An Evaluation of Formative Assessment Probes in a Solution Chemistry Teaching Sequence.

Mark McDonagh

MA (by Research)

University of York

Education

December 2014

Abstract

This study aimed to evaluate the use of diagnostic questions as formative assessment probes in science teaching. Formative assessment practice is still not well developed in most schools. Much practice depends on the formative use of summative tests. In addition, the National Curriculum level descriptors provide an imprecise notion of progression through many science concepts.

The study sought to use more research-based materials to improve the author's formative assessment practice by carrying out an evaluation of a single action research cycle. A backward design approach was used to generate a series of 31 diagnostic questions which could assess understanding in a KS3 solution chemistry topic in a teaching sequence building on progressively more demanding concepts.

The materials were used in the author's school with Year 8 students (n=60). In a 3-4 week teaching programme.

The project provides evidence that the diagnostic questions developed are a highly effective means of gathering rich data of student understanding in a fashion which allows the reshaping of teaching during a normal timetable. The probes were sufficiently sensitive to detect misconceptions in chemistry and provided the teacher with a secure basis for deciding when to repeat the teaching or extend the concepts being taught.

The progression in the teaching sequence was analysed using the frameworks of Bloom and Piaget. Students' performance was shown to be linked to the level of conceptual demand. Evidence was presented which suggested that diagnostic questions could be used as formative assessment probes with summative potential.

Table of Contents

List of Tables

List of figures

Acknowledgements

The author would like to thank Professor Robin Millar for his incisive advice and feedback and generous support.
Thanks also go to Dr Jeremy Airey who stepped in late but without whom the dissertation would never have been finished on time.

Thanks also to the students of St Bede's who were willing co-researchers throughout the project.

And finally, Emma, Edward and William, thank you for your love and forbearance: *Normal service will be resumed shortly.*

Author's declaration

All research presented in this thesis was initiated and conducted by the author with guidance from The University of York. The author is responsible for the research presented in this thesis. No part of this work has been previously published or submitted for any other award at any other institute. All sources are acknowledged as References.

Chapter 1

Introduction

1.1- Background to this study

This research is an evaluation of an approach to improve formative assessment in my own teaching. It is an action research project and involves the development and implementation of classroom materials in a secondary science setting. The aim of the study is to evaluate the extent to which the developed materials can be used formatively and summatively during a series of Key Stage 3 science lessons on solution chemistry.

Wiliam and Black (1996) cite Bloom (1971) and Scriven (1967) in formulating a definition of summative and formative assessment (p. 537). Summative assessment is the judgement made at the end of a period of study to measure the extent of a student's learning of the material covered. Formative assessment is the use of information about the learning to help students improve. This latter type of assessment is less well understood by many teachers (Harrison, 2011; Stewart, 2012).

In many schools, student progress in Key Stage 3 is reported (summatively) using National Curriculum (NC) levels. This normally follows tests or examinations. However, practice has emerged over recent years where reference to ‘levelness’ during teaching has been advocated to improve student learning. I have struggled to use levels formatively in this way. The current study is an attempt to address this difficulty through the development of a research-based approach to formative assessment in my own teaching.

In this introduction I would like to present a background to the current study. This includes a brief description of how NC levels were introduced, how educational policy has linked their summative use to school accountability and how attempts have been made to use NC levels as part of formative assessment.

1.2 - Educational reform and the introduction of levels.

The Education Reform act in 1988 brought radical change to the education system in the UK. A new National Curriculum was introduced together with a national system of assessing pupils in England and Wales. A working party chaired by Paul Black known as the Task Group on Assessment and Testing (TGAT) were given the task of developing the assessments to be used in the new system (Black, 1997). The assessment needed to serve formative and summative purposes. On the one hand, information from assessments was needed for teachers and pupils to provide feedback and improve learning. While on the other hand the assessments were expected to provide data on school performance for a wider public.

TGAT set out four principles to guide the design of the new national system (Daugherty 1995):

- assessments should provide information about a pupil’s achievements in relation to objectives (assessments should be criterion referenced).
- results from assessments should be used to make decisions about a pupil’s further learning (assessments should be formative).

- there should be comparison between classes and schools to share common standards (assessments should be moderated).
- assessments should relate to the expected route through educational development (assessments should relate to progression).

The group saw a way forward in combining formative and summative assessments through using teacher assessments together with external tests. The formative aspect would come from proposed Standard Assessment Tasks (SATs) which would be carried out by teachers. The tasks were to be classroom activities in which pupils could demonstrate performance in a range of subject areas. The subject areas were known as attainment targets: the science curriculum was divided into 22 attainment targets. A set of ten 'levels of attainment' was proposed in each attainment target to profile the progression of pupils through their eleven years of schooling.

The complexity of the proposals - in science there were 354 criterion-referenced statements grouped into ten levels in twenty-two attainment targets - were to prove their undoing. In 1991 the tasks were described by the then Minister for Education, Ken Clarke, as 'elaborate nonsense' and abandoned in favour of written terminal exams which were 'manageable, reliable and summative' (Black, 1997). The formative tasks became summative tests. The 'nightmare' that Black had feared: 'narrowly based external test which would give misleading results whilst constraining teaching' (Black, 1997) (p. 27) seemed to have been realised.

1.3 - Summative assessment and National Curriculum levels

From this point, SATs testing in science at ages seven, eleven and fourteen became the experience of all school children in the UK. The test results were used to track pupils' progress through the first three Key Stages of education with GCSEs forming the final assessment of compulsory schooling. By the end of Key Stage 2 it was envisaged that most children would achieve level 4 and by the end of Key Stage 3 most would achieve level 5 or 6 (DfEE/QCA, 1999) (p. 7).

The use of levels to report on school accountability brought with it unintended consequences. Harlen and Deakin Crick (2003) also found that teachers focussed on the content of the test and used a transmission model of teaching. Teaching which focussed on developing skills was less frequent than that which sought to pass on content knowledge. Classroom time was spent practising tests which led to students becoming demotivated about their learning.

The effect on student thinking has also been discussed in terms of the depth of understanding learners achieve. Harlen & James, (1997), describe deep learning as learning which is actively understood and internalised by the learner and helps to make sense of the world and their experience of it. On the other hand, surface learning is the rote learning of facts which can be reproduced as required. These authors maintain that a focus on summative testing is more likely to result in surface learning taking place, (Harlen & James, 1997).

Children tended to label themselves with the level they achieved. Raey and Wiliam (1999) described how one pupil felt that if she did not perform well enough in the Key Stage 2 SATs to achieve a level 4 or 5 she would be ‘a nothing’ (p349). This could have a negative impact on the motivation of such children. But even children who feel they are successful can ‘become more concerned for ‘what level they are’ than the substance of what they know, can do and understand’ (DfE, 2011) (p. 44).

There were also wider issues with tests. Newton (2009) cast doubt on the reliability of the results from the SATs across the time period from 1996 to 2007. He concluded that results had been misclassified, and a substantial number of students would receive different levels if the test process was repeated. Although no system can be 100% reliable, Newton suggested that a debate was overdue as to how much level of error was acceptable in tests such as these.

Taber (2009) questioned whether the tests could measure understanding in science. Questions designed to assess recall would demonstrate how much factual knowledge had been leaned. However, a question designed to measure application of knowledge depended on how familiar the student was with the question setting. This, Taber suggested was not possible for the test to discern. Ultimately he thought the process of awarding a single overall level as a reflection of an individual’s capability and understanding in science as a nonsense.

Fairbrother (2008) criticised the KS3 Science SATs for lacking validity after examining the types of questions and the coverage of the science curriculum used in the tests over a four year period. He pointed to 90% of marks for the paper being awarded through questions where a single mark was available. Using the Bloom Taxonomy, he argued that this only allowed factual knowledge and comprehension to be assessed. Only 2% of marks were allocated to higher categories such analysis, synthesis and evaluation. Assessment of understanding was prevented by the lack of opportunities for pupils to write at length. Coverage of the science curriculum was not equally weighted and that questions based on scientific enquiry and investigative.

Tymms (2004) challenged the claim that Key Stage 2 test scores from UK primary schools indicated an improvement in standards. The increases in attainment, Tymms suggests, were more likely down to ‘children becoming more adept at taking tests as schools taught test technique.’ Wiliam (2001) made the point that the increased pressure to increase a student’s performance led to teachers spending more time teaching those aspects which would be tested. This gave the appearance that achievement was increasing, while in fact overall achievement across the whole subject may be falling.

Linking test and examination data to school accountability was a central policy to raise ‘standards’. Coe (2013) asserts that the evidence that levels of education have improved over the last 30 years is unconvincing. He cites evidence from international surveys and independent studies to suggest that the three-fold increase in pupils obtaining five GCSEs at grade C or above from 1987 to 2012 is not down to improvements in schools alone. Coe also points out that no rigorous collection of high quality data on attainment in schools has been carried out.

In 2008, Ed Balls unexpectedly announced the end of key stage 3 tests: GCSEs were now to provide the only summative measure of progress in secondary schools. At about the same time the Government’s Assessment for Learning Strategy was launched. This was an attempt to improve formative assessment in schools which inspection data revealed was poorly developed (Ofsted 2008).

1.4 - Formative assessment and the National Curriculum

Black and Wiliam (1998a) published a paper entitled ‘Inside the Black Box: raising standards through classroom assessment (see section 2.3.2). Black and Wiliam described formative assessment as essentially any activity which provided students with information about their performance during instruction so that they could make improvements. This process required that the learning aims were clear, teachers made decisions about student progress and that feedback was given to students to move them forward.

Looking at developments over the past decade it seems to me that the central message contained in the ‘Black Box’ paper was misinterpreted in two ways.

1.4.1 Misinterpretation 1: Using summative tests formatively

TGAT saw the possibility of teachers using information from their own ongoing assessments to make decisions about a pupil’s further learning. Teacher assessment with group moderation formed an essential feature of the original formulation of the 1988 Reform Act (Black 1997). Wiliam (2001) argued that ‘teacher’s own assessments of students are highly reliable because they are based on hundreds of hours of assessment’ (p. 19).

During the early part of the National Strategy for Learning in 2002, training was given to teachers on how to conduct formative assessment. A statement from the training is given below:

- *Teachers use evidence from testing to decide what pupils achieved against their target and plan how to take the next step.* (National Stem Centre, 2014) (slide 1.6)

Despite being described as formative assessment the emphasis was on using summative testing. The interpretation was that teachers would use test data to build a picture of areas within a topic where student knowledge was weak. This was followed by repeated teaching or revision sessions – referred to as intervention or second wave teaching.

Further training materials (DCSF, 2007) required teachers to develop 'script analysis' (p. 7). This entailed spending considerable time entering the individual question scores from student tests onto a spreadsheet to build an overview of where student understanding was weak. This was particularly recommended following 'mock' SATs. These weaknesses were then to be addressed in any teaching which followed to improve the scores for the summer Key Stage 3 results. These teaching 'improvements' were sometimes even applied to other student cohorts.

The Assessment for Learning Strategy (DCSF, 2008) also suffered a similar misinterpretation. (Swaffield, 2011). Two aims from the strategy document (DCSF, 2008) are given below:

- ***every teacher** is equipped to make well founded judgements about pupils' attainment, understands the concepts and principles of progression, and knows how to use their assessment judgements to forward plan, particularly for those pupils who are not fulfilling their full potential*
- ***every school** has in place structured and systematic assessment systems for making regular, manageable and accurate assessments, and for tracking their progress.* (p. 4)

Reading these objectives it was easy to see how they would be translated by many teachers into regular tests. Swaffield quite justifiably feared that not only would the Assessment for Learning (AfL) strategy distort pupils' and teachers' understanding of formative assessment but that it would result in a further wasted opportunity to introduce its practice.

1.4.2 – Misinterpretation 2: AfL using shared learning objectives

The 1999 National Curriculum for science (DfEE/QCA, 1999) contained an appendix of attainment targets. These were statements which set out 'the knowledge, skills, and understanding that pupils of different abilities and maturities are expected to have by the end of each key stage.' (p. 74). These statements were referred to as level descriptors and were published for all subjects. They were used by teachers in deciding the level of attainment of their students at the end of a key stage when national testing was not carried out. This did not apply in science because examinations were to being used to generate levels.

The level descriptors were used in science lessons as learning objectives and were seen as stepping stones to facilitate rapid progress. Coaching to the next level was regarded as 'assisted progression' and was justified because it aimed to help children reach their target level. Because the level descriptors in the National Curriculum document contained quite technical language, more child-friendly versions were written by many schools and local authorities.

Figure 1.1 shows a level ladder of the Solids, Liquids and Gases unit published by Northumberland County Council. Newberry, Gilbert and Hardcastle (2005) published a Levels Mountain (figure 1.2) which gave an overview of how the levels progressed in science from level three to seven. During the National Strategy a consensus grew that if a child was told the difference between what they had to say to be awarded a level 6 over a level 5 this would lead to an improvement in their

My target is:

WHAT LEVEL AM I WORKING AT?

UNIT 7G Particle Model of Solids, Liquids and Gases

Level 3	**Start**	**Check**
I know that some changes are reversible and some are irreversible		
I can identify a solid, liquid and a gas		
Level 4		
I can group materials into solids, liquids and gases		
I can explain why I have grouped them this way		
I know materials are made of particles		
I can explain the arrangement of particles in a solid, liquid and gas		
I can explain reversible and irreversible changes		
Level 5		
I can explain how particles move		
I can relate this to solid, liquid and gas		
I can use the particle model to explain diffusion in gases		
I can use the particle model to explain mixing of liquids		
Level 6		
I can explain the link between particles and energy transfer		
I know the closer the particles, the denser the material		

Figure 1.1 Level ladder from Northumberland County Council

understanding. I cannot find any empirical evidence to support this view. This was not what Black and Wiliam were referring to when they included sharing of learning objectives as a key part of formative assessment.

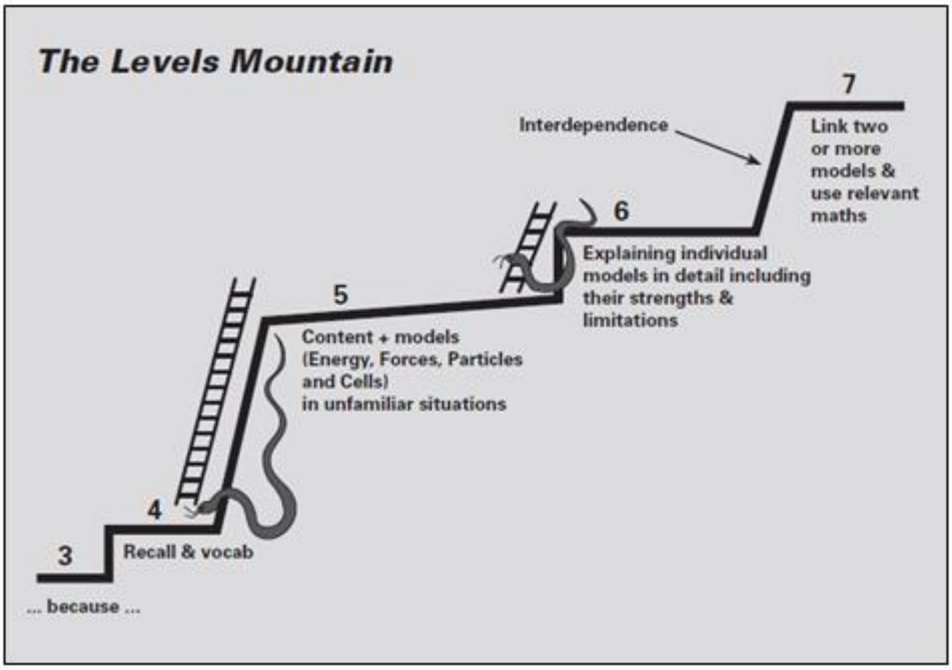


Figure 1.2 The levels mountain (Newberry et al., 2005)

As a classroom practitioner, I struggled to implement formative assessment in my teaching. I used test data to analyse the ‘gaps’ in my students knowledge and understanding and set learning objectives that demonstrated clearly what students had to do to achieve a particular level. But I did not feel that this was bringing the learning gains it should. Ultimately, students were prepared for tests because this was the measure of accountability.

1.5 Carrying out educational research

The research project in this study stems from a desire to investigate formative assessment and to implement it in a fashion more in keeping with the original ideas contained in the ‘Black Box’ paper. In addition I would like to investigate a more useful progression structure currently offered by the National Curriculum levels.

Oversby (2011) recognises that ‘teaching is a research based profession.’ For this reason I think it is important that teachers should take some personal measure of responsibility in addressing issues which they encounter in their own professional practice.

This thesis presents research from a project to evaluate a system of formative assessment using diagnostic question probes. The research was an action research project and involved two classes of Year 8 students in the school where I work. A teaching sequence of lessons on solution chemistry was designed to provide a sequential increase in the cognitive demand of lesson learning intentions. The diagnostic question probes were used to inform decisions about the direction of teaching and to measure progress along the teaching sequence.

1.6 Overview of this dissertation

Chapter two conducts a review of literature of feedback practices and formative assessment. The review also examines the ideas which have been explored for measuring conceptual understanding.

Chapter three lists and describes the research questions. It outlines the reasons for selecting the research strategy and describes the rationale behind the construction of the teaching sequences diagnostic question design. Details of the participants, teaching methods, and data collection are also given.

Chapter four sets out the findings which address the first research question. It presents a series of vignettes which show the use of diagnostic questions provided information to the teacher which allowed the teaching to be modified to provide better learning outcomes.

Chapter five sets out the findings which address the second research question. It presents three case studies of students from the study with an analysis and discussion of their responses. The performance of the students is discussed with reference to the cognitive progression in the teaching sequence.

Chapter six contains the conclusions of the study. The research questions are restated and the main findings summarised. An evaluation of the research methods and strategy are given together with a discussion of the implications for teaching and possible further work.

Chapter 2

Literature Review

2.1 Introduction

The aim of this study is to evaluate a formative assessment approach to teaching a science topic. This chapter explores the recent educational research in relevant areas. Part one will examine the contribution of feedback to assessment and will document the developments introduced by the Assessment Reform Group. The second part will look at a range of models used to measure conceptual understanding which can be used to judge student performance. Finally, student misconceptions in solution chemistry are identified.

2.2 Formative assessment and feedback.

Providing students with feedback 'is one of the most powerful influences on learning' (Hattie & Timperley, 2007) (p. 81). Ramaprasad (1983) defines feedback as:

> '*Information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way.'* (p. 4)

Sadler (1989) used this definition to identify an instructional feedback cycle where a learner:

a) has information about the standard being aimed for,

b) can compare their performance with the standard, and

c) engages with appropriate action to close the gap.

Simply put a teacher needs to tell a student where they are, where they need to get to and how to get there. On the face of it this seems a straightforward process but the reality is that giving feedback is much more complex proposition.

2.2.1 Four levels of feedback

Hattie and Timperley (2007) identified four levels within which feedback was provided to students; feedback on the task, feedback on the process, feedback on self-regulation and feedback on the ego or self. These four types of feedback will be discussed in detail below.

The most effective feedback was that which focussed on the task or the process of achieving a desired goal. To optimise this process, teachers had to state clear learning aims, assess students and assist them towards successful performance. Feedback about performance must refer to how to improve if it is to be effective (Hattie and Timperley, 2007). Giving students grades or numerical score only does not tell students how to improve and does not contribute to further learning (Black, Harrison, Lee, Marshal & Wiliam 2003) (p 46). The timing of feedback has received some attention. It is generally believed that pupils need time to think about a problem. If they are given feedback too early they learn less. (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). And if the feedback is given while they are engaged in the task, it can interrupt and inhibit the learning (Shute, 2008). There is a balance to be struck in the timing of feedback so that it addresses material that is still relevant to the student. Feedback should also be specific and as simple as possible. Shute (2008) reported that vague feedback, caused students to disregard it and complex, lengthy feedback tended to diffuse or dilute the message.

Feedback about self-regulation focused on factors like the commitment and confidence of a student. This had positive outcomes for a student who saw learning as mastering new skills and where mistakes were part of the skill acquisition process. However, less effective learners tended to view performance as something beyond their control and became less motivated if feedback asked them merely to try harder. Feedback directed toward self (praise) was largely ineffective. Kluger and DeNisi (1998) found that comments such as, ‘Well Done!’ or ‘Good girl’ contain little task-related information and could actually have a negative effect on learning. Providing students with normative feedback, that is, feedback which compares the performance of an individual to those of others has been shown to cause those who perform poorly to become demotivated and hence to decrements in achievement (Butler & Nisan, 1986).

2.2.2 Basing feedback on accurate assessment

It is prerequisite for effective feedback that teachers are able to gather reliable data about their students understanding and performance. Some research indicates that this is difficult for some teachers. Schneider and Gowan (2013) found that teachers did not always construct appropriate assessments when measuring student learning and were consequently unable to provide accurate feedback to their students. This observation was also made by Ruiz-Primo and Lui (2013) who examined student science books to investigate teacher feedback practices.

