secondary teachers of science. They report on the use of ‘comment only’ marking in which teachers write comments on pupils’ work rather than giving a grade or mark and this built upon the work of Butler (1998). In their own research, Black and Harrison (2001a) note that the teacher needs to pay close attention to the comments that they give in order to ensure that they are pertinent and relevant. Whilst it was found that this resulted in the marking process taking much longer, teachers reported that they were soon able to identify what was an effective comment which would prompt pupils to move on and develop their thinking and learning. The teachers involved in the study reported that the use of comment-only marking resulted in a much more personalised response and that comments were honed to the recipient, often with references to previous pieces of work. One teacher actively encouraged the pupils to respond to the feedback through the use of a comment sheet where the pupil responded to teacher comments with their own thoughts and evidence of the changes they had made.

Gioka (2006) also looks at the work of teachers teaching science to 11-18 year olds and concluded that feedback only fulfils a formative function when it provides information to help 'close the gap' and that teachers who give feedback also need to allow time for pupils to respond to the comments made. As part of the feedback, it is suggested that questions can be a useful tool for challenging responses and encouraging further thinking.

These ideas are further explored in the research of Markwick *et al.* (2003) who investigated the alternative ways of marking work with an emphasis on AfL techniques. Like Gioka, their work was based on science teaching in a secondary school with comparisons being made across key stages 3 and 4. The study took place over a two year period with the first year being used as a 'control' year in which assessments were carried out in a summative style. In the second year, formative styles of assessment, which included 'open ended questions and comments to guide deeper thinking' (p. 51), were used with the same groups. Although teachers reported making use of summative results to help plan future teaching, the results of end-of-unit tests were rarely used by the pupils themselves to set targets. The results of the study, which involved regular interviews between teachers and pupils, indicated a big increase in the percentage of pupils (from 35 per cent to 85 per cent) who said that they acted upon comments made by teachers between the first and second years of the study. Some pupils indicated that they would also like to receive oral feedback on their work. As in the research of Black and Harrison (2001a), Markwick *et al.* (2003) report that teachers involved in the study found that formative marking took considerably longer to do but felt that it was worth it. Advantages of such marking included being able to provide pupils with higher quality information and being able to differentiate work more effectively. A second strand of this study was the introduction of interviews in which students talked about their science work with their teacher. Pupils reported that this gave them 'a clearer idea about where they were in terms of their potential and how they could continue to improve' (p. 53). Perhaps most important was the fact that several students commented that the interviews 'made them feel as if their science was very important' (p. 53). Markwick *et al.* do point out that this study was relatively small scale and that the pupils were excited to be involved and that 'a more rigorous study would be required to improve confidence in this interpretation' (p. 54) but, nevertheless, the study did show that the changes in the methods of providing feedback 'dramatically affected the way students became involved in their own learning' (p. 54) and through this involvement came an improved enjoyment of science.

In her study of classroom interactions in science, conducted in secondary schools in Singapore, Chin (2006) discussed the inter-relationship of questioning and feedback and proposed four different ways of providing feedback designed to develop pupils' learning. In the first, the teacher affirms the response given by the pupil and carries on to reinforce the response and provide further teaching. In the second, the teacher accepts the response given by the pupil but then goes on to ask a series of related questions to 'probe or extend conceptual thinking' (p. 1326). In the case of an incorrect answer being given, Chin identifies two methods of response. In the first, she suggests that explicit correction is required and a reinforcement of the teaching points. In the other method, Chin suggests that an evaluative or neutral comment is made followed by a 'reformulation of the question or challenge via another question' (p. 1326). She suggests that in this way a constructive challenge is provided which 'forces the student to reflect on and reconsider her answer' (p. 1334). In addition to having a range of response methods, Chin also suggests that when providing teacher feedback in the form of paraphrasing a pupil's response, it can help to verbalise the pupil's thoughts and provide the opportunity to co-construct a response with the teacher and their peers. This can be particularly helpful for pupils with weaker language skills and can help to provide conceptual and linguistic scaffolding which can 'adjust the cognitive and linguistic loads of students' (p. 1336).