2.2.3 Feedback from the student to the teacher

Several authors (Hattie, 2009; Nicol and Macfarlane-Dick, 2006 & Yorke, 2003) raise a further issue concerning feedback. They point to another and possibly more important role of feedback as information a teacher receives from the student. When teachers learn about the expertise their students are developing, it can help them tailor the teaching accordingly. A teacher who has a strong understanding of where a student is in their understanding can plan instruction and shape teaching to support them towards it.

2.3 The contribution of the Assessment Reform Group

2.3.1 The promise of formative assessment

Paul Black and Dylan Wiliam published a seminal review on ‘Assessment and Classroom Learning’ (Black & Wiliam, 1998b). The review examined the findings from about 250 research papers across a wide range of educational settings and concluded that good formative assessment could lead to ‘substantial learning gains.’ The review was followed by a position paper entitled: Inside the Black Box: raising standards through classroom assessment (Black & Wiliam, 1998a) which was published in both the UK and the US in which the magnitude of the possible gains were quantified. The paper reported that:

> *The formative assessment experiments produce typical effect sizes of between 0.4 and 0.7: such effect sizes are larger than most found for educational interventions.* (Black & Wiliam 1998a) (p.3).

To explain the significance of this, an effect size of 0.4 would be the equivalent of moving student's learning forward by 3.5 months in a year. An effect size of 0.7 would be an improvement of 0.78 months (Higgins et al., 2013). It proved to be a hugely influential publication in the UK and, as discussed in chapter one, was used as a basis for (some slightly misguided) National Strategy training. The research had been commissioned by a UK body, the Assessment Reform Group (ARG) which then sought to pursue the potential which the published work promised. The ARG reported an impoverished culture of assessment in many UK schools where it was used mainly for measurement (ARG 1999). In support of their own view, the ARG referred to the 1998 Ofsted review of secondary education in England 1993-1997:

> '*Overall the purpose of assessment is to improve standards, not merely to measure them. Although the quality of formative assessment has improved perceptibly, it continues to be a weakness in many schools.'*
>
> (ARG, 1999) (p. 4)

2.3.2 The development of the formative assessment 'toolkit'

The ARG (2002) referred to formative assessment as Assessment for Learning (AfL) and this term became widely used. AfL was intended to contrast to assessment *of* learning which involved testing. To make the ideas more accessible the group published of a list of ten principles (Table 2.1).

Black and Wiliam's review (1998a) had reported strong evidence for formative assessment improving learning but reflected the weak practice in classrooms. Black, Wiliam and colleagues began work to gather empirical research evidence to support the claims made in 'Inside the Black Box'. The authors acknowledged that the diverse nature of the material which had been used in the review and hence the claims, might lack 'ecological validity' in real classrooms. Wiliam, Black, Harrison and Black (2004) set up a project to explore novel classroom approaches to assessment. Twenty four maths and science teachers in six Oxford and Medway schools were used to conduct mini-experiments using the new methods.

10 Principles of Assessment for Learning:
• Is part of effective planning; • Focuses on how students learn; • Is central to classroom practice; • Is a key professional skill; • Is sensitive and constructive; • Fosters motivation; • Promotes understanding of goals and criteria; • Helps learners know how to improve; • Develops the capacity for self-assessment; • Recognises all educational achievement. Adapted from ARG (2002)

Table 2.1 Ten principles of assessment for learning

The methods involved in the mini-experiments were based on sound, well established educational principles.

- Questioning – involving wait time (Rowe 1974) to deepen dialogue between all members in the classroom.
- Feedback through marking – drawing on the ideas discussed above, comment-only marking which indicated good work and possible improvements were provided.

- Peer and self-assessment – following Sadler (1989) students were encouraged to gain ‘direct evaluative experience’ (p.134) by assessing the work of other students and applying the same criteria to their own work.

Results from this study, (Wiliam, Lee, Harrison & Black, 2004) found that qualitative aspects of classroom practice improved such as classroom dialogue and teachers’ awareness of a shift in their role to allow students to be more active in their own assessment. The quantitative measure, an effect size increase of only 0.35 was lower than predicted. The researchers pointed out, however, that this effect size, if applied across the school, would still raise its ranking into a higher quintile.

A second report presented a deeper examination of evidence collected on two of the original twenty four teachers (Lee & Wiliam, 2005). This work dug down into the practical aspects of the teacher’ work in the classroom using formative assessment and began to identify an issue with professional development. This conclusion addressed another of the original ideas made from Black and Wiliam’s 1998 review. Formative assessment techniques were a means to achieving the deep change needed in current *teacher* practice to improve assessment and learning for all teachers.

By 2005, Wiliam had moved to the US where his contributions continued in the face of the testing regime of the US government’s ‘No Child Left Behind’ policy. Just as in the UK, regular summative testing was seen as the only way to raise standards of education.

A group based at the Educational Testing Service in the US, (Leahy, Lyon, Thompson & Wiliam, 2005) continued developing the techniques begun in the

Oxford-Medway study. These low-tech, low-cost techniques formed the core practices of formative assessment, such as:

- Model good work by students from a previous year to provide an example of what a good performance looked like.
- Give students red, amber and green colour swatches (or traffic lights) and ask them to hold up a colour corresponding to their level of confidence in understanding the learning target.
- Ask 'hinge' questions before to check understanding before moving on from a topic.
- Use a policy of no-hands- up. This prevented students opting out of discussion.
- Use of mini-white boards allowed quick question and answer as a whole class rather than just an individual student.

Wiliam (2011) provides a full list of 53 techniques (p163). The techniques comprised a 'toolkit' for classroom.

2.3.3 The challenges to formative assessment

The learning gains promised by the use of formative assessment techniques were slow to materialise if at all (Ofsted, 2008).

Marsh (2007) identified the pressures of external testing and headteachers' demands impeded the change needed to change teachers' practices. Harlen and Deakin-Crick (2003) reported that where teachers had to deliver results in exams, little use was made of formative assessment: 'summative assessment squeezes out formative assessment'. (p.170).

Swaffield (2011) was critical that much of what teachers in the UK understood to be formative assessment was in fact regular summative testing. This misinterpretation was compounded by the Assessment for Learning Strategy which was rolled out in 2008 (DCSF, 2008).

Heritage (2010) and Shepard (2008) described a situation in the US where large scale interim testing (using tests purchased from international educational publishers such as Pearson) was being misnamed formative assessment and used for accountability.

Perrenoud (1998) argued that the use of specific techniques, such as improving feedback to pupils, was in itself not enough to bring about improved learning. He said that much feedback is like ‘so many bottles thrown out to sea’ (p.87), there was no guarantee that it would be received. Classroom teaching was very complex and improvements in learning involved improving ‘regulation’ of learning by pupils through continual adjustment and adaptation of the teaching based on discussion with pupils. This regulation of learning required very careful planning by the teacher to prepare proactive intervention and to be able to adjust instruction on-the- go (Perrenoud, 1998).

Complexity in managing the process of feedback was also acknowledged by Cowie and Bell (1999), who reported on a two year study of 10 science teachers carried out in New Zealand. While they agreed that formative assessment was integral to the learning process, they emphasised that teachers’ pedagogical knowledge which was pivotal in interpreting and recognising pupil responses was a possible weakness to its development.

The lack of a proper strategy for teacher professional development undermined the possible effectiveness of formative assessment. Teachers who did not really understand the key principles followed the ‘letter’ and not the ‘spirit’ of the approach. Torrance (2007) criticised ‘assessment as learning’ in colleges. This, he explained, was where the learning experience of many students came to be dominated by shared learning objectives and assessment criteria. Webb and Jones (2009) presented findings from a small scale study in which they found that teachers had such different starting points with regard to formative assessment that bringing about the changes required on a school level would prove difficult. It was assumed that teachers could easily adapt instruction based on the principles of formative assessment but a study by Heritage, Kim, Vendlinski and Herman (2009) demonstrated that this was not the case.

Coffey, Hammer, Levin and Grant (2011) agreed with Cowie and Bell (1999) that teachers’ pedagogical subject knowledge was vital in making the interpretations of student responses necessary to make formative assessment work. Coffey et al. (2011) examined some of the transcripts of classroom dialogue from five highly cited published papers which they regarded as influential in the literature of formative assessment. They concluded that without expert knowledge of the particular area of science, it was difficult for a teacher to assess the responses that pupils made. For example, one teacher who was so focussed on the semantic difference between mass and weight had not credited a pupil’s reference to the effects of forces in a flotation problem. The pupil’s response, in fact, revealed a good understanding but Coffey argued the teacher’s focus was not on the substance of the pupils thinking rather that the body of knowledge she had to teach. Coffey argued that no amount of ‘wait time’ or other formative assessment technique would help develop pupil ideas in a situation like this.

2.3.4 A call for more research

Kingston and Nash (2011) re-examined the 1998 Black and Wiliam review and complained that it was being regarded with such deference among many educators that awareness of the limitations of the original literature was being overlooked. The article had initiated huge interest in the field with a promise of large learning gains which proved unfounded in the presented evidence.

Many of the difficulties with the notion of formative assessment were brought together in a review by Bennett (2011.) He listed a range of issues which needed to be clarified before educational professionals should make claims for its effectiveness including:

- Measurement. What was it exactly that formative assessment measured? Many teachers struggled to understand that formative assessment did not have to provide summative information. Teachers had to use skill and judgement to infer meaning in pupil responses.
- Domain dependency. The long list of techniques such as traffic lights and peer assessment had to be applied to a set of materials which progressively built student understanding.
- Professional development. Formative assessment was not just ‘good teaching’. Teachers needed time to engage in an iterative cycle of practice and reflection to adapt the ideas behind formative assessment to their own teaching.

Bennett called for a closer definition of the term formative assessment pointing out that in the research literature, it covered a very wide range of interventions. He asserted that the effect size of 0.7 often attributed to formative assessment was unsubstantiated to date through research. Despite these criticisms, Bennett regarded formative assessment as ‘work-in-progress’ and thought that hard work was still needed to ‘realise its considerable promise’.

2.4 Measuring conceptual understanding

The next part of this literature review addresses the issue of progression offered by the attainment targets in the National Curriculum. The level descriptors are too vague and abstract to structure student progress (DfE, 2011). In this section (2.4), I will explore some alternatives which might provide a more research-based approach to assessing student performance.

2.4.1 The Bloom Taxonomy

In the cognitive taxonomy created by Benjamin Bloom and colleagues, educational objectives were classified into six categories (Krathwohl, 2002) (p32):

- knowledge – recalling or recognizing knowledge from memory.
- comprehension – constructing meaning from instructional messages.
- application – carrying out or using a procedure/applying learned material in situations familiar and unfamiliar.
- analysis – breaking material or concepts into parts or determining how parts relate to one another in an overall structure or purpose.

- synthesis – put elements together to form a coherent or functional whole.
- evaluation – make judgements based on criteria.

The taxonomy is hierarchical in nature which implies that the earlier levels must be mastered before one can move onto the higher levels. The lower levels of the taxonomy relate to learning, understanding and applying knowledge and information. Those that follow are often referred to as 'higher order thinking skills' and require more involved consideration of learning material by a student. It is common to see Bloom categories exemplified by lists of action verbs which are intended to help teachers design learning activities which specify a learning demand they wish to achieve. Such verbs for the lower levels include: list, name, recall, select, describe, use interpret and explain. Higher order verbs include: distinguish, categorise, critique, test, hypothesize and design.

Bloom's taxonomy has been used by examination boards to classify the level at which a question is set (Hunt, 2011; Azar, 2005) and is often quoted in guidance given teacher to structure their questioning during lessons (Gall, 1984; Blosser, 2000).

Sugrue (2002) expressed problems with using the Bloom taxonomy to design instruction and assessment. Depending on who was using the taxonomy she argued, learning objectives could be placed in different categories. In classrooms, Sugrue claimed, there was no practical distinction between understanding and knowledge. Everything above the knowledge level was treated as 'higher order thinking'. Hence, the taxonomy in practice, amounted to two levels.

Oversby (2002) constructed a table of concepts using the Bloom taxonomy to demonstrate how the classification of objectives could be applied to dissolving (table 2.2). The distinctions between the categories do appear to provide a differentiation of concepts which could be useful in assessing student performance.

Level	**How concepts can be used**	**Questions about dissolving**
Recognition	The concept has been met before.	Do you know that dissolving happens when we put sugar into tea?
Recall	The concept can be distinguished from similar concepts and recalled.	When we make tea by putting hot water on a tea bag, the water goes brown. What is the name of the process by which the brown material is extracted from the water?
Comprehension	The concept can be understood in familiar and unfamiliar contexts.	Grass stains can be removed using white spirit dabbed onto clothing. Explain this using the term 'dissolving.'
Application	The concept can be used in new context to develop general ideas.	Caffeine can be taken out of coffee using liquid carbon dioxide. Explain why this takes out only caffeine and not other parts of the coffee?
Analysis	The concept can be used to tease out meaning of a set of data or related concepts.	Sodium chloride is quite soluble in water but hardly soluble in hydrocarbons. Discuss this difference in terms of solvation energy of the chloride and sodium ions and the lattice energy of the sodium chloride solid.
Synthesis	The concept is used to integrate a collection of ideas.	When anhydrous blue cobalt chloride is mixed with water, a pink solution is created. How can this be explained in terms of enthalpy and entropy changes?
Evaluation	The concept can be compared with other related concepts for the same phenomena for explanatory power.	The solubilities of salts containing doubly charged anions and cations are generally lower than the solubilities of salt containing singly charged ions. How far is this true? Can you explain exceptions to this generalisation?

Table 2.2 A modified version of Bloom's taxonomy (Oversby, 2002) (p. 151)

2.4.2 Piaget's stage theory

Piaget postulated that children pass through four qualitatively different stages of intellectual development as they grow from birth to adolescence (Bliss, 2002). The stages are summarized below: (Bennett, 2003a, p. 256)

- Sensori-motor phase – (birth to about age 2) children learn from information which they gather directly through their senses and physical experiences.
- Pre-operational stage – (from about age 2 -7) children reason directly from what they perceive, though their reasoning may not always be logical.
- Concrete operational stage – (from about age 7 - 11) thinking becomes characterized by logic and does not require real objects to be to hand. The characteristic of children at this stage can conserve. i.e. they can see that quantities such as mass and volume remain constant in operations.
- Formal operational stage – (from about age 11 onwards) children become capable of abstract thought and are able to grasp ideas such as those involved in the control of variables and ratio and proportion.

Erickson (2000, p278) cites Driver's criticism of the fixed (or invariant) stages of the cognitive development of children. Driver also questioned the validity of the original work on which Piaget developed his ideas.

Shayer and Adey (1981) reported on findings from the Concepts in Secondary Mathematics and Science Programme This was a large scale research project into the difficulties children had in science and mathematics, (sample size 12,000). The programme's conclusions offered some validation of the Piagetian stages. Six stages and sub-stages were identified in secondary school populations:

- **1** pre-operational
- **2A** early concrete operation
- **2B** late concrete operational
- **2B/3A** transitional
- **3A** early formal operational
- **3B** late formal operational

The research found that only 30% British school children had reached formal operations by the age of sixteen. This suggested that many secondary students found science difficult because they had not yet reached a level of thinking necessary to understand many of the concepts. Shayer and Adey noted that the deficiencies highlighted by the research might also provide the tools to rectify them. These authors developed the Cognitive Acceleration in Science Education (CASE) materials which demonstrated some enhancement in children's cognitive development which led to improved GCSE performance (Adey & Shayer, 1993).

2.4.3 Three level of representation in chemistry

De Jong and Taber (2007, p. 631) cite Johnstone's three levels of representation in chemistry. Figure 2.1 shows the 'triangle of meanings' which displays the three perspectives from which chemistry topics can be viewed or taught. The macroscopic domain refers to phenomena which can be observed and experienced such as colour changes, new products being formed or disappearing. The sub-microscopic domain is based on the particle nature of matter and used to explain the macroscopic phenomena in terms of the behaviour of atoms and molecules. The symbolic domain is necessary because the particles described at the sub-microscopic level are too small to see. All three domains are integral in developing an understanding of chemistry concepts.

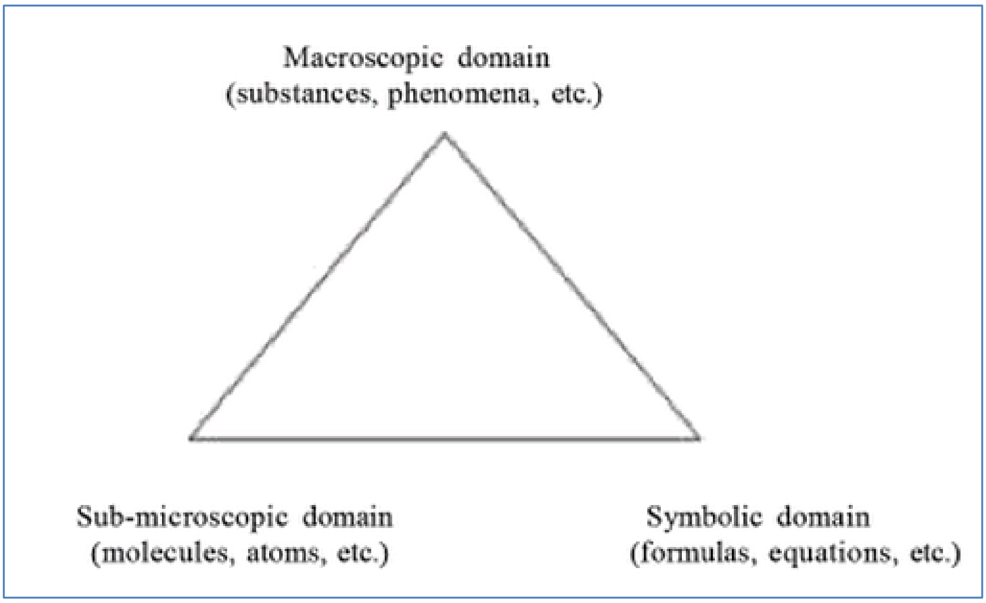


Figure 2.1 The triangle of meanings (De Jong and Taber, 2007, p. 632)

2.4.4 Learning progressions

In 2000, Wilson and Sloane reported on the use of an ‘alternative’ assessment system developed by the Berkeley Evaluation and Research (BEAR) Centre at the University of California. The BEAR system was a generic embedded assessment system which incorporated assessment into day-to-day classroom activities. The method identified progress variables which represented knowledge, skill or competency associated with learning a part of the curriculum. The BEAR approach was to measure the development of the progress variable throughout the course of instruction. This provided information about how knowledge, understanding and skill progressed in students. Wilson and Sloane claimed that this system allowed teaching and assessment to be aligned more closely and provided a context for diagnosing student need. This work led to the idea of Learning Progressions (Alonzo & Steedle, 2009).

Learning progressions are defined as,

> ‘*Successively more sophisticated ways of thinking about a topic that can follow and build on one another as children learn about a topic over a broad span of time.’*
>
> (Songer, Kelcey & Gotwals, 2009) (p 611)

Alonzo and Steedle (2009) reported work done on force and motion. They reviewed science education research in this area to produce a range of concepts which students should understand. They then placed the concepts in order of increasing difficulty including misconceptions (see figure 2.1).

Level	Description
5	Student understands that the net force applied to an object is proportional to its resulting *acceleration* (change in speed or direction), and that this force may not be in the direction of motion. Student understands forces as an interaction between two objects.
4	Student understands that an object is stationary either because there are no forces acting on it or because there is no *net* force acting on it. However, student may have misconceptions related to a belief that the applied force is proportional to an object's speed or motion (rather than its acceleration). Student can use phrases such as “equal and opposite reaction” to justify the existence of no net forces but may not understand this as an interaction. Common Errors: • Motion is proportional to the force acting. • A constant speed results from a constant force. • Confusion between speed/velocity and acceleration.
3	Student recognizes that forces are not *contained* within moving objects; however, student believes that motion implies a force in the direction of motion and that nonmotion implies no force. Common Errors: • Forces are associated only with movement. • Forces are viewed as causing things to move but not causing things to stop. • If there is motion, there is a force acting. • If there is no motion, then there is no force acting. • There cannot be a force without motion. • When an object is moving, there is a force in the direction of its motion.
2	Student recognizes that forces can be caused by nonliving things; however, student may believe that forces reside in within moving objects. Common Errors: • A moving object has a force within it that keeps it going. • A moving object stops when its force is used up.
1	Student understands forces as a push or pull, but believes that only living or supernatural things can cause forces. Common Errors: • Forces are caused by living things. • Forces are associated with physical activity or muscular strength. • Weight, motion, activity, and strength are important in determining an object's force.
0	No evidence or way off-track

Figure 2.2 Learning progression of force and motion. (Alonzo & Steedle, 2009)

The assessment they carried out was in the form of Ordered Multiple Choice (OMC) questions. OMC items are linked to a model of student cognitive development so the answer chosen by a student can indicate a level of understanding. This was a technique this group had piloted in earlier work (Briggs, Alonzo, Schwab & Wilson, 2006). Responses to OMC items were designed to match different levels of understanding. But results in this study proved inconsistent. Pupils who chose a higher level response to one question might choose a low level response to another. The clear progression which the technique predicted was not apparent in the findings.