Feedback forms a crucial part of the AfL model providing pupils with information about where they are and where they need to go to next with their learning. There is a clear link between questioning and feedback with the role of feedback in developing pupils' thinking and learning being vital.

3.2.6 The relationship between formative and summative assessment

A number of articles highlighted the potential tension between AfL and summative assessment/testing. Whilst some articles focused on the impact that summative testing can have on what is taught in science, other articles illustrated ways that summative tests can be used in formative ways.

Harlen (1999) discusses the importance of formative assessment in science particularly in the assessment of scientific process skills. She states that 'Assessment which has a formative purpose is essentially in the hands of the teachers and students and so, theoretically, can be part of every activity in which science process skills are used' (p. 133) and goes on to describe how information can be gathered in the form of observations, questioning, specific task setting and asking pupils to 'communicate their thinking through drawings, artefacts, actions, role play and concept mapping, as well as writing' (p133) which allows for on-going assessment and the opportunity for pupils to show a growing and developing understanding. Harlen discusses how formative assessments can be used to derive summative judgements by reviewing formative evidence against the standards or criteria used to describe levels of performance. She points out the drawbacks to this method, however, indicating that 'it depends for its validity on opportunities having been created for students to show what they can do and on the teacher having collected the relevant evidence' (p. 136). Harlen suggests therefore that 'special assessment tasks' should be made available to allow pupils to show the skills that they have. The combination of a summary of on-going assessment and some well-designed practical tasks is seen by Harlen as 'the best compromise for practical skills' (p. 137) assessment. In contrast, Harlen questions the use of written tasks to assess practical skills (as used, for example, in the Key Stage 2 science tests) and she queries the validity of this type of assessment. Harlen also discusses the impact of summative testing on the curriculum and on formative assessment practices. She notes the evidence relating to the introduction of summative testing at ages 7, 11 and 14 and the narrowing effect that this has had on teaching, as she says 'When the assessment stakes are high for students and teachers, what is tested inevitably has a strong determining effect on what is taught' (p. 137).

It is this concern with how test content affects what is taught in class (the notion of teaching to the test) that is raised by Peacock (2005) in an interview with Naylor and Keogh. They describe how their work has revealed that some teachers perceive that rote learning, rather than thinking, is rewarded. This is illustrated by the quotation of one teacher who said that '...if I get children thinking, they'll do worse in the tests, because the mark scheme discriminates against thinking children' (p18). According to Peacock (2005), Naylor and Keogh suggest that the introduction of the National Curriculum has resulted in the decline of imagination and creativity in the primary classroom, but they do concede the positive result of its introduction being that teachers do now 'expect to teach science' (p. 18). Keogh recognises that removal of the science tests may cause some teachers to stop teaching science but cites Scotland where the removal of testing does not appear to have had a negative impact. She suggests that removal of testing may 'free up people's thinking and allow them to be more adventurous' (p. 19).

In contrast, Daws and Singh (1999) describe some research in which teachers made use of tests to develop formative work. Pupils in a secondary school were invited to look at test papers that they had taken in previous lessons and to highlight 'key words'. The teacher then took the pupils through the mark scheme for each question inviting pupils to interject and question the reasoning behind the 'official' answers. Pupils were also invited to explain alternative approaches to answering the questions and to discuss responses. The teacher involved believed that this method of looking at the test questions helped pupils 'to understand ... the ways in which examiners pose questions and expect them to be answered' (p. 72) and that it also increased the pupils' 'sense of control over the assessment process and confidence in tackling exams' (p. 72). Pupils involved in this work reported that they felt confident when providing their answers and explaining their reasoning because they knew 'neither condemnation not ridicule would ever result. Rather a frank, supportive a diagnostic discussion would occur, because all pupils realised this was in everyone's interest' (p. 72). This highlights the value of the activity and the need for the development of an appropriate climate of trust and respect. The teacher involved felt confident that pupils understood the purpose of the activity as being to 'improve their understanding of topics and problem-solving strategies...to maximise their *future* exam performance' (p. 73) thus using assessment formatively in order to develop their future learning.

Black and Harrison (2001b) also describe ways in which teachers at a secondary school used summative assessments to develop formative work. They report on two teachers who use completed test papers to encourage pupils to compare answers and to devise their own mark schemes based on a group consensus and understanding of the question asked. Although this posed some difficulty for lower ability pupils, the teachers report that more able pupils are able to realise that 'the best answer is not necessarily the majority one, but the one that can be best justified' (p. 46). Another method of incorporating formative and summative assessment is to encourage pupils to discuss a topic (i.e. one that they are revising for a test) and to write questions based on it. These questions are then made available for other students to answer with the original authors responsible for providing the correct answers. These methods encourage peer discussion and critical thinking about the types of questions that might be asked in summative tests and to consider how they could be answered.

3.2.7 Review of materials and teaching tools

Several articles described successful use of particular AfL tools and techniques or specific published materials designed to support AfL in science.

In the section on peer- and self-assessment, the work of Stow (1997) was described in relation to concept mapping being used for self-assessment purposes. The use of concept maps is also explored by Atkinson and Bannister (1998) in a study which sought to compare the use of concept maps and annotated drawings. Concept maps are recognised as allowing for links to be shown between different concepts and for misconceptions to be identified. Annotated drawings have been identified as offering the 'the opportunity for an alternative form of expression to children who may well hold ideas, but who find it difficult to express them in words or to recognise the links between them' (p. 3). The small-scale study focused on the use of these two assessment methods with key stage 1 children. At the start of a new topic, a number of key words were identified and the children were invited to produce a concept map. They were also directed to produce a drawing relating to the topic, with annotations added by an adult. Scores were then allocated to each method according to the inclusion of key words, accuracy of annotation and so on. Atkinson and Bannister identify two patterns in the data, the first being that the older the child the more 'correct ideas' they are able to represent. This finding was true for either method. The second finding indicated that as the children's age increases, concept maps become a more useful tool for allowing children to show their understanding. An additional finding was that more able children, regardless of age, tended to do better using concept maps to show their understanding. Conversely, less able children tended to score more highly using annotated drawings. Although this is a very limited scale study, Atkinson and Bannister conclude that concept mapping can be a very useful assessment tool but note that there should not be over-reliance on one method of assessment as 'children can demonstrate varying levels of understanding when asked to respond in different modes' (p. 5).

Articles describing the research of Keogh and Naylor and their associated development of materials to support AfL are fairly prevalent and the development of concept cartoons is discussed explicitly in section 2.2.2 on Misconceptions and classroom climate. In their articles Keogh and Naylor (1998, 2004, 2007) generally describe how particular tools and techniques can be used to promote AfL and make reference to things such as the development of an appropriate classroom climate, development of talk, the value of revealing understanding and misconceptions and so on. In one article, Keogh and Naylor (1998) describe the ways in which they updated their work in developing concept cartoon following feedback from teachers. As a results of working with children using their earlier versions of the concept cartoons, Keogh and Naylor identified a number of changes that were needed including the need to provide a range of alternative views rather than just one, the need to shift from negative to positive statements and ensuring that a scientifically acceptable viewpoint is included in the options. Feedback from teachers, according to Keogh and Naylor, indicates that concept cartoons provide good access to children's ideas, even from those who are usually reluctant to reveal their thinking. They indicate that there is evidence that concept cartoons can promote conceptual change, promote discussion and development of ideas and can provide an accessible and motivating start point for a science activity. One of the main successes of the concept cartoons is their usefulness for children with learning difficulties due to the minimal text demands and visual impact. Although the article can be viewed as a promotional device, it does demonstrate that the materials are developed around the principles of AfL with children's learning in science being at the heart of the work that they do. Further examples of this occur in other articles, such as Keogh and Naylor (2004), in which they reiterate the value of their materials by commenting that 'finding out and taking children's ideas into account has an important place in primary science' (p. 18) which they believe can be achieved through the use of concept cartoons. As the result of running a three year project based on 'Active Assessment' which involved about six thousand teachers in professional development courses, Naylor and Keogh (2007) have reflected upon the principles underpinning their project and what they have learnt. Through their work they have attempted to 'describe the connections between thinking, learning and assessment and familiarise teachers with a wide range of Active Assessment strategies' (p. 73). In summary, they refer to approaches that 'actively involve learners in purposeful assessment activities and potentially result in further learning' (p. 73). Within the article, Naylor and Keogh demonstrate how their materials embed the background principles of AfL and refer to much of the research that has gone on in this area. They describe, for example, how the materials and their continuing professional development (CPD) training help teachers to make use of peer- and self-assessment techniques, questioning, subject specific assessment examples and consider ways to personalise learning. They conclude that it is an unrealistic task to expect teachers to be able to differentiate all tasks but that by 'involving pupils more in understanding and negotiating learning targets helps to make the process more manageable as they take responsibility for their own learning' (p. 78).