Further work on learning progressions on density and floating was reported by Shavelson (2008). The approach to the taught material was much the same but the assessment strategies were more sophisticated. Shavelson listed four types of knowledge which could be assessed; declarative, procedural, schematic and strategic. Each type of knowledge would be assessed using a different approach. The results in this study were again inconclusive as teachers struggled to implement the activities exactly as the researchers had intended.

Interest in learning progressions remained high and other articles were published, for example, about scientific modelling (Schwarz et al., 2009), genetics (Duncan, Rogat & Yarden, 2009) and carbon cycling (Mohan, Chen & Anderson, 2009) among others (for a full list see Duschl et al., 2011) However, learning progressions remained a hypothetical model of student learning. Assessments were carried out in an attempt to validate or confirm the progression. With a small shift in emphasis, there came a realisation that learning progressions could act as a support for the techniques of formative assessment.

2.4.5 Combining learning progressions and formative assessment

Heritage (2008) discussed how learning progressions might be used for instructional planning and formative assessment. As they delineate progression from novice to more expert performance, the proposed trajectory provides a clear view of the building blocks of learning – what to teach and what is to be taught next. She asserted that they are ideally suited to eliciting evidence and providing feedback which are two important strands of formative assessment.

Learning progressions allow assessment information to be gathered in a systematic, planned way. Teachers know what and when to assess. Even if an informal assessment opportunity arises, teachers are aware of the overall trajectory and can base judgments within that framework. Teachers are more likely to *notice* student thinking and be able to spot errors and misconceptions during teaching (Bennett, 2011). Success criteria (Leahy et al. 2005) or 'what a good one looks like' (Sadler, 1989) are already set by the learning goals. The learning goals form a strong basis for determining a good performance. Coffey, et al. (2011) argued that formative assessment should not be domain-general practice. Learning progressions could be assessed through the use of content specific resources. Feedback, which we have already seen is important for motivation and self-regulation could be timely and specific. The issue of not being clear about what to base feedback on (e.g., Schneider & Gowan, 2013), is dealt with effectively. Student performance could be measured using tasks which provide information about student understanding. Teachers can use the continual stream of evidence about how pupils are heading towards a desired goal to adjust further teaching and assessments if necessary. Feedback is easier to provide because it deals with chunks which are of a manageable size. Finally, the feedback does not involve students comparing themselves with others in the class; they can understand their own performance in relation to a goal.

Despite this compelling logic, however, questions remain about learning progressions. Commentators refer to it as an idea that still has a long way to go. Shavelson and Kurpius (2012) raise some important questions. Does the research literature provide an exhaustive picture of the concepts and misconceptions that pupils can develop during a taught science topic? Is there one way to proceed through a topic or are there multiple routes that cognitive development could occur in students? What about the number of levels in a given topic or the grain size of the steps. These remain questions which research has still to answer.

Duschl et al. (2011) cite Driver as saying;

> *"No claims can be made about the pathways in thinking individual students follow."* (p.171)

Duschl et al. (2011) add that learning progressions must be viewed as longitudinal studies over many years rather than the short teaching sequences which have been investigated so far.

2.5 Probing for understanding

This section examines the literature for techniques of diagnostic assessment probes.

Wylie and Wiliam (2006) extended the idea of using hinge or diagnostic questions. In the formative assessment 'toolkit' these questions would be asked for mini-white board or index card responses. The aim was to find out how many of the students in a class could supply the correct response demonstrating readiness to move on. Between fifty to one hundred individual test items were written. These were simple multiple choice questions with a single correct response and distractors which matched possible misconception students might hold. Asking a *single* question did not present teachers with a time-consuming task and did give teachers a clear insight into what students were thinking (Ciofalo & Wylie, 2006).

Keeley and Eberle (2008) developed Curriculum Topic Study (CTS) probes in science teaching. These were activities which might involve a single multiple choice questions or an extended written response to a science problem. The approach was informed by two sources; the science concepts students needed to understand and the ideas research showed they tended to hold. Many of the test items contained a two tier assessment with a multiple choice part and a further question to explain the choice (Treagust, 2006).

Both of the above approaches use very logical reasoning to plan their assessment probes based on good educational research data about student ideas in science. There is however no published material to show how well this approach works.

The York Science Project is a more recent initiative. Embedded formative assessment materials have been developed for use in Key Stage 3 science and piloted in 35 UK schools (Millar &Whitehouse, 2012). This is a work in progress and has not yet produced any published data. Following the trialling of materials the resources will be launched on a website platform for wider use by teachers.

2.6 Student understanding in solution chemistry

Designing formative assessment around learning progressions may provide a more effective strategy to elicit information about student thinking (Heritage, 2008). To date, no work has devised embedded formative assessment probes for the concept of dissolving. Understanding dissolving well at Key Stage 3 is important as it provides a foundation for more demanding ideas in science such as: electrolysis, molarity and osmosis.

There is a large literature of the misconceptions that students hold concerning the use of the particle model in schools' science (Driver, Squires Rushworth & Wood-Robinson, 1994) This project aims to investigate the use of formative assessment probes in a solution chemistry topic. For this reason it is necessary to identify the possible alternative frameworks or misconceptions that students are likely to have. The research literature reports the following which students have difficulty in understanding concepts related to dissolving:

- Students use the words melt and disappear when referring to dissolving (Holding 1987; Driver et al., 1994).
- Students in Year 8 (13 years old) depict solute as bits drawn in and distributed without water particles. A continuous view of water is prevalent; shown by shading or 'wavy' lines (Driver et al 1994, Nakleh, 1992).
- Solute is located at the bottom and not evenly distributed throughout the mixture (Prieto, Blanco & Rodriguez 1989).
- Sugar would not pass through a filter paper during filtration (Devetak, Vogrinc & Glazar 2009).
- Dissolving requires stirring (Prieto et al, 1989; Ebenezer & Erikson 1996). Dissolving is similar to changing state at a particle level (Calyk, Ayas & Ebenezer 2005).

- Mass is lost during dissolving (Holding, 1987).
- Stirring makes dissolving occur faster as solvent particles have more energy (Ebenezer and Fraser 2001).
- Water particles absorbs the solute particles (Prieto et al. 1989).
- Holding (1987) and Kapabinar, Leach and Scott (2007) also make the observation that students rarely mention particles without being prompted or if they are referred to it can be loosely such as bits or small pieces of solute.

Assessment probes designed to reveal these alternative views among students will play an important role in assessing the overall development of students' ideas about different aspects of dissolving.

2.7 Summary

This literature review has highlighted some of the areas which would be useful to address the research questions. Formative assessment needs to focus specifically on the science substance of the learning. Feedback from learner to the teacher is a powerful form of feedback which produces learning gains. Formative assessment seems to be useful in conjunction with a learning progression. Measures of conceptual demand which preceded the National Curriculum are supported by research evidence. Planning or embedding specific questions to act as probes looks to be an efficient way of measuring student understanding.

These ideas will form the basis for my approach to answering my research questions.

Chapter 3

Methodology

3.1 Introduction

The aim of this project is to improve the formative assessment in my own teaching. In chapter 2, I highlighted the main features of formative assessment and its development and difficulties over previous years. I also looked at the conceptual hierarchies that have been identified. In this chapter I will set out my research questions and how I plan to collect data to answer them.

3.2 Research questions

The work aims to address two main research questions:

RQ1 To what extent can diagnostic questions be used to manage the collection of accurate, real-time data about pupil understanding to shape teaching during a typical science topic?

RQ2 To what extent can student responses to diagnostic questions be used to indicate their individual performance on a cognitive progression?

The task of assessment has to be manageable for a classroom teacher working in a busy classroom during a full teaching day. As Black and Wiliam (1998b) describe:

‘Teachers have to manage complicated and demanding situations, channelling the personal, emotional, and social pressures of a group of 30 or more youngsters in order to help the learn immediately and become better learner in the future.’ (Black and Wiliam 1998b) (p.140).

RQ1 is about balancing the teacher’s roles in advancing the teaching while measuring its impact. By real-time I mean that the assessment is carried out lesson by lesson, idea by idea. Carrying out assessment during the teaching is a challenge which the diagnostic questions will address in this research. Accurate means that the information included in the student responses should tell me clearly about students’ thinking. The diagnostic questions focus on specific scientific concepts and misconceptions. This is in contrast to some formative assessment techniques which ask students to indicate a general feeling of whether or not they understand. By using the information gathered from the diagnostic questions I will make decisions on how to proceed with the next lesson or part of it. A typical science topic would take between ten to fifteen hours of teaching to complete.

RQ2 seeks to build a picture of the performance of individual students during the course of the lesson series. Normally, summative testing is carried out at the end of instruction to assess student performance. This takes up a lot of additional time and produces a score which may not be very helpful in indicating what a student knows or understands. The teaching sequence will follow a carefully constructed learning progression which will build on progressively more challenging ideas. I had no fixed view at the start of this research how I would frame a summative decision on student learning. The research literature reviewed in chapter two has provided some possible alternatives in terms of Bloom’s taxonomy, Piaget’s stages and the conceptual demands found within a learning progression. The emphasis for RQ2 is to see what extent the diagnostic questions can provide information towards a decision: the judgement will ultimately come from the teacher. RQ 2 is not about attributing a decision based on a pre-determined score or proportion of work completed. For that I can always carry out a test.

3.3 Deciding on a research strategy

3.3.1 Rejecting an experimental research design

Measuring the effectiveness of a teaching intervention may be carried out using an experiment (Yin et al., 2008). Groups could be randomly assigned to one condition or another and the performance of the different groups could be measured using pre- and post-tests to indicate how students had progressed. Experiments can provide results which can be argued to be objective, reliable and valid. As a science teacher with a science background this is an approach with which I am comfortable.

Yin (et al., 2008) reported on an experimental study on the impact of formative assessment. Reflecting on the reasons why the study did not provide the evidence expected they conceded that:

> *"it is much more difficult to conduct a 'perfect' experimental study in education than in natural sciences or a psychology laboratory, because it is almost impossible to control for many factors."* (p.354)

Truly experimental research in the social sciences is described by Hakim (2000) as virtually impossible and not realistic. She goes on to say that no type of study is inherently better than another ***in so far*** that the design should be selected according to the purposes of the research and the issues to be addressed. (My emphasis).

The purpose of this research is to collect data on learning and use it formatively. I felt that this was beyond my skill as a researcher to conduct a comparison without the use of test data. The data generated will be indistinguishable from the classroom work children produce in lessons normally. I felt that an approach which focussed on changing practice and which measured any impact was more suited to this type of investigation. This research will be an evaluation carried out in the context of an action research project. Below I will briefly describe each and give some reasons why I feel they are suited to this research project.

3.3.2 Carrying out an evaluation

The main reason for carry out an evaluation in this study is to answer the questions:

1. *Is the programme achieving its goals*
2. *Does the programme have an effect?* (Cronbach, 1987) (p11)

Different methods of evaluation exist which could describe this research:

- Summative evaluation Stake, (1986) in Bennett (2003b) (p10) seeks to answer questions about what relationships exist between the goals of a programme and its outcomes.
- Goal free evaluation Scriven 'undertaken without reference to any statements of outcomes produced by the programme developers.' Bennett (2003b) (p 30)
- Parlett and Hamilton refer to evaluation as illumination (1976) where taking into account the wider contexts in which educational innovations function the

> primary concern is with description and interpretation rather than measurement and prediction (p.88)

This study is a goal free evaluation because I will look at the outcomes and compare them to the needs which I have identified for the research to address. The research will have both intended and an unanticipated consequences. Scriven voiced concerns that the side effects tend to be played down (Scriven, 1991). I will attempt to evaluate the actual effects (positive or otherwise) of the research study on my students and on my professional practice as a teacher.

3.3.3 Action research

Action research is research undertaken by practitioners (e.g. teachers) for the purposes of improving their practice and understanding (Kemmis, 2009). It is usually carried out while the practice is being performed. Its aim is to address problems and tackle them in whatever way seems most appropriate. This approach is suitable as it is an idiographic methodology which considers the ways individuals create, modify and interpret the world. Elliott (1991) describes action research as “the study of a social situation with a view to improving the quality of action within it” (p69). Action research is not characterised by particular data collection techniques used but by the attitude to the knowledge developed. The emphasis is on explanation and understanding rather than objective reality.

Action research involves cycles of action and critical reflection. A formulation of action research cycles is shown below:

- Identify a concern.
- Think of a possible way forward.

- Try it out.
- Monitor the action by gathering data.
- Evaluate progress by establishing procedures for making judgements about what is happening.
- Modify the practice in the light of the evaluation. (Adapted from McNiff and Whitehead, 2011).

If the intended purpose of the research is to evaluate classroom activities the best way to collect data from students is to use the written material they generate in lessons. This work requires the researcher to be involved in what is taking part and not as an external observer.

A research design such as the one I have indicated might arguably lack the objectivity of an experimental approach but there are features which make the chosen strategy a better one with which to proceed.

This principal source of the data will be collected in a natural setting. Students will be working in their normal timetabled science lessons with their normal teacher. The students' work will provide a rich source of descriptive data which will be analysed by the researcher (teacher). A classroom teacher is a key instrument in the research and this means that catching essential meaning and intention is maximised.

Challenges to validity in the context of this work arise from issues of honesty in the reporting and researcher bias. Another factor which could be considered is the small sample size. However the scope and timescale of the project is realistic in terms of what can be managed.

Because of the limits of time, this project it will only contain one cycle.

3.4 Constructing the research instruments using backward design

3.4.1 Backward design

In designing embedded assessment activities that demonstrate understanding we 'begin with the end in mind' (Covey, (1998) cited in Wiggins and McTighe (2005, p. 1). In their book, Understanding by Design, Wiggins and McTighe (2005) point to the 'pointlessness' of many school experiences which although cover the curriculum content, only accidentally lead to insight or achievement. They suggest an alternative approach, namely 'backward design.' This begins with the question; what will students understand at the end of the teaching and what evidence would show this? This is an approach which has been adopted by other groups (Leach & Scott, 2002; Whitehouse, 2014). The stages of backward design are as follows:

Stage 1 - Identify desired results.

Stage 2 – Determine acceptable evidence

Stage 3 – Plan learning experience and instruction.

3.4.2 The stages of backward design in this project.

Following this model the assembly of classroom research materials went through several stages:

- Firstly a learning progression or *teaching sequence* had to be constructed using a suitable research-informed curriculum of solution chemistry. The misconceptions which pupils have in this area were also identified and included as part of the teaching sequence.
- These ideas were collated to stage a '*scientific story'* (Ogborn, 1996 in Leach and Scott, 2002, p. 122). The resulting 'solution chemistry story' summarised the knowledge I would want a student to learn during the teaching.
- From the scientific story, a set of individual *learning intentions* were written so that operational *learning outcomes* (what could students *do* who understand this work?) could be identified.
- Finally the *evidence of learning items* were written and set into the classroom activities to be used by students.

Each of these stages will now be discussed more fully in the sections immediately below.

3.4.3 Identifying the range of content to be taught

The teaching sequence was constructed using strands from five sources: The national curriculum statements for key stage 3 (DfEE/QCA, 1999); the national curriculum level descriptions developed by Russell and McGuigan (2003a, 2003b) at the University of Liverpool; the learning progression devised by Johnson and Tymms (2011) on the concept of substance; education research on student misconceptions in solution chemistry and the science numeracy work of Lenton, Stevens and Illes (1999).

The 1999 national curriculum provides an appropriate range of science content within the topic of dissolving. It lists the following statements that students should be taught:

- How particle theory of matter can be used to explain the properties of solids, liquids and gases, including changes of state, gas pressure and diffusion.
- That mixtures are composed of constituents that are not combined.
- How to separate mixtures into their constituents using …appropriate methods.
- That when physical changes take place, mass is conserved.
- About the variation of solubility with temperature, the formation of saturated solutions, and the differences in solubility of solutes in different solvents. (DfEE/QCA, 1999) (p. 32)

The level descriptions by Russell and McGuigan (2003a/b) (appendix 1) give guidance as to what level of demand the National Curriculum statements. The authors have written additional statements, which appear together with the original National Curriculum level descriptions to provide a finer detail of 'levelness' than the National Curriculum statements alone. Johnson and Tymms (2011) published an analysis of item difficulties on the concept of a substance. These workers carried out an analysis of student responses to questions on different aspects of substance chemistry. Aspects (such as change of state or chemical reaction) were scored and a learning progression based on the scores was constructed. Ideas that students found more difficult were scored higher on a range from zero to 84. The region on mixtures and dissolving was used to produce a guide for the chemistry of solutions along increasing conceptual challenge and is shown in table 3.1.

Description of concept	**Score of difficulty**
Distinguish between dissolving and not dissolving.	21
Distinguish between melting and dissolving.	28
Recognise dissolving for coloured solutions.	31
Intrinsic motion in liquid state.	31
Recognise solutions as mixtures.	33-44
Salt crystal residue forms on evaporating salt solution.	34
Mass does not change with dissolving.	45
Filtering will separate a suspension not a solution.	54
A substance’s particles can explain dissolving.	60-65
Air and water can form a mixture.	62-66

Table 3.1 Learning progression for mixtures and dissolving (Adapted from Johnson & Tymms, 2011)

A significant amount of educational research has been carried out into the misconceptions children bring into science lessons (see section 2.6). Reviewing the literature has produced a wide range of misconceptions which need to be addressed in a teaching sequence based on forming solutions. These are summarised in table 3.2.

Description of misconception	Reported in
Dissolving describe as 'melting, and 'disappearing'.	Holding, 1987 Driver et al., 1994
A continuous view of solvent (water) is prevalent.	Driver et al., 1994; Nakleh, 1992
Dissolved solute sinks to the bottom of the container.	Prieto et al., 1989
Dissolved sugar will not pass through a filter paper during filtration.	Devetak et al., 2009
Dissolving requires stirring.	Prieto et al., 1989 Ebenezer & Erickson, 2005
Mass is lost during dissolving.	Holding, 1987
Stirring makes dissolving occur faster as it gives solvent particle more energy.	Ebenezer & Erickson, 2001
Water absorbs solute	Prieto et al., 1989

Table 3.2 Summary of difficulties students have with dissolving

An ability to handle numbers and interpret graph contributes to a understanding in dissolving and solutions as with many areas of science. (Lenton, Stevens & Illes, 2000). This project will involve assessment of students' performance of the following mathematical operations:

- tabulating experimental data.
- constructing line graphs.
- interpreting line graphs.
- using a formula to carry out a calculation.

3.4.4 Staging a scientific story

The three strands: backward design, content range and student misconceptions, were used to stage a science story. The story of the solution chemistry progression is shown in appendix 2. The story provides an overview of all the science content which will be taught. It is also effective in allowing the content to be arranged into sections which group similar ideas together which comprise learning intentions, i.e. the concepts I want the students to learn (table 3.3).

Section of solution chemistry story	Learning intention
1	• Solids are soluble or insoluble. • Solutions are transparent and possible coloured. • Dissolving can occur if some solute is left. • Dissolved solute can be detected by colour change or taste.
2	• Dissolving happens because solute particles separate and mix with solvent particles. • Solution can be retrieved by evaporation. • Dissolving and melting are not the same physical process.
3	• Stirring is not required for dissolving to occur.
4	• Mass is conserved but not volume.
5	• There are solvents other than water.
6	• A powdered substance will dissolve faster that lumps.
7	• Solutions can become saturated. • Solubility depends on solute and temperature.
8	• Gases can dissolve. • Gas solubility depends on temperature.

Table 3.3 Groups of learning intentions identified from solution chemistry story.

A summary of the five sections in the solution chemistry story, the learning progression of Johnson and Tymms (2011), the National Curriculum level description and the student dissolving misconception are is shown in appendix 3.

3.4.5 Writing intended learning outcomes

The next step was to write operational, or intended learning outcomes for each learning intention. Wiliam (2011) reminds us that '…students do not learn what we teach. If they did, we would not need to keep gradebooks. We could instead, simply record what we have taught.' (p. 47). Intended learning outcomes are an effective means of students demonstrating mastery of anything they have been specifically taught. For example; where students have been taught (learning intention) that the rate of dissolving changes with the available surface area of a solute, students might be asked to write a prediction (learning outcome) based on information given to them. The intended learning outcomes for each of the learning intentions is shown in appendix 4.

3.4.6 Writing evidence of learning items

This is the final stage in this sequence and produces the lesson resources which students will use in lessons. As a classroom teacher I frequently write worksheets to support students in lessons. This is a consequence of having limited access to textbooks. Worksheets are a flexible resource and can be tailored to specific needs. The 'evidence of learning' items form part of the classroom activities presented to students on worksheets during each lesson. Each item has an individual number for ease of analysis. The worksheets supply students with all the instructions for practical work where this is required and spaces for experimental results. All the worksheets used with students in this study are compiled in appendix 5. The evidence of learning items fell into six categories:

- observation/prediction with explanations;
- 2 stage multiple choice questions; (Treagust, 2006)
- short response questions;
- drawing diagrams;
- constructing a table for data/drawing line graph/reading line graphs;
- calculations based on a given procedure.

In all 31 items were written. For ease of communication in this thesis they will be referred to as diagnostic questions (or DQ.) The diagnostic questions were designed to produce responses from students which were short and specific. Students' completed worksheets with answers were collected and scanned for storage as a portable document file (pdf) format so that assessed work could be returned. All student responses were assessed by the author.

3.5 Participant and methods

3.5.1 Participants

The participants of this this study were year 8 pupils (n=60: m=27, f=33) in a comprehensive school in North Lincolnshire in the UK. The students were between thirteen and fourteen years of age. The students comprised two teaching groups and were taught at different times but by the same teacher in one hour lessons, three times per week. The two classes were designated as top sets at the school. Despite this, the students' Year 7 SATs in English and mathematics ranged from low level four (4c) to high level five (5a) which revealed the spread ability among students to be reasonably broad. A small number of the students were of ethnic origin and spoke English as a second language (n=17). There were no students involved in the study who were registered as having any specific special education need.

3.5.2 Teaching methods

The diagnostic questions were used in this study with students in their normal timetabled science lessons. Lesson activities were designed to produce learning experiences with which pupils would be familiar. One advantage of carrying out this style of research is that it is in a natural setting so Hawthorne effects[1] should be minimised.

Students worked in the classrooms as they normally would. They sat in groups of four forming two working pairs in which they also carried out practical work. One difference from normal teaching was that they were not required to use their science books. All the materials for writing on were provided by the project. This caused concern in some students who felt that they might not learn as much because they were not writing in their book.

Questioning and discussion in lessons followed some of the activities from the formative assessment 'toolkit' (mentioned in section 2.3.2). 'Wait time' (Rowe, 1974) was given before and after questions by the teacher. Students were asked to discuss ideas together and engage in exploratory talk before sharing with the whole class (Mercer & Dawes, 2008). Mini-whiteboards (an 'All-student-response system', Wiliam, 2011, p. 92) were made use of to obtain quick feedback from student when needed. A policy of no-hands (Leahy et al., 2005) was used to ensure that all students were involved in classroom dialogue.

Any science concepts necessary for student understanding but which were not part of the solution chemistry story, (e.g. diffusion, DQs 14-16) were reviewed before teaching at an appropriate time.

Three resources were used to model dissolving to students: Stuff and substance CD-Rom (Science Enhancement Programme (SEP), 2005), Stuff and

---

[1] The Hawthorne effect is defined by Zimbardo (2007, in Cohen, Manion, & Morrison, 2011) as where the context of the intervention could affect the outcomes and behaviours of the participants.

Substance: ten key practicals in chemistry (Johnson, 2011) and the dissolving activity from Sunflower Science (Sunflower Learning Ltd., 2006.)

A range of teaching strategies was employed: guided practical work; front of lab demonstrations; PowerPoint® presentations including using photographs of the students' own work; BBC bite size or YouTube video clips on a relevant aspect of the teaching; and Sunflower Science computer animations.

3.5.3 Data analysis

Student responses to the diagnostic questions were marked and assessed by me as the classroom teacher as I would do with any work student produced. No mark scheme was written because I felt it was better to use my own judgement to interpret answers. Application of rigid marking criteria could mean that some responses received no credit more because of issues with literacy than science. Finding meaning in the communications students make is part of the everyday activity of teachers. It is also consistent with an action research setting that a normal means of assessment is used. I will, however, refer to student responses as 'correct' to indicate that in my judgement the student demonstrated a sufficient degree of accuracy in a diagnostic question answer.

3.5.4 Reflective Diary

Throughout the time when the lessons were taught I kept a diary. This was an important part of the research process. The diary was used to record what was covered in discussion, what questions and comments students had made in the lessons and how I felt the activities had gone.

It also became the place where I recorded where teaching had progressed to. The DQs provided a large number of responses which had to be processed quickly. I used the diary to indicate which DQs had been answered together with tallies of student responses to give an overview. Any individual responses that could be used to stimulate further thinking in subsequent lessons were also noted.

It was also a means of recording some evaluative comments after certain lessons activities had been used. This will provide an important contribution to the evaluation of the effectiveness of the diagnostic questions which have been written and in answering the research questions.

The author has also kept a record of photographs of practical equipment set-ups, practical results, mini-white board responses, main class white board use and computer screen-capture from the lessons.

3.6 Ethical considerations

Before beginning the research project it was necessary to consider any ethical implications of the work. Taber (2007) reminds us:

> *'Whenever researching one's own students, it is important to prioritise the 'ethical imperative' and to try to contextualise the enquiry within an ethical framework that ensures students know their involvement:*
>
> - *is voluntary;*
> - *is safeguarded by confidentiality;*
> - *is not linked to any kind of formal class assessment; and*
> - *may be cancelled at any moment by their choice, and without detrimental consequences* (Taber, 2007) (p. 140)

This information was shared with the students who participated in the research.

In addition, I discussed two issues concerning the study with the senior management at my school: informed consent from student or parents/carers and the changes to the normal Key Stage 3 curriculum.

It was decided that written informed consent form students or parents/carers was not required because the data which the research would generate was within the normal course of teaching. No student work cited in the dissertation would lead to a student becoming identified.

The changes to the curriculum were considered within the range of changes that school science departments make on a regular basis. For example, if the department chose to deliver a new KS3 scheme (which would be a bigger change) staff would not seek permission from senior teachers or parents/carers.

To comply with the university's regulations on research ethics I completed and submitted the Department's Ethical Issues Audit Forms.

Chapter 4

Using diagnostic assessment to shape teaching.

4.1 Introduction

Formative assessment presents teachers with considerable challenges. My first research question asks whether the diagnostic questions designed in this project can contribute to effective formative assessment. In this regard, information has to be collected rapidly and has to be of sufficient quality and detail (accuracy) to allow the teacher to make decisions about next steps in teaching - which could be the next morning following an afternoon lesson. The diagnostic questions have been written with these needs in mind. In this chapter I will relate a series of vignettes from lessons which detail how pupil work was assessed, what feedback it provided and how this enabled teaching to be reshaped to improve student learning.

4.2 Recognising dissolving

In the first activity, students were given a range of substances and asked to test them to find out if they dissolved in water. The substances were salt, sugar, copper sulphate, chalk powder, glycerine, liquid green food colouring and wax pellets. Students were asked to record their observations and identify what had dissolved. The diagnostic question (DQ 1) asked them to give reasons for deciding whether a substance had dissolved or not. The practical was followed by a whole class discussion to share what had been observed. The discussion raised two issues. Firstly, many students had used the term ‘disappeared’ when they reported that a substance had dissolved. It is common for students to misuse this term in this context (Holding, 1997) and it was also seen in the written responses where 52% of students in one class stated that solute ‘disappeared’. It was possible during the discussion to challenge this and the term ‘transparent’ was accepted as a better alternative. Secondly, some students had not observed glycerine, green food colouring or copper sulphate dissolve. This may be because these first two substances were liquids and the latter two produce coloured solutions which may have led to a degree of uncertainty. The time in the lesson allowed for a demonstration to show that they did dissolve to produce transparent solutions, clear or coloured.

A closer look at the students’ written answers to DQ 1 after the lesson revealed a third issue that the discussion had missed. Around a quarter of students recorded that chalk had dissolved. This indicated a misinterpretation of the colour change. Students had recorded chalk as having dissolved because it had ‘turned the water white.’ During the lesson a photograph had been taken of all the test solutes in boiling tubes of water. The next lesson was started by showing the picture of the chalk suspension (Figure 4.1) and asking the question ‘Has it dissolved? This gave rise to a discussion which reinforced the agreed rule that solutions were transparent as well as tackling the error that chalk was soluble.

During the first lesson and before the second, the specific focus of the diagnostic question helped uncover student errors and misconceptions. As a teacher I was confident that we had achieved the learning intention that ‘*solutions are transparent and possibly coloured*’ and were ready to move on.

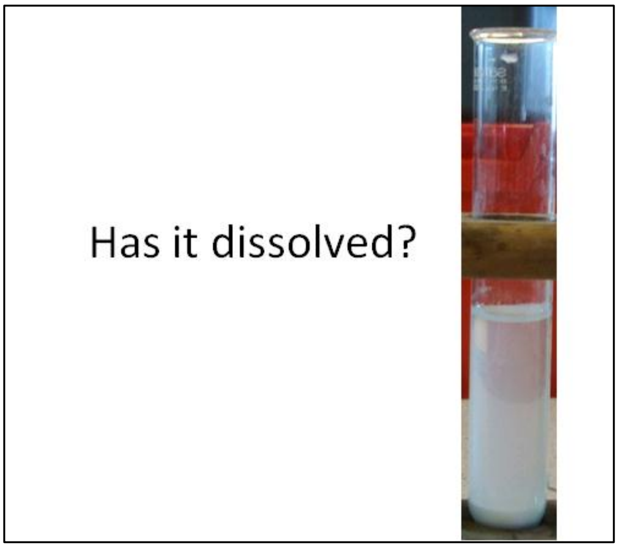


Figure 4.1 Slide of chalk suspension.

4.3 Using a particle model to explain

To support students in forming their explanations in the activities that followed, it was useful to revisit the particle idea of dissolving beforehand. Three resources were used to do this as detailed in section 3.5.2 (SEP, 2005; Johnson, 2011; & Sunflower Learning Ltd., 2006). The learning intention was to understand that '*dissolving happens because solute particles separate and mix with solvent particles.*' Students carried out tasks which would allow them to visualise the behaviour of particle during dissolving. Examples of the images the students used in these activities are shown in figures 4.2 and 4.3.

The activities and diagnostic questions which followed were designed to probe students' understanding of this process by applying their understanding to:

- filtration, and
- evaporation of a solution.

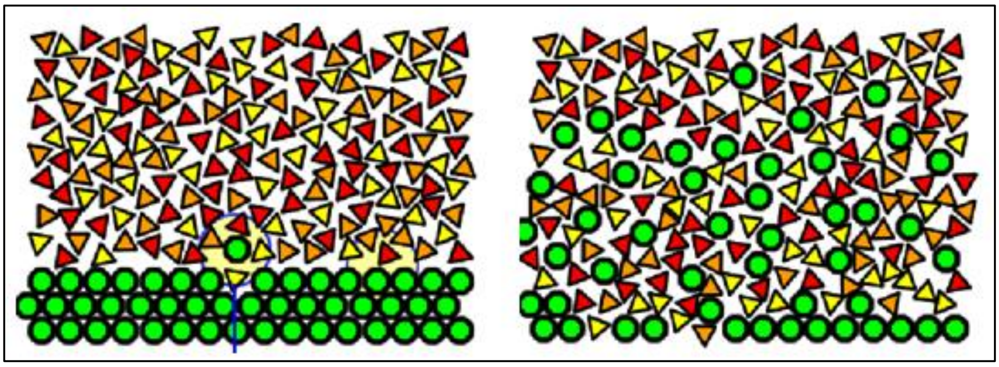


Figure 4.2 SEP materials showing dissolving of solute (circles) in solvent (triangles).

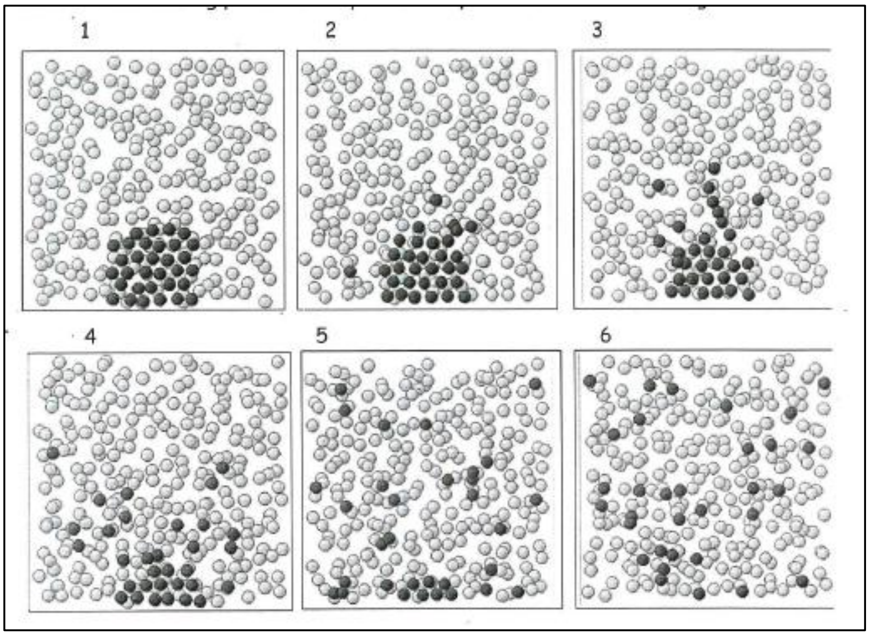


Figure 4.3 Sequence showing dissolving taken from Sunflower Science

4.3.1 Filtering solutions

Students carried out a simple procedure pouring a salt solution and a chalk suspension through separate filter papers. They were asked to draw what they observed. The diagnostic questions (DQ 5 and 6) asked for an explanation of the differences observed on the filter papers. One further diagnostic question (DQ 7) asked students to predict how the result would be different if the holes in the filter paper were smaller.

Many students (82%) recorded the expected observation that a white residue was seen on one filter paper which they identified as chalk but no residue on the other filter paper where the salt solution had passed through. A smaller number (60%) offered an explanation by linking this observation to the salt being dissolved and the chalk being not dissolved. No students applied the particle model of dissolving to the problem. When asked to predict what would happen if a filter paper with smaller holes was used, (DQ7), about a third of students predicted that salt would be left on the filter paper.

The student responses indicated that more time was needed to explore how the particle model could be used to explain the observation more fully.

This item was revisited in the next lesson. After further instruction and discussion students were asked to draw on a blank sheet of paper the particles of salt and chalk as they interacted with the filter paper. This activity produced two types of drawing. One type of drawing (done by 27% of students) is shown in figure 4.4. Drawings like this revealed that students visualised the chalk 'particles' as bigger than the holes in the filter paper and the salt particles as small enough to pass through. Later discussion confirmed these students did not visualise the chalk pieces as composed of lots of smaller particles.

The other type of drawing was similar to figure 4.5 (done by 31% of students). These demonstrated a clearer understanding that the salt particles were separate from each other because they had dissolved but the chalk pieces still

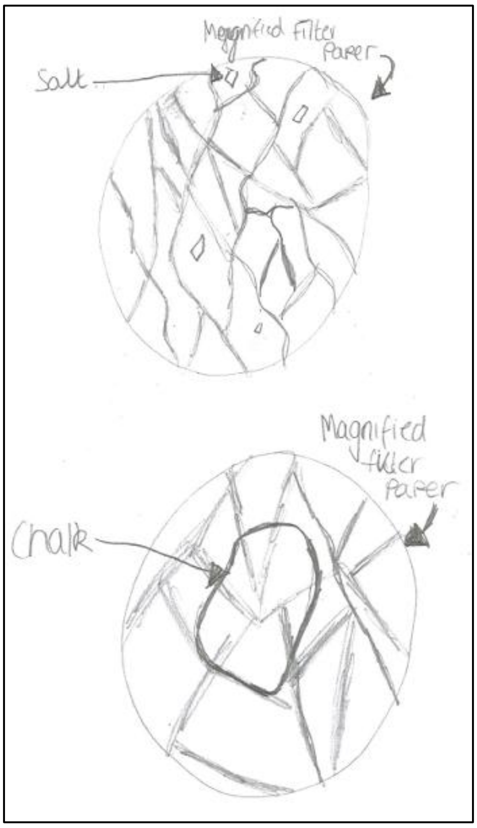


Figure 4.4 Student drawing of salt and chalk ‘particles’ on filter paper.

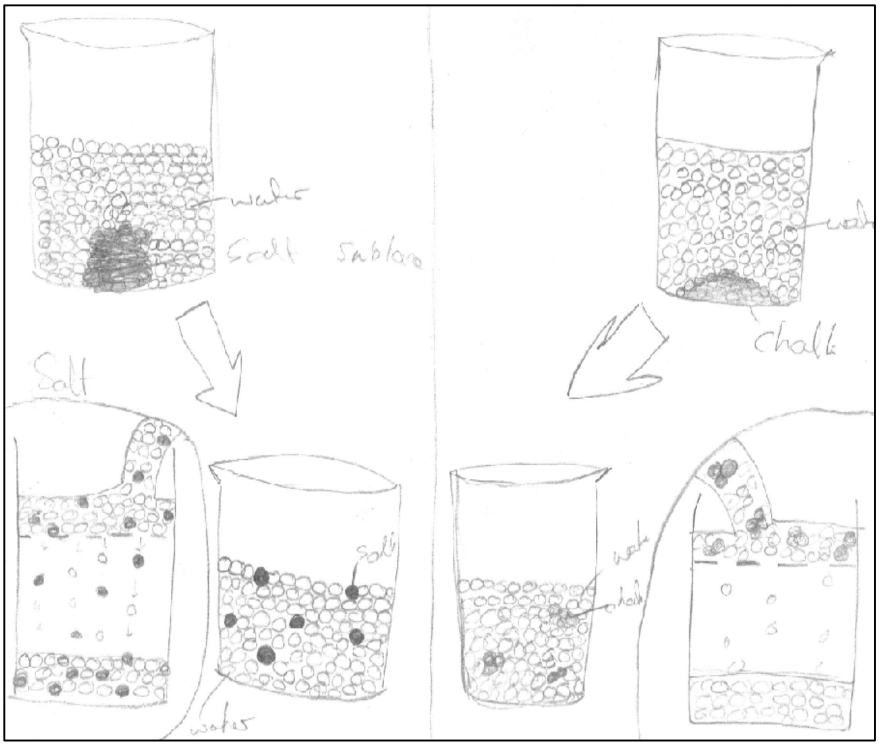


Figure 4.5 Student drawing of chalk, salt and water being filtered.

comprised lots of particles joined together. These students could explain that the chalk was seen on the filter paper because the groups of chalk particles were too large to go through the holes. Also these students saw the water as composed of particles which also passed through the holes in the filter.

This was the first instance where I used the feedback to make a decision to change from the planned teaching. Holding (1987) and Kapabinar et al. (2007) observed that students do not tend to give explanations involving particles unless prompted. Despite the explicit references to the particle model in the teaching that preceded it (figure 4.2, & 4.3), student responses were related to their (macroscopic) observations. The drawing activity was quickly improvised to assess further teaching. Responses second time around were more developed but they revealed a further misconception held by some students that chalk consists of particles which are too big to pass through a filter.

This created a dilemma of what to do: repeat whole class teaching, invite students back at another time, pair students who hold the misconception with others who could explain to them. The work was planned to take place in one lesson but ended up taking double this time because concepts had to be revisited. I took the decision to move on to the next lesson because the ideas of particles was to be covered again in later lessons. But the diagnostic questions had revealed some unexpected thinking among the students.

4.3.2 Evaporating salt solution

Students heated salt solution in a watch glass over a beaker of boiling water. They were asked to draw the particles of water and salt as they were heated in the watch glass (DQ 12). The learning intention was to demonstrate that '*solute can be retrieved by evaporation*' and this can be explained using the particle model. Figure 4.6 shows a typical representation of what students drew: a patch of salt in the centre of the watch glass with the water evaporated.