Nott and Suckling (2003) describe how a summative assessment tool can be used for AfL purposes in their article about *Testbase*. They describe a project, entitled '*Testbase: Assessment for Learning*' (TAfL), in which the national key stage 3 tests and associated reports produced on CD-ROM (which form the *Testbase* materials) were used to develop formative assessment items. The study reports that the writing and trialling of 'TAfL' items was undertaken by a relatively small number of people and were mostly trialled by year 7 and 8 pupils. Trial schools were provided with a list detailing the characteristics of formative assessment. There were also five main methods which were suggested for developing TAfL items and AfL ways of working – no marks (the use of traffic lighting or comment only marking), no hands (ask questions directly to pupils by name), show and tell (use of whiteboards to display and answer), questions (encouraging comparisons of a range of responses) and wait time (teachers do not seek an immediate response). Item writers were encouraged to choose a test question and to then adapt it to be used in a formative way, for example, asking pupils to re-write the mark scheme in a child-friendly way, encouraging pupils to traffic light their understanding of particularly words or concepts as a question is discussed and so on. These items were sent to teachers to adapt as appropriate and to then trial the items and provide feedback. Nott and Suckling report that 'pupils were motivated by the TAfL items', that they were 'intellectually challenging and seen as worthwhile by the pupils and the teachers' and that 'Contrary to research that indicates that teachers are conservative ... the TAfL project shows that teachers have not lost their professional judgement and creativity about what is worthwhile and how to develop and teach it' (p. 63). The aim of the project to 'close the loop between the original items and the data and summaries of pupils' performance' (p. 66) was seen to be met by this method of intervention and it was noted that some teachers are continuing to re-write and adapt their own TAfL items. Indeed, advice on how to write TAfL items is provided on the ***Testbase*** website.

Cheung (2006) reports on the development of a 'technology assisted formative assessment tool'. The research is based on chemistry teaching in secondary schools in Hong Kong. Cheung explains the importance of formative assessment for good chemistry teaching with regard to the need to provide feedback to students on current achievement and to the teacher in terms of their teacher effectiveness. The article details the development of a *Windows*-compatible computer system called Test Construction Support System (TCSS), a web based application that can be used by pupils and which stores information for teachers. The aim was to provide a tool for teachers to source pupil misconceptions and to construct their own tests. At the time of the article, there were three topics covered by TCSS. Pupils were required to do some multiple-choice self-assessment exercises at their own pace at the end of a period of study. There are two modes – self-learning and quiz mode – and these provide the pupils with different options. In the self-learning mode feedback is given to the pupil following the completion of each question – either positive reinforcement of a correct answer or advice to look at the 'instant feedback' and attempt the question for a second time. A 'hints' button is also available. Pupils answer between ten and sixteen questions in this format and they are thought to aid self-assessment. In 'quiz mode' pupils answer a series of items which are then computer scored. Information, based on statistical analysis of the items, is then made available to the teacher providing details about which questions were correctly answered and highlighting any misconceptions that the pupil might have. There is also the option for the teacher to construct their own test from the item bank. TCSS also offers a 'students' misconceptions database' which provides teachers with common misconceptions in the three topic areas. Cheung reports that teachers involved in a trial of TCSS found it to be useful for 'providing teachers with good examples of formative assessment' (p. 66) as well as allowing for identification of pupil misconceptions. Similarly, pupils in the trial reported that they 'gained a better understanding of the chemical concepts after attempting the multiple-choice questions' and that the feedback, hints and explanation functions were useful in the self-learning mode.

Keeley *et al*. (2005) report on an American initiative to use formative assessment probes. They explain the, often frequent, issue of pupils coming to a science topic with pre-conceived ideas and the teacher not tapping into these before the teaching of the topic. The role of the probes is to try and ascertain the pupils' understanding of the topic before teaching begins and to help identify misconceptions. This is achieved via a two-tiered assessment probe – one tier which consists of multiple choice questions which provide common misconceptions as the distractors and the second tier which requires pupils to answer an open response question which requires them to explain their thinking. Keeley *et al.* explain that the probes can provide a 'quick class snapshot' and that the results can be quickly analysed and 'used to design instruction using strategies that explicitly target their students' ideas and guide them through a conceptual change' (p. 18) and go on to state that the probes can be used before and during a topic in order to 'help make students' thinking visible to themselves, their peers and their teacher....[providing] feedback that can guide modification and refinement in thinking' (p. 18). The article emphasises the role of the probes in allowing for teaching to be adjusted according to the demands of the needs of the class and pupils within it and whilst the article explains that there are (or will be) published probes available, Keeley *et al.* state that teachers can design probes themselves in order to tailor them for their own use.

Although the articles relating to particular tools are mostly 'self-promotional', they do highlight ways in which particular resources can be developed to support particular aspects of AfL. It should be borne in mind that none of these articles include an objective review of the products.

4 Summary of findings from the literature review

Although assessment for learning has many generic features which are applicable to a wide range of subjects and key stages, there are some features of AfL which can be specifically honed for science teaching and learning. Research in this area focuses on a range of strategies for use within the classroom and much is made of the need to develop an appropriate classroom climate in which pupils are able to express their ideas in a non-threatening environment. The main thrust of AfL in science is the need to find out what pupils know, what they don't know and what they partly know – their misconceptions – and to develop teaching that will move their understanding on. The literature explores the need for a range of questioning, the importance of talk and discussion and the provision of appropriate feedback, all of which can involve and contribute towards self- and peer-assessment. The use of specific tools such as concept maps and concept cartoons can also be used to draw out pupil understandings – and, perhaps more importantly, misunderstandings.

5 References

Asoko, H. and Scott, P. (2006). 'Talk in science classrooms.' In: Harlen, W. (Ed) *ASE Guide to Primary Science Education.* Hatfield: Association for Science Education.

Assessment Reform Group (2002). *Assessment for Learning: 10 Principles* [online]. Available: http://www.qca.org.uk/libraryAssets/media/4031_afl_principles.pdf [10 July, 2009].

Atkinson, H. and Bannister, S. (1998). 'Concept maps and annotated drawings: a comparative study of two assessment tools', *Primary Science Review*, **51**, 3–5.

Black, P. and Harrison, C. (2001a). 'Feedback in questioning and marking: the science teacher's role in formative assessment', *School Science Review*, **82**, 301, 55–61.

Black, P. and Harrison, C. (2001b). 'Self- and peer–assessment and taking responsibility: the science student's role in formative assessment', *School Science Review*, **83**, 302, 43–49.

Black, P. and Harrison, C. (2004). *Science Inside the Black Box.* London: nferNelson.

Black, P. and Wiliam, D. (1998). *Inside the Black Box*. London: nferNelson.