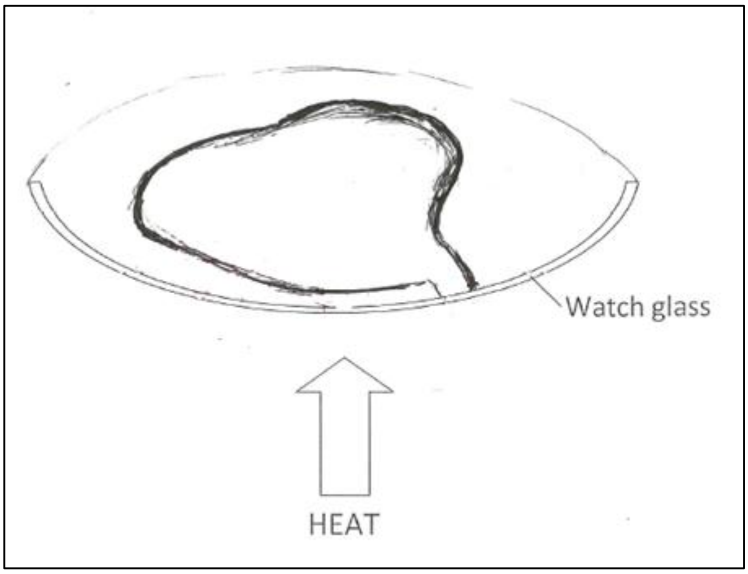


Figure 4.6 Student diagram of salt on a watch glass

The tendency for pupils not to refer to the particle model unless prompted (Holding, 1987; Kapabinar, 2007) had emerged again. It was necessary to revisit DQ12 to probe students' understanding more deeply. At the start of the next lesson the particle nature of the salt solution was revisited and the drawings were redone as a 'before' and 'after' diagram. Students now typically drew the salt solution (shown in figure 4.7) as a mixture of salt and water particles. The salt and water particles

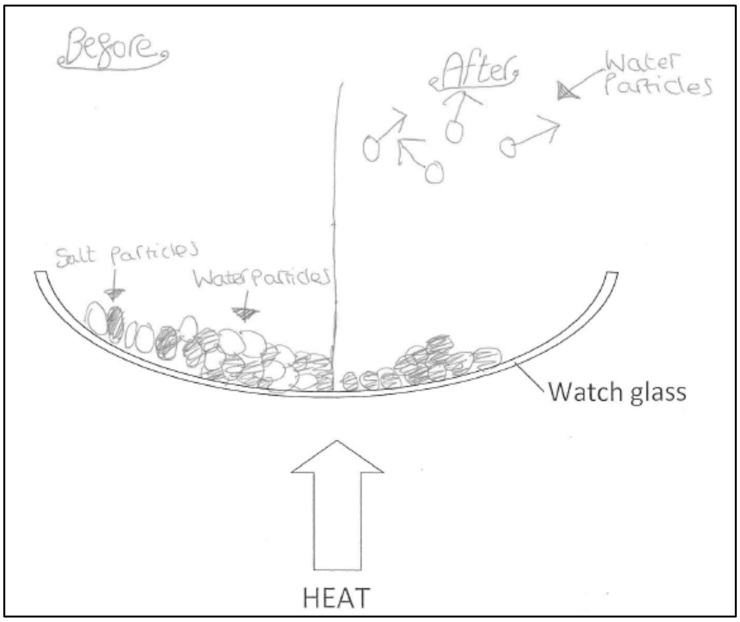


Figure 4.7 Student diagram showing evaporation of salt solution.

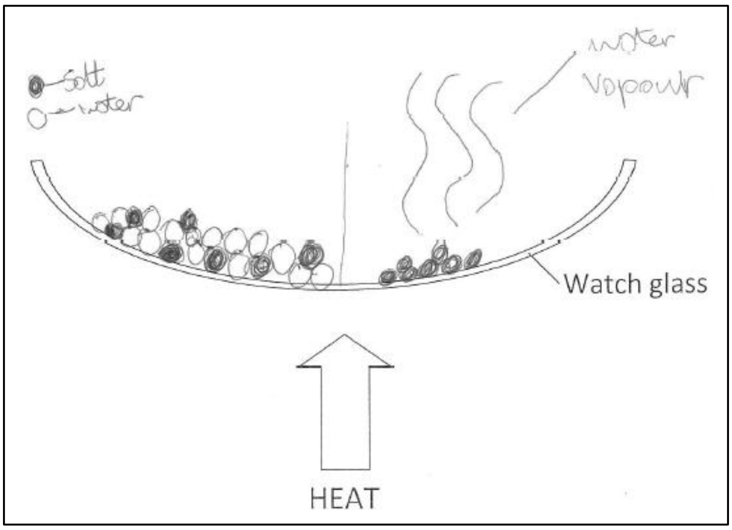


Figure 4.8 Drawing of water vapour represented using wavy lines

after evaporation were also correctly represented as solid and gas respectively. Some students represented water vapour as ‘wavy’ lines (figure 4.8) even after they had shown the solution and solid salt using particles. Nakleh (1992) suggested ‘wavy’ lines were a way in which students showed continuous representation.

These responses revealed some interesting student thinking. Did the particle and wavy lines display a half-way position between macroscopic and sub-microscopic representations? Or had students drawn both macroscopic and sub-microscopic representations to be sure of getting the right answer? I did not pursue either of these questions but the possibility was there.

To summarise, the re-teaching and reassessment of these ideas based on the initial answers given to the diagnostic questions DQ 5 & 6 and DQ 12 allowed some students to be moved forward in their learning. In both activities it had been necessary to remind students to explain in terms of particles. It was clear that about a third of students held a misconception that chalk did not filter because its particles were too big (figure 4.5) and another third had not responded to the extra time given to the filtration work. On the whole students had identified salt as the solid left on the watch glass after evaporation but many struggled to represent the solution, salt and evaporated water using the particle model without further help. Both activities had taken double the planned time to teach.

It is important to note that I was fully aware which students were struggling with the application of the particle model. In the course of my normal teaching I tell students the answer and ensure that the exercise books look as they should. Neatly presented diagrams copied from textbooks may conceal a real lack of understanding. Using diagnostic questions provided accurate, timely not to mention unexpected feedback from students

4.4 Is stirring needed for dissolving?

Probably all students are aware that stirring speeds up dissolving but a common misconception is that stirring is needed for dissolving to occur at all. The learning intentions aimed to address this misconception. Students were to gain an idea of the *intrinsic motion* of the water particles and how the dissolve solute *diffused* throughout the solution.

Following pre-teaching which re-capped on the ideas of diffusion, students watched a demonstration of a ‘crystal’ (small tablet) of potassium permanganate being placed at the bottom of a large beaker of cold water. They were asked to predict and draw a diagram to show what would happen to the crystal if it was left like this for a long time – say a week (DQ 14). They were also asked to predict whether the outcome would be different if the water was stirred over this time.

In response to the first part of the question where the water was unstirred, 66% of students drew a solid surrounded by an area of diffuse colour in a continuous solvent (for example, figure 4.9). Most students therefore understood that the solute would dissolve slowly and diffuse through the solvent. Drawings of the stirred beaker, where these were done (83% in one class left this blank) showed no solid and the colour evenly mixed through the water. Only 17% attempted a diagram using particles but nearly half of these contained inaccuracies.

Driver et al. (1994) have noted that the continuous view of a solvent (where matter is not represented by particles) is prevalent among students of this age. It is also likely that students drew their responses as macroscopic representations because the prompt for using particles in this question was missing. The learning intention

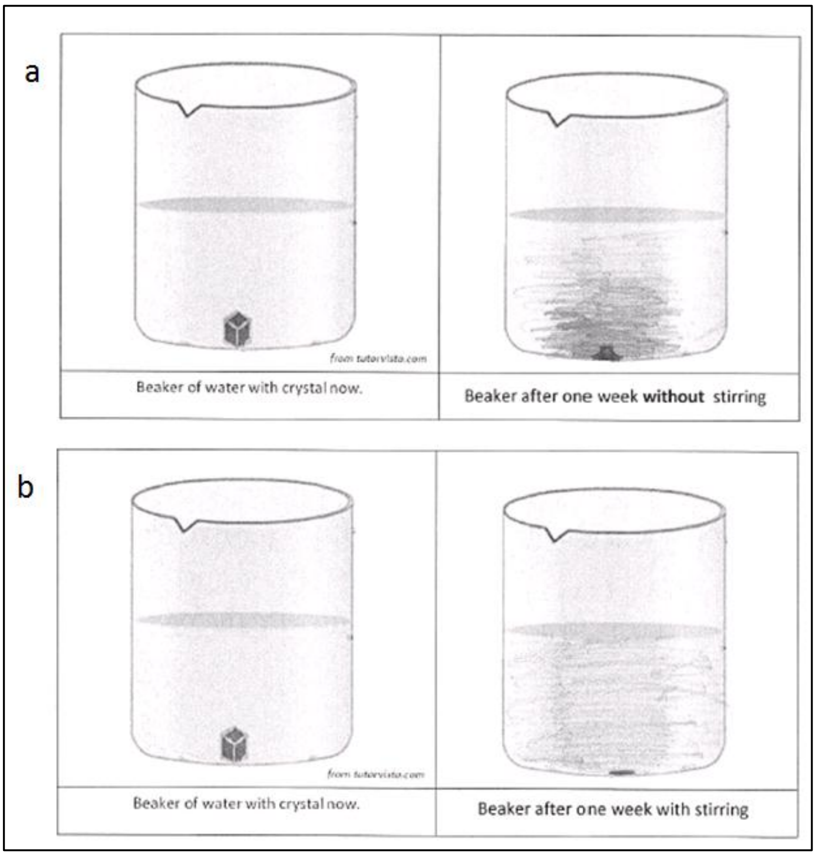


Figure 4.9 Student's drawings of solute dissolving with and without stirring

required students to use a concept of the intrinsic motion of water. This was unlikely to be achieved if student thinking was not in terms of particles.

Further instruction involved an improvised role-play out in the playground with students acting as particles of solvent (no blazers) and solute (blazers). Role-play is a technique I picked up during my teaching practice which provides a handy analogy of the kinetic theory of matter (Mc Sharry & Jones, 2000). This activity emphasised the intrinsic motion of water particles which could dislodge solute particles from the solid. The solute was then free to diffuse which led to a concentration gradient forming around the solid. The concentration gradient slowed dissolving because it prevented the water molecules getting near the solute. Stirring removed the concentration gradient and hence caused faster dissolving.

Following the role play, two diagnostic questions (DQ15 and 16) were used to elicit students' explanations of dissolving with and without stirring. A large proportion of students (85%) now referred to the motion of water particles interacting with the solute and causing dissolution, a typical comment is given in figure 4.10. About half of these (41%) were able to describe the effect of the concentration gradient slowing the rate of dissolving, a good example of a student response in shown in figure 4.11. This was a very large improvement on the previous attempts and revealed that many were able to use the idea of intrinsic motion and diffusion.

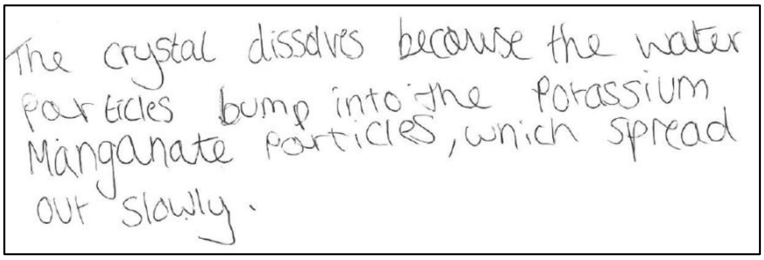

The crystal dissolves because the water particles bump into the Potassium Manganate particles, which spread out slowly.

Figure 4.10 A student's description of dissolving without stirring.

Less successful, was the description of why stirring increased the speed of dissolving. Only about a third of students managed a clear description of how stirring removed dissolved particles from around the solute. A number of students responses (31%) contained the misconception (Ebenezer & Fraser, 2001) that stirring moves the solvent particles faster giving them more energy which speeds dissolving .This is a difficult misconception to counter especially in light of the students' experience of the role play where a faster solvent student would have dislodged more solute-students.

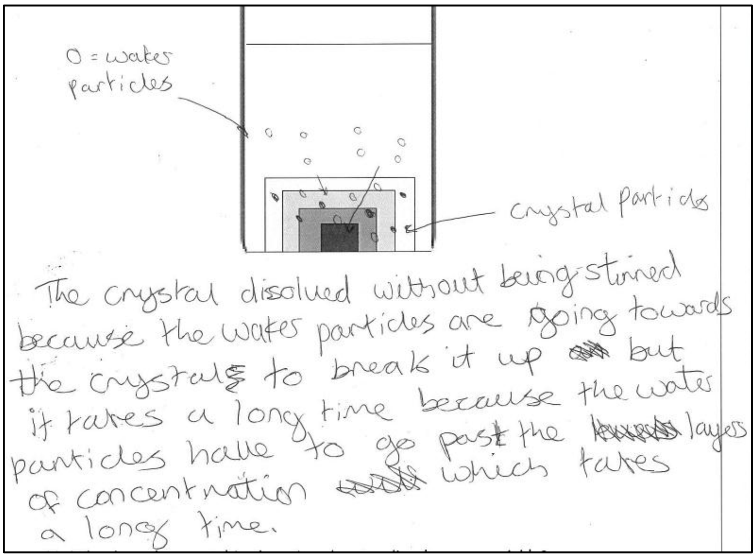


Figure 4.11 A student's description of concentration gradient around a solute.

I resorted to a role-play to spend more time developing the students' understanding of particle behaviour in the dissolving process. Students (should) already know about the motion of particles in liquids and new learning came from applying that to dissolving. Most pupils made the link following the role play.

The idea of the concentration gradient emerged during the improvised role play. While some student-solvents were busy breaking off student-solutes from the block, other student-solvent complained that they could not get to the front. This was used an opportunity to explain the effect of solute near the crystal impeding access of water particles. This is an advanced idea which involves the concept of equilibrium. It was a difficult concept to apply for over half the students.

Most struggled to explain the effect of stirring by removing the concentration gradient. A further misconception arose: that stirring speed up dissolving because the water has faster particles. The macroscopic movement of water and the sub-microscopic motion of water particles may well have become confused by considering the role-play. Following the analogy, stirring the student-solvent particles would certainly have made them move faster.

4.5 Where does the solute go?

This section describes how diagnostic questions revealed students could successfully use a ‘gap’-filling model to explain the conservation of mass but not volume during dissolving. Students were then challenged to account for the phenomenon of saturation using this model. The work led to an attempt to explain saturation using a more demanding inter-particle forces of attraction model.

The activity began with a demonstration published by the Institute of Physics (2004). 200g of salt was added to a 1 litre volumetric flask and the flask was topped up to the 1l mark with water. The mass of the flask was measured and it was inverted several times. The volume was seen to be reduced by about 30ml. When asked to predict if the mass had changed, most students (96%) correctly stated that the mass would remain the same. This was a pleasing finding because students can conceptualise a loss of mass in physical reactions such as dissolving (Holding, 1987).

When asked to explain what had caused the volume to reduce (DQ 18), over half the students (55%) said that the salt particles were sitting in-between the water particles. This view was confirmed using DQ 19 where 74% of students selected and explained a model which best fitted this. Figure 4.12 shows a representative piece of student work.

The responses to this diagnostic probe revealed that many students were secure with this mental picture of solute interspersed with solvent. It was therefore

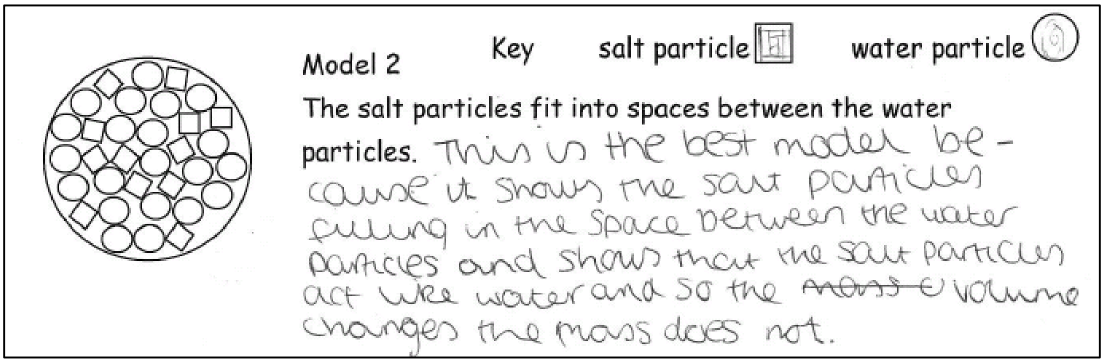


Figure 4.12 Student explanation for selection model showing loss of volume in dissolving.

considered a reasonable idea to extend student thinking later in the project following work on saturation.

Oversby (2000) points out that the ‘gap’-filling model of dissolving has limitations and can lead to misunderstandings. A simple demonstration is often used at the primary level to illustrate dissolving by mixing dried peas and salt (Oversby, 2000). The salt slips between the peas and this is used to model the solute-solvent interaction. He highlights two ideas which cause problems:

- Liquids are incompressible and so cannot have large spaces between the particles.
- Molecules in substances such as sugar are much larger than water molecules making it difficult to fit in the ‘gaps’ and yet sugar is soluble.

A better explanation of dissolving would involve an understanding of the forces of attraction between the particles of solvent and solute.

The project lessons on saturation were designed to assess numeracy skills involved in interpreting graphs of substance solubility at different temperatures (figure 4.13). The general view among the pupils was that saturation occurred when the ‘gaps’

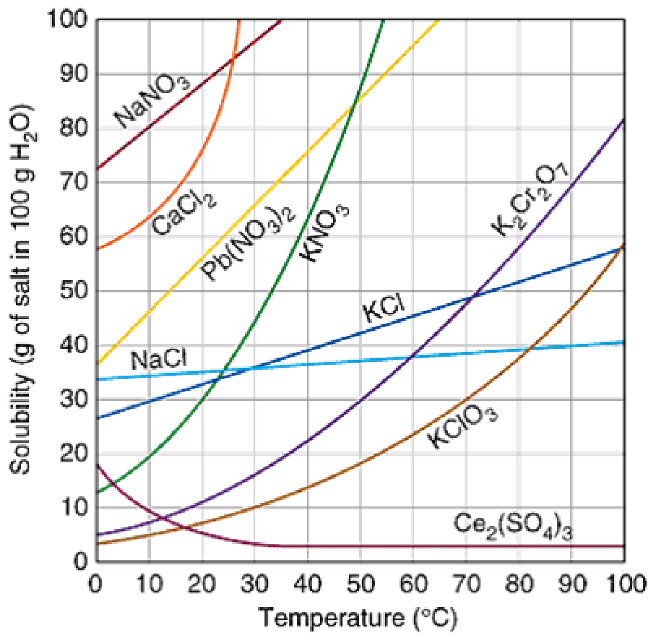


Figure 4.13 Solubility curves of different substances.

between the particles of solvent became filled. During class discussion it was suggested that when a substance was heated, the 'gaps' between particles could expand and this could account for higher solubility because the 'gaps' would be larger. The following questions also emerged from class discussions. If saturation occurred because the spaces in between the solvent molecules were full, should solubilities for different substances not be comparable? Why did different substances vary so much in their solubility? Not only this but why did a substance such as Cerium Sulphate become less soluble as the temperature increased?

To address these emerging questions an additional activity was designed. A diagram was used to focus discussion on how solids were held together by strong forces of attraction (figure 4.14). Solvent particles were able to 'attract' the solute particle away from the solid. In insoluble solids the forces between solute particles

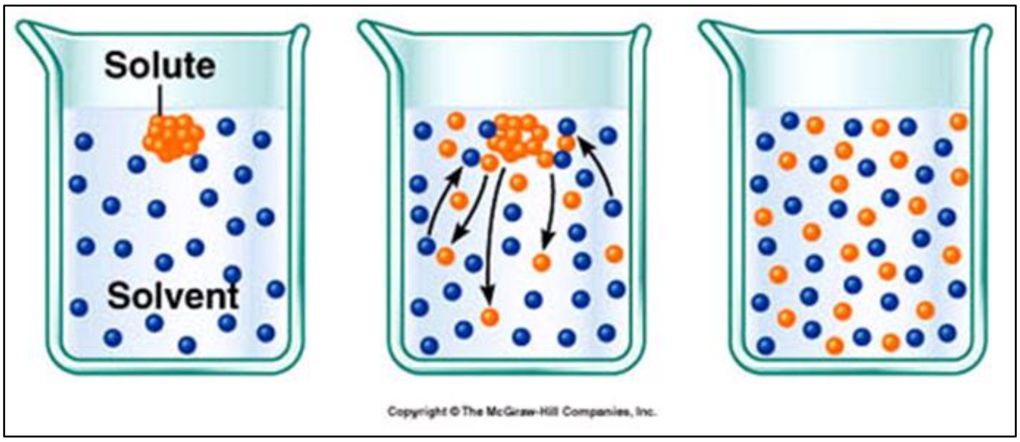


Figure 4.14 Diagram to illustrate break-up of solute particles during dissolving.

were too strong to break. Saturation happened when the solvent particles were no longer able to attract and hold the solute particles away from the solid.

Following this instruction, students were asked (DQ 28s) to use the idea of forces of attraction to explain why:

- Solids dissolve in liquids.
- Adding more solute produces a saturated solution.

In response to the first bullet point, a quarter of students persisted with the idea that solute particles merely mix with the solvent and sit between the 'gaps'. However, 62% were able to demonstrate an understanding of the new idea that the

28s Use the idea of forces of attraction between particles to explain the following:

Solids dissolve in liquids.

The solids have a force and stick together but the water particles have a stronger force and drew them out of the group and start mixing as more water particals drew them in also.

Figure 4.15 Description of dissolving based on forces of attraction.

solvent particles can overcome the forces of attraction between the particles of the solute. This is exemplified in the student answer shown in figure 4.15. While many students acknowledged the limitation of the ‘gap’-filling model to deal with saturation, only a few were able to articulate the more demanding concept of solute-solute and solute-solvent forces of attraction. Responses to the next bullet point revealed that most pupils had reached a limit in their understanding for the moment.