Budd Rowe, M. (1974). 'Relation of wait-time and rewards to the development of language, logic and fate control: part II', *Journal of Research in Science Teaching*, **11**, 4, 291–308. Cited in Harlen, W. (2006d). *Teaching, Learning and Assessing Science 5–12*. 4th edn. London: Sage Publications.

Butler, R. (1987). 'Task–involving and ego–involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and performance', *Journal of Educational Psychology*, **79**, 4, 474–482. Cited in: Harlen, W. (2006d). *Teaching, Learning and Assessing Science 5–12*. 4th edn. London: Sage Publications.

Butler, R. (1988). 'Enhancing and undermining intrinsic motivation; the effects of task–involving and ego–involving evaluation on interest and performance', *British Journal of Educational Psychology*, **58**, 1–14. Cited in: Black, P. and Harrison, C. (2001a). 'Feedback in questioning and marking: the science teacher's role in formative assessment', *School Science Review*, **82**, 301, 55–61.

Cheung, D. (2006). 'TCSS: a new computer system for developing formative assessment', *School Science Review*, **88**, 322, 61–70.

Chin, C. (2004). 'Students' questions: fostering a culture of inquisitiveness in science classrooms', *School Science Review*, **86**, 314, 107–112.

Chin, C. (2006). 'Classroom interaction in science: teacher questioning and feedback to students' responses', *International Journal of Science Education*, **28**, 11, 1315–1346.

Daws, N. and Singh, B. (1999). 'Formative assessment strategies in secondary science', *School Science Review*, **80**, 293, 71–78.

Department for Children Schools and Families (2008). *The Assessment for Learning Strategy*. London: DCSF [online]. Available: www.teachernet.gov.uk/publications [3 June, 2008].

Evans, N. (2001). 'Thoughts on assessment and marking', *Primary Science Review*, **68**, 24–26.

Gioka, O. (2006). 'Assessment for learning in physics investigations: assessment criteria, questions and feedback in marking', *Physics Education*, **41**, 4, 341–346.

Hargreaves, E. (2007). 'The validity of collaborative assessment for learning', *Assessment in Education: Principles, Policy & Practice*, **14**, 2, 185-199.

Harlen, W. (1999). 'Purposes and procedures for assessing science process skills'', *Assessment in Education*, **6**, 1, 129–144.

Harlen, W. (2004). 'Talking and writing: have we got the balance right?', *Primary Science Review*, **83**, 17–19.

Harlen, W. (Ed) (2006a). *ASE Guide to Primary Science Education.* Hatfield: Association for Science Education.

Harlen, W. (2006b). 'Assessment for learning and assessment of learning.' In: Harlen, W. (Ed) *ASE Guide to Primary Science.* Hatfield: Association for Science Education.

Harlen, W. (2006c). 'Teachers' and children's questioning.' In: Harlen, W. (Ed) *ASE Guide to Primary Science.* Hatfield: Association for Science Education.

Harlen, W. (2006d). *Teaching, Learning and Assessing Science 5–12*. 4th edn. London: Sage Publications.

Harlen, W. (2007). 'Holding up a mirror to classroom practice', *Primary Science Review*, **100**, 29–31.

Harrison, C., Drozdowskij, J. and Westhead, K. (2001). 'Formative assessment in primary classrooms', *Primary Science Review*, **68**, 19–22.

Harrison, C. and Harlen, W. (2006). 'Children's self– and peer–assessment.' In: Harlen, W. (Ed) *ASE Guide to Primary Science Education.* Hatfield: Association for Science Education.

Keeley, P., Eberle, F. and Farrin, L. (2005). 'Formative assessment probes: uncovering students' ideas in science', *Science Scope*, **28**, 4, 18–21.

Keogh, B. and Naylor, S. (1998). 'Teaching and learning in science using Concept Cartoons', *Primary Science Review*, **51**, 14–16.

Keogh, B. and Naylor, S. (2004). 'Children's ideas children's feelings', *Primary Science Review*, **82**, 18–20.

Keogh, B. and Naylor, S. (2007). 'Talking and thinking in science', *School Science Review*, **88**, 324, 85–90.