Over half of the group had been able to visualise solvent-solute interaction leading to dissolving, but only 14% could adequately explain how the solvent became saturated using this idea. One student wrote:

> A solvent particle can only connect to so many solute particles before they have no room to connect to anymore. When all of the solvent particles are taken, solid particles are left without being dissolved.

I must emphasise at this point that the above description of the lesson is very far removed from the normal lesson that I would plan and deliver to a Year 8 class. The student responses to the diagnostic questions on the ‘gap’ filling model prompted me to take a decision to extend the work towards a concept I would simply never consider with this year group. The graphs were intended for work on numeracy but they provided a stimulus for questions which intrigued students. Once they realised the ‘gap’-fill model did not fit the data we were on a journey together to try and find a better explanation. This was not seen as extra work by the class.

Although for the majority, the reality was too much to understand for the time being I believe when they come across this concept in later years there will be a greater motivation to engage with it again.

4.6 Do gases dissolve?

Diagnostic questions (DQ 25, 25 & 27) on dissolving and saturation, demonstrated that most students were clear about the idea that solubility of solids tends to rise at higher temperature. This section was designed to introduce and assess students with the contrasting behaviour of gas solubility.

4.6.1 Solubility of gases

The activity began with a simple experiment where students observed a beaker of water as it was heated gently. At 10°C or 20°C intervals, students were asked to record their observations. All students noticed that bubbles of gas formed on the thermometer and on the inside of the beaker. Students were asked to identify the gas and explain where the bubbles came from (DQ 29). The responses were as follows:

The bubbles were evaporated water	25%
The bubbles were hydrogen and/or oxygen	15%
The bubbles were (an unspecified) gas	11%

From these student responses, it was clear further instruction was needed. The behaviour of gas particles and the differences with particles in a solid were reviewed. The main difference highlighted was that gas particles have much more energy than particles in solids and so are further apart and move around much faster. The SEP

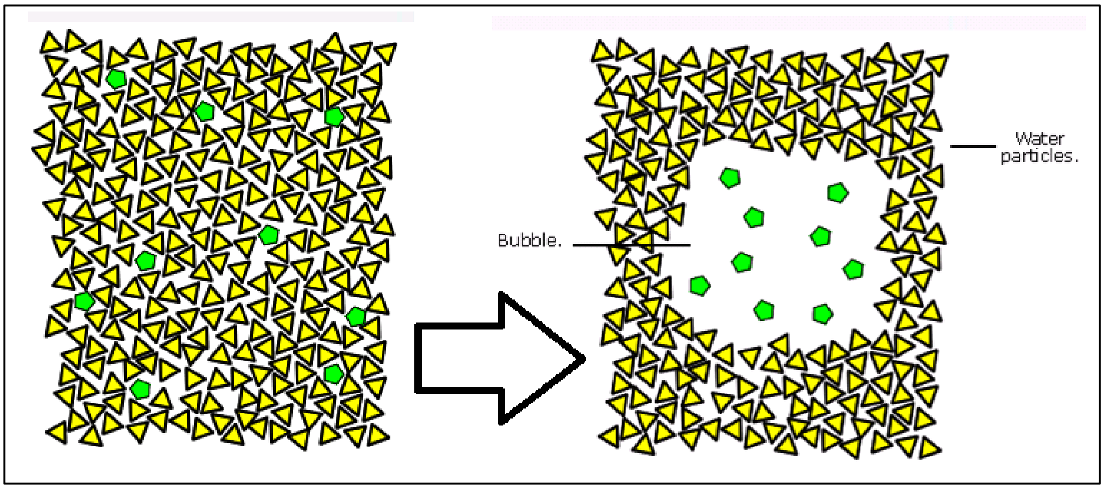
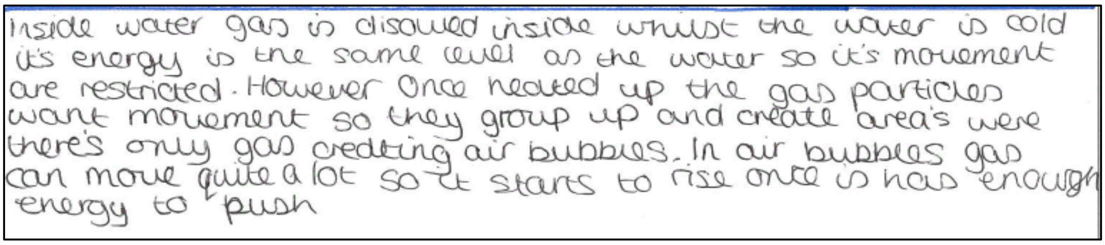


Figure 4.16 Model of gas solution and bubble formation (SEP, 2005).

materials were used to model how gases might be represented forming a solution. From this starting point, students had to consider what happened if the particles of gas were heated. It was important for students to recognise that the gas particles at higher temperature were more likely to form areas where they could move around more. This activity caused the formation of bubbles. Figure 4.16 shows a diagram used to support student understanding of the formation of bubbles in heated water. Following this instruction students were given a diagnostic question 29 again. The responses revealed that students now thought that:

- the gas was dissolved in the water - 71%
- the gas particles move faster at higher temperatures – 64%
- the gas particles had enough energy to escape from the solvent water – 54%



Inside water gas is disolved inside whilst the water is cold it's energy is the same level as the water so it's movement are restricted. However once heated up the gas particles want movement so they group up and create area's were there's only gas creating air bubbles. In air bubbles gas can move quite a lot so it starts to rise once is has enough energy to push

Figure 4.17 Student explanation of bubble formation in warm water.

Figure 4.17 gives a good example of student responses to DQ29 following further instruction. Some of the wording is a little unclear but the idea behind dissolved gas particles gaining energy from being heated and forming a bubble is there.

To probe student understanding four diagnostic activities were used (DQ30). The first question asked pupils to use line graphs to identify the relationship between solvent temperature and the solubility of four gases which 82% did correctly. The next two were questions which asked students to apply their understanding of the relationship shown in the line graphs. Less than half (40%) were able to point out that fish may struggle to live in water at 30°C because insufficient oxygen would be dissolved. Students fared better when asked why fizzy drinks should be served from the fridge and 63% could explain that more fizz required more gas to be dissolved which occurred at lower temperatures.

That more substance dissolves when the temperature of the solvent reduced is not an idea readily apparent to most children. The concept of gases in solution is one of high demand (Johnson & Tymms, 2011). The diagnostic questions indicated that many students had been able to use this concept well. I intended to end the topic on a high with the dramatic Diet Coke and Mentos demonstration (Coffey, 2008). But I saw a final opportunity to extend the students' thinking before we finished. I planned a lesson on gas pressure and devised some new diagnostic questions which were not on the planning list.

4.6.2 – Gas pressure

Students at Key Stage 3 should be aware that gas particles are spread far apart but can be compressed. The GCSE physics course requires a quantitative treatment of gas pressure so a clear mental model at KS3 of pressure and its effects would be desirable. The behaviour and arrangement of gas particles was revisited and discussed.

The diagnostic questions were not prepared on a paper sheet for students but were displayed on the class board:

1. *How does pushing the plunger down change the pressure of the gas and the amount which is dissolved?* using figure 4.18.
2. *Use your understanding of how gasses dissolve in water to explain the information in this graph.* using figure 4.19.

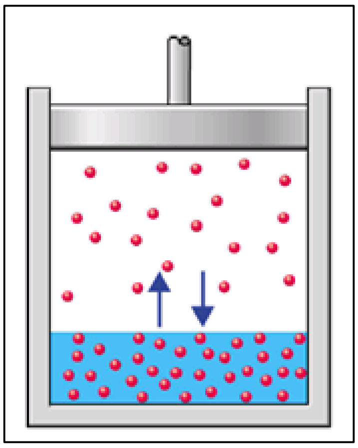


Figure 4.18 Sealed container with gas and solvent (from www.studyblue.com)

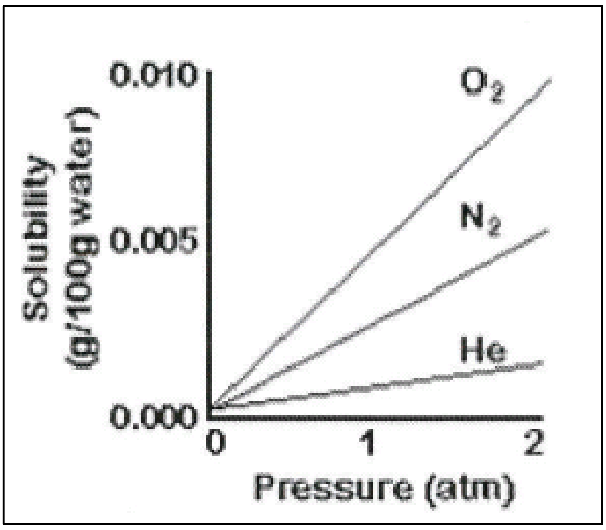


Figure 4.19 Graph of gas solubility against pressure (from chem.latech.edu)

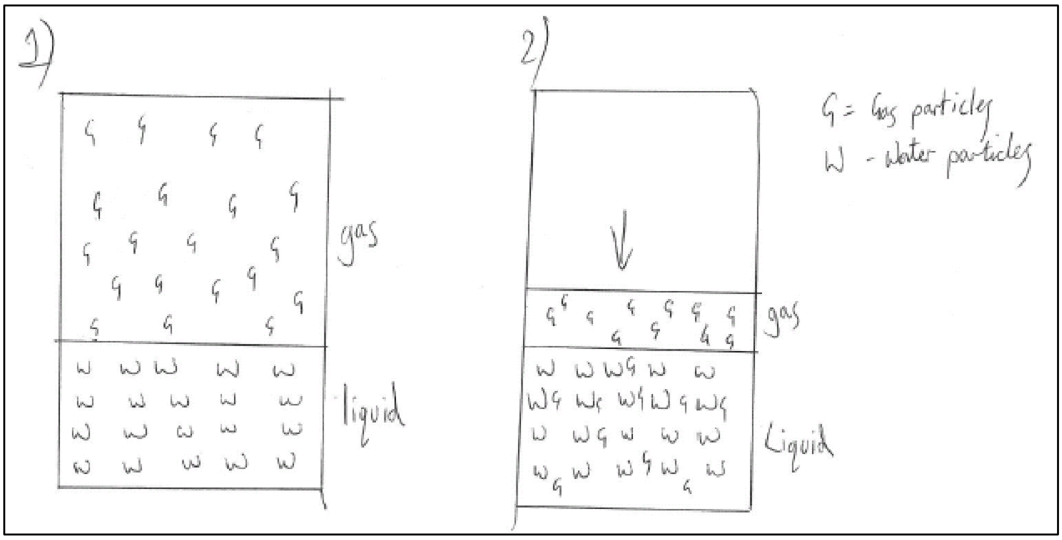


Figure 4.20 Student diagram showing gas (G) dissolving under pressure.

Only a small number of students made the connection between pressure and the quantity of gas dissolved. Figure 4.20 shows one student’s representation of the gas pressure being increased: the gas (G) is clearly visible among the (no-longer continuous) solvent (W) as the pressure was increased. This student explained what happened by saying that:

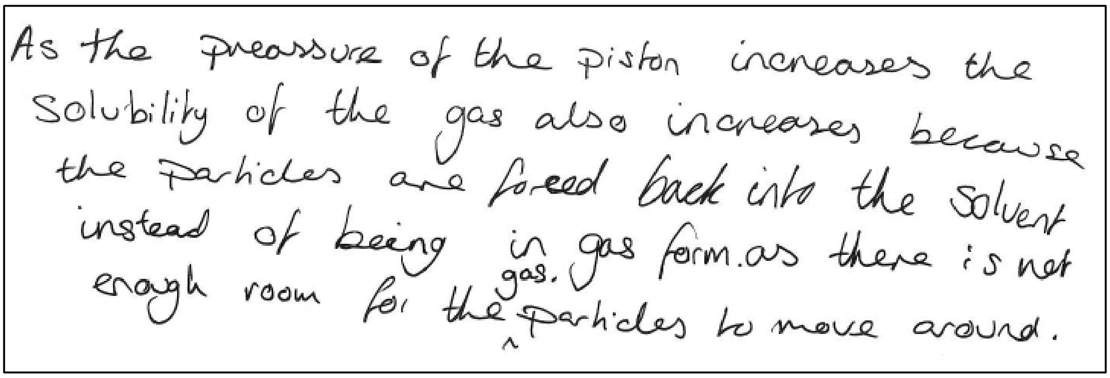
more is dissolved because it is compressed.

The effect of gas pressure on dissolving is a difficult concept because it involves equilibrium. Figure 4.18 shows particles of gas with arrows indicating that they enter and leave the solvent. Although equilibrium applies to saturated solutions of solid solute it has not been encountered by students in this project. The student's explanation of their diagram in figure 4.20 is an accurate description of the relationship but is limited in its detail. Their diagram in figure 4.20 reveals that the student did not think the gas particles enter the water without the plunger being pressed.

I would judge that the student does not really understand the idea of equilibrium based on these responses.

Students fared a little better on a second diagnostic task based on linegraphs of the solubility of three gases against pressure (figure 4.19). While some noticed that oxygen was the most soluble (14%), a third (34%). could describe the relationship between the pressure and the solubility. A third of students explained the relationship in terms of what was happening to the particles of gas as the pressure increased.

Figure 4.21 shows a typical example of a student explanation. It is interesting to see that this student was still thinking about the piston from the previous task.



As the preassure of the piston increases the Solubility of the gas also increases because the particles are forced back into the solvent instead of being in gas form. as there is not enough room for the gas. particles to move around.

Figure 4.21 Student description of gas particles as pressure is increased.

This section of work (4.6) was carried out over three one-hour lessons which I felt was a long time to spend on a single idea. But the progress that many students made was considerable. Following the experiment, the bubbles were identified and in some cases explained (figure 4.17) as dissolved gas coming out of solution, the relationship between gas solubility and temperature was describe and applied in novel setting and then the effect of pressure on dissolving. Again this is very different from how my normal teaching progresses. The diagnostic questions made it possible to gauge how many students were following the teaching and that informed any decision to modify or extend the teaching.

4.7 Discussion of chapter 4

In this chapter I have written about the use of diagnostic questions as formative assessment probes to collect data from students during lessons. I have detailed several short episodes from the lessons to provide evidence of how the diagnostic questions can be used to shape teaching. The findings in this chapter provide evidence that diagnostic questions are sensitive tools which can identify a range of student responses in a number of contexts.

The use of science vocabulary is important to provide precise communication. The diagnostic questions highlighted a need early on in the project to guide students to use specific words to describe dissolving. The change to planned teaching was minimal as short discussions help clarify student ideas.

The use of the particle model in science is a source of considerable challenge for students (Driver et al.,1994). Students' macroscopic treatment of observations was identified and teaching was modified and repeated to pursue the teaching aims again. This occurred several times throughout the lessons. Evidence presented here shows that it was effective in helping students engage with more demanding thinking in their use of the particle model.

A misconception about stirring was identified following one of these repeated teaching episodes. Although the misconception was not addressed in this project the fact that it was revealed indicates the utility of the diagnostic questions. The dilemma to repeat teaching or to move on occurred several times. This is a decision which teachers have to judge on a case by case basis. But having a clear view of student thinking through diagnostic questions would surely lead to better choices.

The diagnostic questions provided a basis for the teacher to extend work towards concepts that are not normally taught to Key Stage 3 students. Where student knowledge was secure, the option was available to advance student thinking, in difficult concepts such as saturation and dissolving and gas pressure. Teaching that 'moves beyond the limit of the curriculum' is viewed favourably by the inspectorate (Ofsted, 2013) (p. 42). If extension activities for individuals had been available, responses from the diagnostic questions would have quickly identified those more able students who would have benefited. The students who struggled were also identified quickly by the diagnostic questions and similarly could have benefitted from alternative tasks.

Shavelson referred to formative assessment on-the-fly (2008) by which he referred to the spontaneous interactions which occur in lessons to clarify student thinking. Perrenoud (1998) said that carefully planned lessons must remain sufficiently adaptable to cater for unsuccessful learning. The improvised nature of some of the activities reported here resulted directly from the information gained from the diagnostic questions. Not only were the impromptu activities useful in meeting learning need, they added a splash of spontaneity to an otherwise constrained curriculum.

Overall the diagnostic questions were effective in providing rich feedback from the students to the teacher. The teacher was able to make decisions quickly and target specific difficulties to move learners on.

Chapter 5

Examining the Learning Progression

5.1 Introduction

RQ 2 asked to what extent the diagnostic questions could be used to indicate student progress along a cognitive hierarchy. This involves using the diagnostic questions towards a more summative purpose. The research approach allows a close examination of what students did well and identified particular ideas students found difficult. The data will be analysed by examining three case studies of student responses to different groups of diagnostic questions which reflect a range of performances across the participants.

5.2 – Organising the data

5.2.1 Student responses on the learning progression

The teaching sequence was devised to present less challenging concepts in earlier lessons with the scientific ideas becoming 'more sophisticated' as the scheme progressed. Hence I expected to find a pattern in the research data indicating that student's correct responses became less widespread as the lessons were taught. Correct responses from all participants to each DQ (see section 3.5.3) were summated to produce a percentage score for each item. A bar chart was constructed to show the percentage correct scores for each DQ in the order in which it was

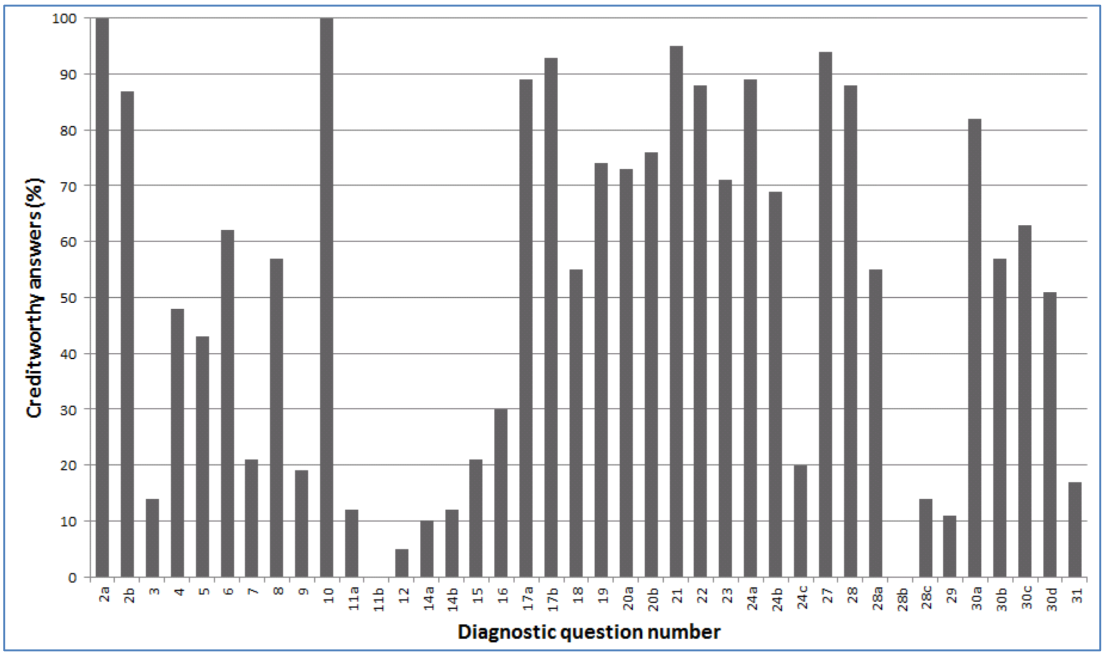


Figure 5.1 At-a-glance view of diagnostic question performance

encountered in the teaching (figure 5.1). It is clear from this graph that the expected pattern has not emerged and that student performance varied from item to item throughout the sequence of lessons.

5.2.2 Three case studies of students' performance

In order to analyse this data and find patterns within it which described student learning, I rearranged the DQs on the x-axis (figure 5.2). The items were set out in order of percentage score to produce a pattern of attainment. The DQs were placed into three groups which received roughly similar scores. The first group (high scoring items) ranged from 100% to 69%, the second group (intermediate scoring items) ranged from 68% to 40% and the third group ranged from 30% to 0%. The boundaries between the groups are arbitrary, especially between the high and intermediate groups, but this approach facilitated an analysis of student performance.

The use of the DQs does not just provide a summative score of correct

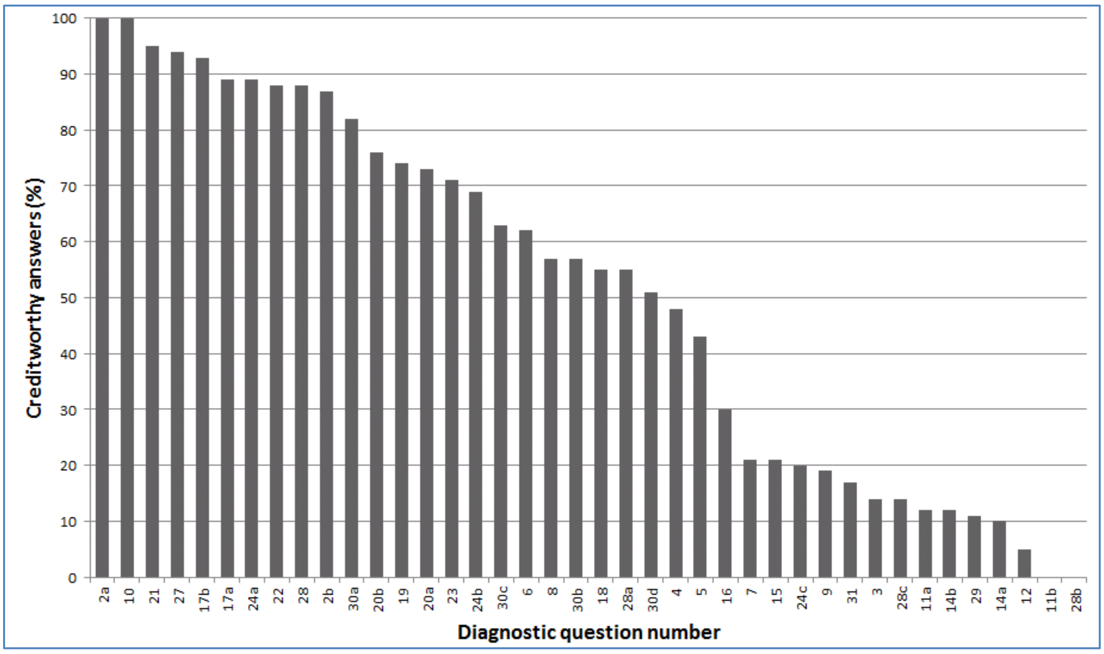


Figure 5.2 Sequenced view of diagnostic question performance

answers but reveal what students can and cannot do in the learning activities. To present examples of the types of responses participants gave in each of the three groups of diagnostic question, I propose to conduct case studies of three students. Each student was selected on the basis of being representative of students in an upper, middle and lower ability range. NfER CAT scores and KS2 English and mathematics SATs grades were used as criteria to assign students to each band. Student A is upper ability (CAT score 120, KS2 SAT: mid to high level 5), student B is middle ability (CAT score 110, KS2 SAT: low level 5/high level 4) and student C is lower ability (CAT 97, KS2 mid-level 4).

By analysing the data in this way I have attempted to build a profile of how students of different ability performed in response to a range of questions of varying challenge.

5.3 Analysis of students A, B & C responses on high scoring items

5.3.1 Recognising dissolving in water

At the start of a topic on solutions it was necessary to establish important key terms which signalled a substance had dissolved. Students were given a table of five substances and a description of each after they were added to water (DQ 2a & b). Both student A and B correctly identified the terms *transparent* and *colour change* as descriptions of two substances which formed solutions. In explaining their choice one student wrote:

> Student A: Substance c became clear. Substance b went transparent and became a clear blue colour.

Student C recognised dissolving had occurred and selected substances b and c but did not refer to the terms transparent or to the colour change to explain their choice:

> because the substance has disappeared and is not visible as it has mixed with the liquid.

The discussion with students about better alternatives to terms like 'disappeared' is covered in Chapter 4.

This item was at the start of the teaching sequence and it was expected that students would perform well.

5.3.2 Recognising salt is the residue from evaporating saline

Another activity in which the three students scored highly was the identification of salt formed on a watch glass after the evaporation of salt solution

(DQ10). The question was preceded by a practical activity heating saline solution which students interpreted well. Only student C did not name the white substance as salt but wrote instead:

clear crystals appeared on the watch glass.

The reflective diary records that during later discussion it was accepted that the white crystals were salt by the few students who had not explicitly stated it.

5.3.3 Conservation of mass but not volume

The learning intention in lesson five, near the middle of the teaching sequence, was that during dissolving, mass is conserved but volume is not. This is described as a more demanding concept by Johnson and Tymms (2011) and by NC levels. In addition, Holding (1987) reported that many students believe that mass is lost during dissolving.

Students were asked to predict and observe the changes in mass and volume during a demonstration where the mass and volume of salt and water were measured before and after mixing (DQ 17 a & b). Student responses indicated that they had made the distinction between the behaviour of the two quantities. Their observations were as follows:

Student A
a) mass - It didn't weigh any more or less. It stayed the same mass.
b) volume - The water level went down
Student B
a) mass – the mass of the salt and water decreased by 1g during the experiment
b) volume – the total volume decreased at the start it was at 1000ml and at the end it was well below.
Student C
a) mass – during the experiment the mass of the solution decreased from 1314 to 1313g
b) volume – the volume went down from 1000ml to less

There was a question as to how significant students B and C thought the loss of 1g was. The reflective diary records that in discussion which followed the activity, all students expressed agreement that the mass remained constant.

A follow-up diagnostic question asked students to select from four possible models to explain the loss of volume during dissolving (DQ 19). Students had to recognise that the most suitable choice was one which described the salt fitting into 'gaps' between the water molecules. Three distractors were based on alternative concepts. Students A, B and C recognised the correct model.

5.3.4. Recognising non-aqueous dissolving

Students carried out a practical where they mixed organic and inorganic solute and solvents. A simple picture model was used to visualise the results obtained (DQ 21). Students were asked to use letters to represent particles of substances such as ethanol, oil, and water. Figure 5.3 shows how student B completed the task. Oil (O) floats on water (W), Water and ethanol (E) mix as do oil and ethanol and the

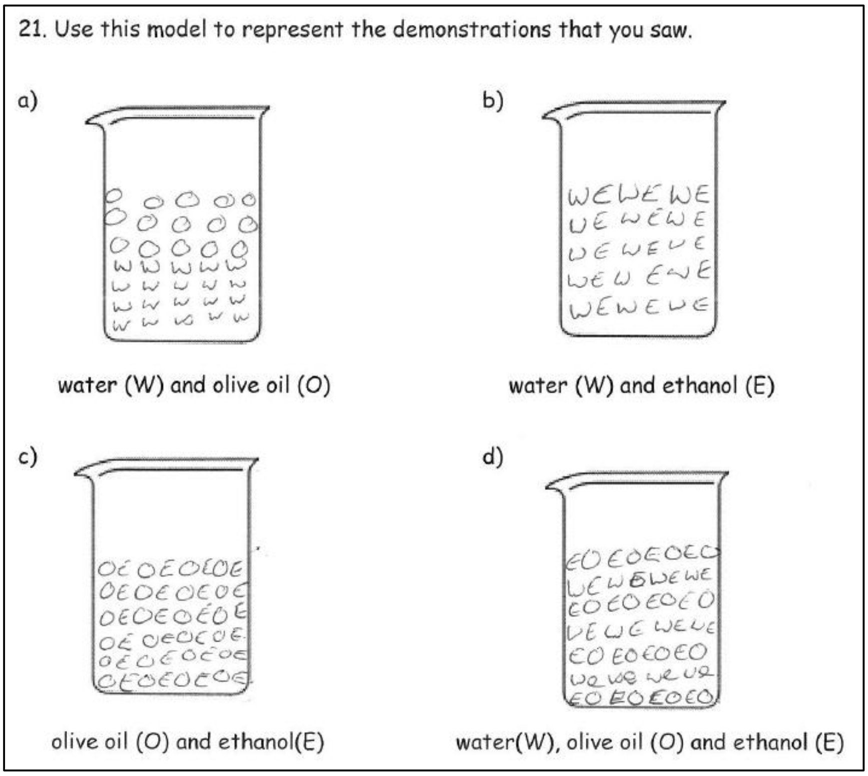


Figure 5.3 Student B representation of different combinations of organic substance and water

emulsion formed with all three substances can be represented using a combination of the letters. The work from students A and C was similar. In the two-part question that followed (DQ20 a & b), students needed to consider the use of the term insoluble in a non-aqueous context. Student A and B responded that paint was not ‘insoluble’ because it could dissolve in white spirit and ethanol was not a suitable ‘solvent’ for nail varnish because their particles would not mix. It was not clear how student C transferred an understanding of the terms insoluble and solvent to this new setting. Student C’s response to DQ 21 was correct but to DQ 20 selected:

a) paint was insoluble because did not dissolve in water, and
b) that nail varnish is too heavy to dissolve in ethanol.

Although the use of the terms in a novel situation caused student C some difficulty, the completion of the diagram model demonstrated some transfer of understanding to a non-aqueous context.

5.3.5 Basic mathematical operations

Students carried out an investigation into the rate of dissolving of jelly cubes cut into smaller and smaller pieces. They were then required to draw line graphs to show the rate of solubility with surface area using experimental data (DQ 23). The skills involved in selecting scales and interpreting the patterns shown are considered level 6 in the National Curriculum which would suggest a high level of demand.

All three students drew accurate line graphs.

Student A described the relationship between surface area and the speed of dissolving (DQ 22) by saying:

> [The] more surface area the faster it dissolves. The more surface area the easier it is for the water particles to break down the solute.

Student A and B offered an explanation based on their model of the dissolving process which involved the idea of a larger area of solute on which the solvent can act. Student C described a similar relationship but explained it in terms of the jelly being broken up:

> It would take less time to dissolve the same mass if it is smaller because if it is smaller it would be quicker to break it apart.

Two questions (DQ 24a & b) required students to read the line graph to estimate an x value for y and vice-versa. Student C did not record any clear responses to these questions. Students A and B both provided accurate answers.

Reading line graphs was also asked in DQ 27 on solubility. All students were successful in determining correct values for solubility and temperature when provided with the other variable. Moreover, all three students were able to use the procedure below to calculate the solubility and mass of solute in a series of problems.

$$\text{solubility (\%)} = \frac{\text{mass (g) x 100}}{\text{volume of solvent (ml)}}$$

5.3.6 Summary of section 5.3

The lack of an expected pattern as discussed in section 5.2.1, leaves a question as to how these results can be understood. The literature review surveyed several frameworks for measuring conceptual understanding. In this and in subsequent summaries I will refer to appropriate research-based methods of categorization in order to interpret and make sense of the data.

In this section, the three students demonstrated facility in the activities and diagnostic questions they undertook.

They identified dissolved substances by recognising and recalling that solutions are transparent even if coloured. This was an expected result. This activity was the first in the teaching sequence and therefore intended to be the least challenging. In terms of the Bloom taxonomy recall and recognition are the lowest levels of cognition.

They recognised that salt form evaporated saline which is consistent with Johnson and Tymms (2011) who scored this as 34 on their scale 0-84. Also a Piagetian interpretation would describe the conservation of mass of the salt as early concrete: 2A (Shayer and Adey, 1981. p82).

Some of the activities occurred later in the teaching sequence. Both the loss of volume and use of organic solutes/solvents required the use of a simple particle model. In Bloom this would be at the explain level as it requires a cause and effect relationship to be understood (Anderson & Krathwohl, 2001, p. 67). Describing a simple model of novel organic solutes and solvents (DQ 19) involved recognition of the correct diagram.

The use of an equation to perform a calculation would be classified by Bloom as application as it is the executing of a procedure (Anderson & Krathwohl 2001, p. 67). Shayer and Adey (1981) classified simple mathematical operations as late concrete 2B (p. 76). Choosing scales and interpreting data in graphs is ‘levelled’ at 6/7 in the National Curriculum (high performance at the end of Year 9, DfEE/QCA, 1999, p. 7). The responses to the activities presented here suggest that students were not confronted with excessive challenge.

The observations made by the students were also macroscopic in nature (De Jong & Taber, 2007) which would also suggest that they were less challenging.

5.4 Analysis of students A, B & C responses to intermediate scoring items

5.4.1 Identifying dissolved solute

When asked to indicate ways a sugar solution could be identified from a list of five possible options (DQ 4), students A, B & C made the following choices:

- student A: taste, evaporation
- student B: taste evaporation, filtration
- student C: colour of solution, evaporation, filtration.

This is a first instance in these findings where student answers were not entirely based on observable results. Students A and B noticed that the option taste would identify dissolved sugar. This response could be attributed to their own experience. The other options provided required students to use their own mental picture of dissolving. All three students recognised that the solution could be evaporated to reveal solute, but B and C incorrectly selected filtration which show that there was some error in their thinking about either the process of dissolving or filtration

5.4.2 Filtering salt solution and chalk suspensions

Following the selections made in section 5.3.1, the students filtered a salt solution and a chalk suspension and were asked to explain the results (DQ 5). Students A and C struggled to interpret the results they found. Their answers showed they were unable to distinguish or identify the materials left on both filters. Student B

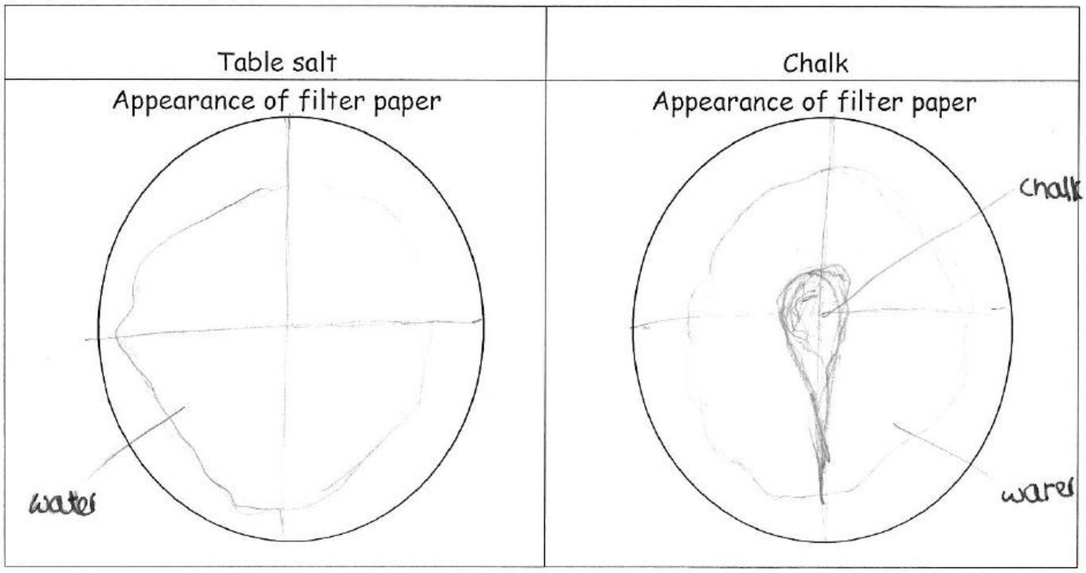


Figure 5.4 Student B drawing of filter paper

had more success and drew the diagram shown in Figure 5.4.Student B wrote the following explanation:

> The chalk one still had stuff on it because it hadn't dissolved so it would separate in the filter paper whereas with the salt one nothing is shown because the salt has dissolved into the water so it can't be separated.

Student B identified chalk deposited on one filter paper but the other had no solid on it. The description was based on observation. The response attempted to explain the differences between the salt and chalk by saying the salt had dissolved and the chalk had not. Student B did not explain how the particles were involved in causing this outcome.

This question required the students to move from an observational response at a macroscopic level to one which required an understanding of what is happening at a sub-microscopic level (De Jong & Taber, 2007). This type of question seems to differentiate between students who can explain their observations in terms of a particle model and those who cannot.

5.4.3 Explaining non-conservation of volume

Students had observed the conservation of mass but loss of volume during a demonstration. All three students selected a correct response to a multiple choice question (DQ 19) on how volume is lost during dissolving (section5.3.3). In DQ 18, the students were asked to explain the loss of volume in their own words. Student A wrote:

> some molecules fit between other molecules.

The student correctly used the term molecules unprompted. Although they did not directly link the loss of volume to the molecules 'fitting' together they appeared to show understanding.

Student B wrote:

> because the salt had moved into the gaps in the water and was no longer a solid.

Student B had an idea of 'gaps' existing in the solvent. Student B focussed on this particular feature to explain the reduction in volume but used slightly confused terminology in referring to the salt as no longer being a solid. Student C tried to use the idea of particles fitting together but was confused by a technical aspect of the demonstration stating:

> the released air bubbles would have moved the particles around into tiny gaps.

The responses of the three students to DQ 19 suggested that they had a clear understanding that solute fits in-between solute particles. Having to express this idea in their own words revealed that this idea was easier to recognise than explain.

5.4.4 Applying ideas using gas solubility

An item towards the end of the teaching sequence required students to transfer their thinking from solid solute to gas solute. Students A, B and C had been able to describe the relationship of gas solubility and temperature using a line graph as opposite to that for many solids (DQ 30a). The ability to apply an understanding of this relationship was tested in two questions

30 b) Why would a fish find it difficult to survive in water at 30°C?

30c) Why are fizzy drinks better when they are served from the fridge?

Student A wrote:

a) there wouldn't be a lot of oxygen in the water.

b) They are colder which means that it is more soluble meaning that it has more gas in it.

The answer to part a) is creditable answer but it is only through reading part b) that it is evident that student A made the link between greater gas solubility and lower temperatures. Students B and C did not seem to grasp this relationship.

Student B wrote:

a) less oxygen it escapes through the bubbles.

b) it is better because the drinks have been in a cold area so as it comes out the room temperature hits it so the gases its heated which makes it fizzy.

Student B's first response was based on the experimental observation that bubbles formed when water is heated in a beaker and not on the relationship shown in the line graph. In the second response, student B linked the fizz in the drink to the sudden increase in temperature when the fizzy drink is removed from the cold.

Student C wrote:

- not enough air
- there will be more air in the drink so it will be fizzier

Student C generalised the gases in the water to be air. Although they made a quantification of the amount of gas in each scenario they did not link solubility and temperature. This is an example of where a markscheme would have penalised students in this study (see section 3.5.3). Student C's ideas were expressed as well as could be managed and some meaning was put across. But a markscheme which looked for specific terms or phrases would probably have awarded little credit.

5.4.5 Summary of section 5.4

The activities in this section required more abstract thought from the students. Tasting the sugar in a solution was within the experience of many students but predicting the effect of evaporating or filtering a solution required students to see the solution in a different way. Student reasoning needed to move away from macroscopic descriptions to explanations using a sub-microscopic mental picture.

Student B explained the chalk was filtered because it was not soluble. Shayer and Adey (1981) describe this as late concrete thinking (2B) because it is linking a cause and effect. If student B had discussed how particles caused the observed effects this would be classified as early formal thinking (3A) (Shayer & Adey 1981, p. 81). This was seen in student A and B's explanations of the loss of volume demonstration. The Sunflower (2006) and SEP (2005) resources were used to help students' visualisations but transferring images of models from animations and pictures into operational thinking is difficult for students (Naah and Sanger, 2013).

Some of the activities could be analysed in the terms of Bloom taxonomy. When explanation was required over recognition in the loss of volume activity, students scored less well. The question on gas solubility tested students' ability to apply the knowledge gained from the graph.

The construction of the line graph involved setting out an axis which is an example of proportionality. Shayer and Adey (1981) classify this as early formal thinking -3A.

The activities presented more challenge to the three students. Student C did least well in terms of a score but the diagnostic questions reveal some indication of what thinking is going on. There is no link visible between these activities and their order on the teaching sequence. The level of difficulty does seems to link to the move away from macroscopic thinking and the lower levels of Piagetian and Bloom'ian' categories.

5.5 Analysis of student A, B & C responses to low scoring items

5.5.1 Explaining the filtration of chalk but not salt solution

Students had examined the results from the filtration of salt solution and chalk suspension. Some students failed to make accurate observations or to link the observation to the behaviour of particles (section 5.4.2). This activity was repeated and student drew a diagram of the process of filtration at a particle level. The students were asked how the results would be different if the holes in the filter were made smaller (DQ 7). The responses are shown below:

> Student A: I think the results won't be different because the water and salt can still pass through as they're the same size
>
> Student B: Because the filter paper holes are small so if the salt crystals were small it would be held up in the smaller holes
>
> Student C: no response

Student A used an image of the water and salt particles being similar in size and did well to notice that if water could pass through the holes in the filter then so should dissolved salt. The diagram drawn by student B (figure 5.5) shows they grasped the idea that there is something about the difference in particle size which caused the effect on the filter paper. Student B did not understand the true scale of the relative sizes of the filter paper and the water and salt particle: from their drawing it is easy to see how salt particles could be caught in the filter if the holes if the filter were slightly smaller. Student B missed the model's prediction that the water particles would also be blocked.