Khwaja, C. and Saxton, J. (2001). 'It all depends on the question you ask', *Primary Science Review*, **68**, 13–14.

Leakey, A. (2001). 'Fantastic feedback', *Primary Science Review*, **68**, 22–23.

Lindsay, C. and Clarke, S. (2001). 'Enhancing primary science through self– and paired–assessment', *Primary Science Review*, **68**, 15–18.

Macro, C. and McFall, D. (2004). 'Questions and questioning: working with young children', *Primary Science Review*, **83**, 4–6.

Markwick, A., Jackson, A. and Hull, C. (2003). 'Improving learning using formative marking and interview assessment techniques', *School Science Review*, **85**, 311, 49–55.

McNair, S. (2004). '"A" is for assessment', *Science and Children*, **42**, 1, 18–21.

Mercer, N., Dawes, L., Wegerif, R. and Sams, C. (2004). 'Reasoning as a scientist: ways of helping children to use language to learn science', *British Educational Research Journal*, **30**, 3, 359–377. Cited in: Asoko, H. and Scott, P. (2006). 'Talk in science classrooms.' In: Harlen, W. (Ed) *ASE Guide to Primary Science Education.* Hatfield: Association for Science Education.

Millar, R. and Hames, V. (2002). 'EPSE Project 1: Using diagnostic assessment to improve science teaching and learning', *School Science Review*, **84**, 307, 21–24.

Naylor, S. and Keogh, B. with Goldsworthy, A. (2004). *Active Assessment: Thinking, Learning and Assessment in Science*: Sandbach: Millgate House Publishers.

Naylor, S. and Keogh, B. (2007). 'Active assessment: thinking, learning and assessment in science', *School Science Review*, **88**, 325, 73–79.

Newton, P., Driver, R. and Osbourne, J. (1999). 'The place of argumentation in the pedagogy of school science', *International Journal of Science Education*, **21**, 5, 553–576. Cited in: Keogh, B. and Naylor, S. (2007). 'Talking and thinking in science', *School Science Review*, **88**, 324, 85–90.

Nott, M. and Suckling, S. (2003). 'Using Testbase for assessment for learning – the TAfL project', *School Science Review*, **85**, 311, 57–67.

Peacock, A. (2005). 'The PSR interview: simple, purposeful and problematic', *Primary Science Review*, **89**, 17–19.

Qualter, A. and Taylor, J. (1999). "Brilliant Erin! Brilliant!': setting a risk–taking climate', *Primary Science Review*, **58**, 22–23.

Sato, M., Coffey, J. and Moorthy, S. (2005). 'Two teachers making assessment for learning their own', *The Curriculum Journal*, **16**, 2, 177–191.

Stark, R. and Gray, D. (2001). 'What science do they think they know?', *Primary Science Review*, **68** , 8–11.

Stow, W. (1997). 'Concept mapping: a tool for self–assessment', *Primary Science Review*, **49**, 12–15.

Van Zee, E., Iwasyk, M., Kurose, A., Simpson, D. and Wild, J. (2001). 'Student and teacher questioning during conversations about science', *Journal of Research in Science Teaching*, **38**, 2, 159–190.

Watkins, C. with Carnell, E., Lodge, C., Wagner, P. and Whalley, C. (2001). ‘Learning about Learning enhances performance’, *NSIN Research Matters*, **13**

Wells, G. (2008). ‘Dialogue, inquiry and the construction of Learning Communities.’ In: Lingard, J., Nixon, J. and Ranson, S. (Eds) *Transforming Learning in Schools and Communities: the Remaking of Education for a Cosmopolitan Society*. London: Continuum.

Additional Reading

Bell, B. and Cowie, B. (2001). 'The characteristics of formative assessment in science education', *Science Education*, **85**, 536–553.

Black, P. (1998). 'Formative assessment: raising standards inside the classroom', *School Science Review*, **80**, 291, 39–46.

Black, P. (1998). *Testing: Friend Or Foe*. London: Falmer Press.

Black, P. and Wiliam, D. (1998). 'Assessment and classroom learning', *Assessment in Education*, **5**, 1, 7–73.

Cowie, B. and Bell, B. (1999). 'A model of formative assessment in science education', *Assessment in Education*, **6**, 1, 101–116.

Spavold, Z. (2005). 'Using formative assessment to raise pupil motivation: a small classroom–based study', *School Science Review*, **86**, 317, 119–123.

Appendix 1: Search strategy for literature review

Search strategy

Alongside this desk review of resources was a systematic literature review of research into assessment for learning in science, using a range of different educational databases. Search strategies were developed using terms from the relevant thesauri (where available) and/or free-text searching. The same search strategies were adhered to as far as possible for all the databases.

The key words used in the searches, together with a brief description of each of the databases searched, are outlined below. Throughout, (ft) has been used to denote free-text search terms. The searches were defined by the following criteria:

- Material published between 1997 and 2008
- Focusing on key stage 2, but including other key stages where appropriate
- Research carried out in the UK primarily, but also extending to Australia, New Zealand and the USA

Australian Education Index (AEI)

AEI is produced by the Australian Council for Educational Research. It is an index to materials at all levels of education and related fields. Source documents include journal articles, monographs, research reports, theses, conference papers, legislation, parliamentary debates and newspaper articles.

#1	Science education	#10	Assessment criteria (ft)
#2	Science curriculum	#11	Learning outcomes (ft)
#3	#1 or #2	#12	Learning objectives (ft)
#4	Assessment for learning (ft)	#13	Self evaluation
#5	AfL (ft)	#14	Self monitoring (ft)
#6	Formative evaluation	#15	Self regulation (ft)
#7	Rich questions (ft)	#16	#4 or #5… #15
#8	Self assessment (ft)	#17	#3 and #16
#9	Peer assessment		

British Education Index (BEI)

BEI provides bibliographic references to 350 British and selected European English-language periodicals in the field of education and training, plus developing coverage of national report and conference literature.

#1	Science education	#14	AfL (ft)
#2	Science curriculum	#15	Formative evaluation
#3	Science teaching	#16	Rich questions (ft)
#4	Science tests	#17	Self assessment (ft)
#5	Practical science	#18	Peer assessment
#6	Scientific literacy	#19	Assessment criteria (ft)
#7	Chemistry education	#20	Learning outcomes (ft)
#8	Physics education	#21	Learning objectives (ft)
#9	Science experiments	#22	Self evaluation
#10	Science clubs	#23	Self monitoring (ft)
#11	General science	#24	Self regulation (ft)
#12	#1 or #2 or…..#11	#25	#13 or #14 or…#24
#13	Assessment for learning (ft)	#26	#12 and #25

The Educational Resources Information Center (ERIC)

ERIC is sponsored by the United States Department of Education and is the largest education database in the world. It indexes over 725 periodicals and currently contains more than 7,000,000 records. Coverage includes research documents, journal articles, technical reports, program descriptions and evaluations and curricula material.

#1	Science education	#10	Assessment criteria (ft)
#2	Science curriculum	#11	Learning outcomes (ft)
#3	#1 or #2	#12	Learning objectives (ft)
#4	Assessment for learning (ft)	#13	Self evaluation
#5	AfL (ft)	#14	Self monitoring (ft)
#6	Formative evaluation	#15	Self regulation (ft)
#7	Rich questions (ft)	#16	#4 or #5… #15
#8	Self assessment (ft)	#17	#3 and #16
#9	Peer assessment		

Author searches

Black, P.

Wiliam, D.

Handsearches

Primary Science Review, School Science Review and *Science Education*, as key journals in the field, were handsearched to identify articles that had been missed in the database searches.

www.nfer.ac.uk

National Foundation for Educational Research
The Mere Upton Park
Slough Berks SL1 2DQ
Tel: 01753 574123 Fax: 01753 691632
Email: enquiries@nfer.ac.uk