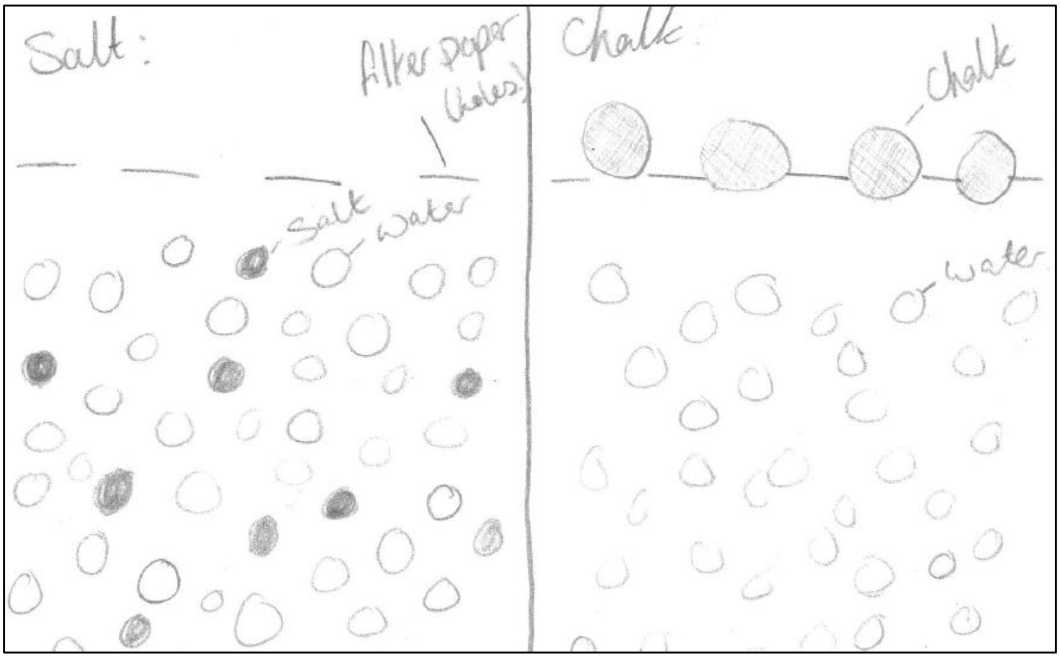


Figure 5.5 Model used by student B to represent salt water and filter paper

5.5.2   Explaining Saturation

The ‘gap’-filling model of the interaction between solute and solvent was used as a model to account for the loss of volume during dissolving (DQ 19).However it is inadequate to account for why solutions become saturated. A better model to explain saturation would refer to the forces of attraction between particles (DQ 28s, part c). Student B wrote:

> Adding more solute eventually produces a saturated solution because the water particles struggle to find the solute as the other solute particles block the way for the water particle to reach.

Student B used the idea of a concentration gradient from another question as slowing down dissolving. Students A and C were not able to provide any answer to this question.

5.5.3. Explaining the effect of stirring

Another example which required students to manipulate a mental model was the explanation of a solute dissolving without stirring (DQ15) and how stirring caused dissolving to occur faster (DQ 16). Students had observed a demonstration of potassium permanganate pellets (crystals) dissolving in stirred and unstirred beakers of water and had taken part in a role-play of the process before attempting their explanations. The teaching aims were to realise that: a) the *intrinsic motion* of water was sufficient to cause dissolving and b) the *concentration gradient* which built up around the solute could be removed by stirring and allow the water better access to the solute.

Student A wrote:

> DQ15.The water particles are hitting the crystal so that it breaks into bits which help water particles get in between so this is how the crystal will mix without being stirred.
>
> DQ 16. when you stir it you help it break apart so the water particles mix

In the first explanation, student A made a connection between the intrinsic motion of the water particles and how this physical interaction removes solute from the solid. Student A did not use the model to explain stirring but referred to the mechanical action of stirring water breaking up the solid.

Student B wrote:

DQ 15. The crystal particles spread out but because it is more concentrated the water particles cannot get through to the main crystal

DQ 16. It dissolves quicker because you are moving the crystal particles around letting the water particles get to the crystal particles helping it dissolve quicker

Student B did not refer to the interaction of solvent and solute causing the solute to dissolve but ‘saw’ the solute particles spreading out in the solvent. Student B did not use the idea of intrinsic motion explicitly but described its visible effect. This student used the idea of a concentration gradient slowing the solvent’s action on the solute. But because the interaction between solvent and solute is not clear, the effect of stirring was not fully explained.

Student C wrote:

DQ15. the crystal dissolved slower than being stirred and the colour of the crystal started to spread across the beaker.

DQ16. It dissolves faster because the water particles are moved around easier and they hit the crystal particles more often.

Similar to Student B, the word ‘spreading’ is used to denote the diffusion of the solute. Student C’s response to DQ 16 suggests that they were aware that the water and solute interact in some way but their use of the model is limited. Student C offered a simple relationship between stirring and the rate of dissolving or ‘spreading’. This student did not demonstrate a view of intrinsic motion of the water but described their observation of the demonstration.

5.5.4 Comparison of evaporation with crystallisation

Holding (1987) and Driver et al. (1994) have reported that students often use the word ‘melt’ to describe the dissolution of solid in water. Being able to compare the differences or similarities between changing state and formation of a solution would provide good information about students’ ability to use the particle model. One of the practical activities involved the evaporation of salt solution leaving solid crystals in a watch glass. I felt this was a good opportunity to assess pupils’ thinking on how the formation of a solid following evaporation might be similar or different to the formation of solid during freezing. The practical activity was followed by DQ 11 parts a & b:

11 a) In what way is this process similar to freezing?

The responses of the three students are shown below:

Student A: the process changes the water: water → gas same as freezing the water changes water → ice.

Student B: they are both solidifying solids.

Student C: because we are turning liquid back into a solid.

A student who understood the particle model well might have responded by describing how in both cases the solid particles form a regular pattern and hold each other with strong forces of attraction. Student A noticed that freezing and evaporation are changes of state but misunderstood that the question was asked about the salt which formed. Students B and C did refer to the formation of salt but did not use the particle model. Student C also made the error of comparing a liquid solution mixture) to a liquid (single) substance.

The second part of the task reversed the question:

11 b) In what way is this process different to freezing?

The student responses were as follows:

Student A: Uses heat instead of cold temperature

Student B: Heating instead of cooling. Removing the water to make it a solid whereas when you freeze it you keep the water – just frozen

Student C: Because we evaporate the water.

My expected response here would have referred to particles of the same substance losing energy during freezing compared to solute particles re-joining upon loss of a solvent. Despite the question referring to the salt, student B made an attempt at describing liquid water becoming solid during freezing perhaps because it is easier to conceive of water freezing than salt. Both students A and B pointed to freezing happening at cold temperatures which reflects a limited experience of a range of substances' melting and boiling points. All three students found it difficult to use the particle model to make the comparison asked by the question.

5.5.5 Recognising dissolving when solute is still present.

Another question all students had difficulty with was whether they could account for solute still being present at the bottom of a beaker following addition of solute and stirring (DQ 3). The question carried four options and asked for a reason for choosing the option. The correct one prompted students to recognise that a solution is clear even with undissolved solute. All three students selected the wrong choice but gave slightly different explanations:

Student A: The sugar did not dissolve because the pupil did not stir it enough.

Student B: If he had stirred it for longer, the sugar would have completely dissolved.

Student C: As he just left it, it sank to the bottom

Both students A and B pupils missed a cue in the question that a *large* amount of sugar was added and C missed the information that the sugar had been stirred. But I do not believe details contained in the question were the main obstacle to answering this question correctly. That some solute had dissolved while some remained visible seems to have caused students difficulty.

5.5.6 Using a line graph to calculate gradient

In a previous section most students had correctly used line graphs to read off values and many had correct constructed and interpreted linegraphs. DQ 24c required students to use a line graph to estimate the increase in surface area which would halve dissolving time. This is a more complex task where two measurements need to be taken on both axes. Student A gave an incorrect response:

double the surface area to make the jelly dissolve faster

Students B and C did not give a response to this question.

5.5.7 Summary of section 5.5

The diagnostic questions reported this section proved challenging for our case study students and for all the participants. The activities required a substantial degree of abstract thought to explain observed phenomena using and manipulating the particle model.

Predicting that dissolved salt would be unaffected by smaller holes in a filter paper and describing how stirring removes a concentration gradient to increase the rate of dissolving involved sophisticated mental imaging. This is formal thinking. Students had to imagine, in DQ3, that some solute had dissolved because it could not be seen. Shayer and Adey (2011) class this as early formal (3A) (p. 31) because it is an indirect interpretation of what they saw. Saturation is classed as late formal (3B) because it requires multiple variables to be considered in describing the interplay of intermolecular forces which result in saturation (Shayer & Adey, 2011, p. 93). The use of the graph to calculate gradient also uses multiple variable and requires formal thinking.

In terms of the Bloom taxonomy, the thinking required by these tasks is high order. Solid formation from evaporation and freezing is a comparison where students need to analyse the different conditions and tease meaning out of the information they are given. They must then integrate ideas to synthesise a new possibility which they have not encounter before.

5.6 Discussion of Chapter 5

Figure 5.2 shows a clear trend in the data collected by the diagnostic questions when they were assessed by the teacher. Careful consideration was made to the construction of the teaching sequence and this was not the pattern which was expected. By analysing the trend that did emerge, it was possible to group tasks which differentiated student facility.

The case studies were carried out to simplify the task of summarising and reporting results. Each student roughly represented a level of ability within the group. Over the three sets of diagnostic questions their responses were interpreted in terms of the conceptual frameworks of Bloom and Piaget. But this begs the question 'Now what?'

RQ2 was concerned with gathering evidence to measure progress. The diagnostic questions provided a rich source of information and revealed much of the thinking that was going on in students' minds. How to summarise this data becomes a task in itself. The diagnostic question data can be used in a variety of ways.

At one extreme, each of the 31 questions could be allocated a score and the total score of correct responses given by the student could be given as a number or grade. This is very similar to testing and reduces knowledge to scores which do not tell us much about students' capabilities.

By looking at an overall performance indicated by the diagnostic questions, a student could be placed into a category of Bloom or Piaget. This would be similar to what is written above. Instead of a score or grade, a child would receive a category this too would give very little information about what they could do or what they understood.

An alternative might be to construct a grid of the curriculum content contained in the teaching sequence. This could be completed by the student based on their performance in the diagnostic questions. Perhaps a code, 3 for a strong performance and 0 for inadequate performance could provide an overall profile across the teaching. This approach is similar into a method reported by

McArthur(1978) cited in Leahy and Wiliam (2011) called the S-P technique. The important feature of this assessment is that the student has information be able to reflect on their learning.

I stated in chapter three that the purpose of RQ2 is not to provide a summative score but to interrogate the information provided by the diagnostic questions to see if the information they provide can inform a summative judgment. In a sense the use of diagnostic questions pulls away from the tendency to simplify student performances as a single grade or level. Providing student with detailed feedback in summative judgements in a range of curriculum-related tasks is something that diagnostic questions may make a possibility.

Chapter 6

Conclusions

6.1 – Introduction

In this final chapter the finding will be summarised and briefly discussed in the light of the research question. An evaluation of the study will be done where the relationship between the findings and the research strategy will be examined. Some final thoughts are included in the final section about the possible implications of this research.

6.2 – Summary of the main findings

Research question 1 (RQ 1) asked to what extent the diagnostic questions could be used to manage the collection of accurate, real-time data about pupil understanding to shape teaching. The results presented here provide evidence that using diagnostic questions is a highly effective means of collecting data in a busy classroom setting. The diagnostic questions focussed the assessment firmly on the learning intentions. They were sufficiently sensitive to detect misconceptions, incorrect use of vocabulary, and low order thinking among students. The student responses provided a rich source of assessment data which could be processed quickly by the teacher and provided a rich snapshot of student thinking. Several scenes were presented which detailed the use of student data which the teacher had acted on to change the course of the instruction often at very short notice. The response from the teacher was occasionally improvised but the quality of the assessment data kept the learning on track.

Research question 2 asked to what extent the diagnostic questions could indicate a performance on a cognitive progression. The student responses were analysed and showed that the design of the cognitive demand did not produce a pattern which could be interpreted in terms of progression. Student responses were compared to the developing stages of Bloom and Piaget where a link could be drawn between the conceptual level and the degree of success observed in answering individual items. Students tended to do less well in questions where thinking involved abstraction or higher order thinking. The quality of the data provided by the diagnostic question was considered as the starting point for a style of summative assessment. The level of understanding across a range of content could be indicated by the performance on diagnostic questions and used to provide a more useful measure of summative feedback.

6.3 – Evaluation of the methodology and strategy

It is important that I reflect on the limitations of the design of the project at this stage of the project. I stated in chapter three that this type of study was a goal-free evaluation (Scriven, 1991). This involves not just looking at the intended consequences but also the unforeseen effects and side effects.

The scope of the research was quite ambitious. When I first envisaged the project I planned a scheme where Bloom's taxonomy would be used to plan a teaching sequence and students would be measured as they progressed along it. The realities of carrying out good research I understand now would make this a huge undertaking. Assessment is vast field to get to grips with in the short space of a Master's research program.

The studies which examined the misconceptions in the chemistry literature tended to focus on one concept. This was not a realistic option for me. As a classroom teacher I had to maintain a level of teaching appropriate to the needs of my Year 8 students. I feel that there were aspects of this project that stretched the focus of the research. The materials took a long time to research and develop and I feel this detracted from the development of a good research approach.

How good was the learning progression? A considerable amount of work was put into developing the teaching sequence. Duschl et al. 2011 (section 2.4.5) view learning progressions as work which should be based over many years. The lack of an expected result in terms of the learning progression should not then come as a surprise. Carrying out a short intervention such as the one in this study does not lend itself to measuring long term effects. The constraints of the project coupled with my own distorted view of National Strategy teaching (that progress can happen in a lesson) meant that that progress was never really going to be measured.

The findings in chapter 5 could be questioned for their validity. The analysis of the data for this research question was carried out by grouping question responses together on the basis of their score. This may have distorted the appearance of the findings.

It is important to note that this study was carried out with a sample of 60 children and in one area within science which depended on student ability to use the particle model. What can be taken from the findings of this work are given above but it is important not to generalise. The same pupils may find other areas of science personally more interesting and this may lead to different levels of performance.

Time limited the study to one cycle of an action research method. The question remains if the results would be different if the materials were modified. Teaching the materials in the project produced a lot more of what could be regarded as data that just the diagnostic question results. What would happen if certain types of feedback were used? What would be the difference if the DQs were repeated in silence? What effect would introducing an aspect of self or peer assessment have? These questions were raised but could not be answered. A longer project would have been useful to address further questions in subsequent action research cycles.

The data analysis was carried out by me only. An action research project would be better with an element of collaboration but I was keen to start and did not organise this beforehand so it was not part of the project. This also prevented any moderation of the student responses. Questions remain as to what another teacher would have thought creditworthy. Would using the DQs with another class produce results which would similar conclusions? The level of demand in some of the questions in retrospect seems quite high. During the lesson it felt appropriate. Good relationships with students and their willingness to pursue learning was a good mix in the classroom. But when it came to reading the answers it was clear I was sometimes getting a little carried away

Also I cannot be certain how much of the material I collected from students was entirely their own work Students sat in groups of four and were encouraged to talk to each other. They did not complete the DQs under test conditions and I did not expressly warn against plagiarism or helping each other. I recorded in my reflective diary that the forms of words students were writing were sometimes very similar to what I or another student had said earlier in the lesson. Was this understanding or just a child with a good memory?

But would the project have made a difference? What performance would be different in the students who were taught by me compared to those in another class if they were to be tested and compared? I commented to other staff during the project that the real benefit of working with the students this way would only be felt in year 10 when they had to learn the topics on the GCSE curriculum that depended on a good understanding of solution chemistry.

Has it affected the motivation of students? Many of the issues associated with summative testing in chapter one spoke of the poor practice resulting from teaching to the test. Has the use of diagnostic questions gathered all this data and still enabled lessons to be all about learning? If student motivation was measure in terms of engagement with the task I could point to the extent to which students completed written tasks. Very few were left blank. The DQs were designed with large blank spaces Student took full advantage of the opportunity to engage in extended writing – a fine response to Ofsted's concern that students are 'set tasks that do not allow them to reveal their understanding of a science concept or idea.' (Ofsted 2013, p32) Providing small spaces in future work may help focus students' ideas before they put pen to paper. Students said they like the lessons being about the same topic over a length of time. Often science lessons move through the content very quickly. Last year I taught solubility in about four hours, this year that duration was nearer fifteen.

I did not make any other attempts to collect data from the students. I was minded to write a short questionnaire after the teaching had begun but time did not allow that to happen. Students did seem to enjoy the project. On the face of it, it looked no different to my normal teaching. They commented positively on the increased time they were given to stay in one topic science and the extra time given to go back and re-cover tasks which they thought difficult. They also saw the extension work as a bonus – something different from learning for a test.

The decision not to finish with a final test was the correct one. This decision made students very curious. Much of their experience in schools has been learn-test-result-move on. But as I state in section 6.2 there was much more value in scrutinising the responses to the DQs than producing a summative score or a meaningless single level (Taber, 2009).

Could it be done like this all the time? Diagnostic questions produce a large quantity of data from a large group. The teaching time for both groups overlapped for a short time and I felt under pressure to keep up with the paperwork. But I would argue DQs provide such a rich source of data about student thinking and they are not so unwieldy, that 1 or 2 assessments per week on important concepts would be possible in normal teaching. Even one-off diagnostic questions produce some worthwhile material as Bennett (2003a) says ‘written DQs share many of the advantages or written questionnaires as a research tool they enable a lot of data to be gathered rapidly from a large sample.’ (p.29)

6.4 – Implications of findings learning and further work

Since the commencement of this study, the decision has been taken to remove the official use of levels and level descriptors from in the assessment of school pupils from September 2014. This decision followed the report by the National Curriculum review expert panel (DfE, 2011).

Hence I find myself in the rather distinct position: that of writing the conclusion to a study which sought an alternative approach to National Curriculum levels just as they vanish from schools. To be clear, the irony is not lost on me. Understandably, there is concern among teachers that a system the profession learned to use as a shared vocabulary no longer exists to inform children or parents about attainment.

There will be no single system imposed for ongoing assessment and schools will be left to create their own. I would say that there are two ways to proceed. The first is to carry on with an emphasis on testing and allow the experience of results to dominate the learning experience of children.

Alternatively there are some timely comments which could begin to shift the emphasis towards learning:

1. The expert panel (DfE, 2011) stated that quality assessment needs precisely defined constructs are clear about what pupils should know and understand.
2. The expert panel also recommended curriculum Programmes of Study consist of discursive statements of purpose with anticipated progression. This would place an emphasis on mastery learning and allow learners to move on when ready. The expert panel gave the example of high-performing jurisdictions

3. The new key stage 3(*sic*) curriculum for science (DfE, 2013) described sequences of knowledge and concepts. It emphasised that it was vitally important pupils develop ‘secure understanding of each key block of knowledge and concepts in order to progress to the next stage. Insecure, superficial understanding will not allow genuine progression.’ (p. 2).
4. Brill and Twist (2013) proposed that teachers needed to develop a culture of discourse to reach a much better understanding high quality assessment

When we compare these comments to the TGAT’s four principles on which the National Curriculum was to be based: criterion referenced, formative, progressive moderated, one might wonder if these ideas are entirely new.

Schools will now have to think carefully about how they cater for assessment up to GCSE. The dependency on summative assessment will not improve learning as we have seen in the evidence presented in chapter one. A focus on what students know, understand and can do using formative assessment principles is a sensible way forward.

I find myself surprised at giving some of the last words in this thesis to Ofsted. In their recent advice to inspectors (Ofsted, 2014) they advised their members when considering how accurate assessment data in schools was to ask how assessment could allow teaching to be modified to meet students’ needs. Also it asked how assessment revealed what students know, understand and can do.

With regard progress of pupils inspectors will consider how well:

> *pupils’ strengths and misconceptions are identified and acted on by teacher during lessons, and more widely to:*
>
> – *plan future lesson and teaching.*

– *remedy where pupils do not demonstrate knowledge or understanding of a key element of the curriculum.*
– *deepen knowledge and understanding of the most able.*

(p. 4)

It is difficult to see how these could be achieved through regular summative testing.

I began this thesis with a question about how I might improve my understanding of formative assessment and improve the learning of my students. I feel I have come a long way on this journey even though I have barely scratched the surface of the issues of assessment and learning.

The final words go to Gregory Cizek, a co-editor of the Handbook of Formative Assessment (Cizek, 2010). Musing on the challenges that lie ahead for all teachers trying to provide the best learning experience for their students through developing effective formative assessment, he concludes with a comment which I feel puts the challenge into context:

> *'Development of high quality assessment events, the purposeful integration of formative assessment into classroom activities and the iterative use of formative assessment results to inform the next instructional steps for teachers and refinement of learning goals for students. Each of these requires considerable investment in time to plan and conduct the activity…* (p.12)

Teaching and learning sequence

The sequence below describes the likely progression in pupils' knowledge and understanding in the part of the national curriculum covered by this unit.

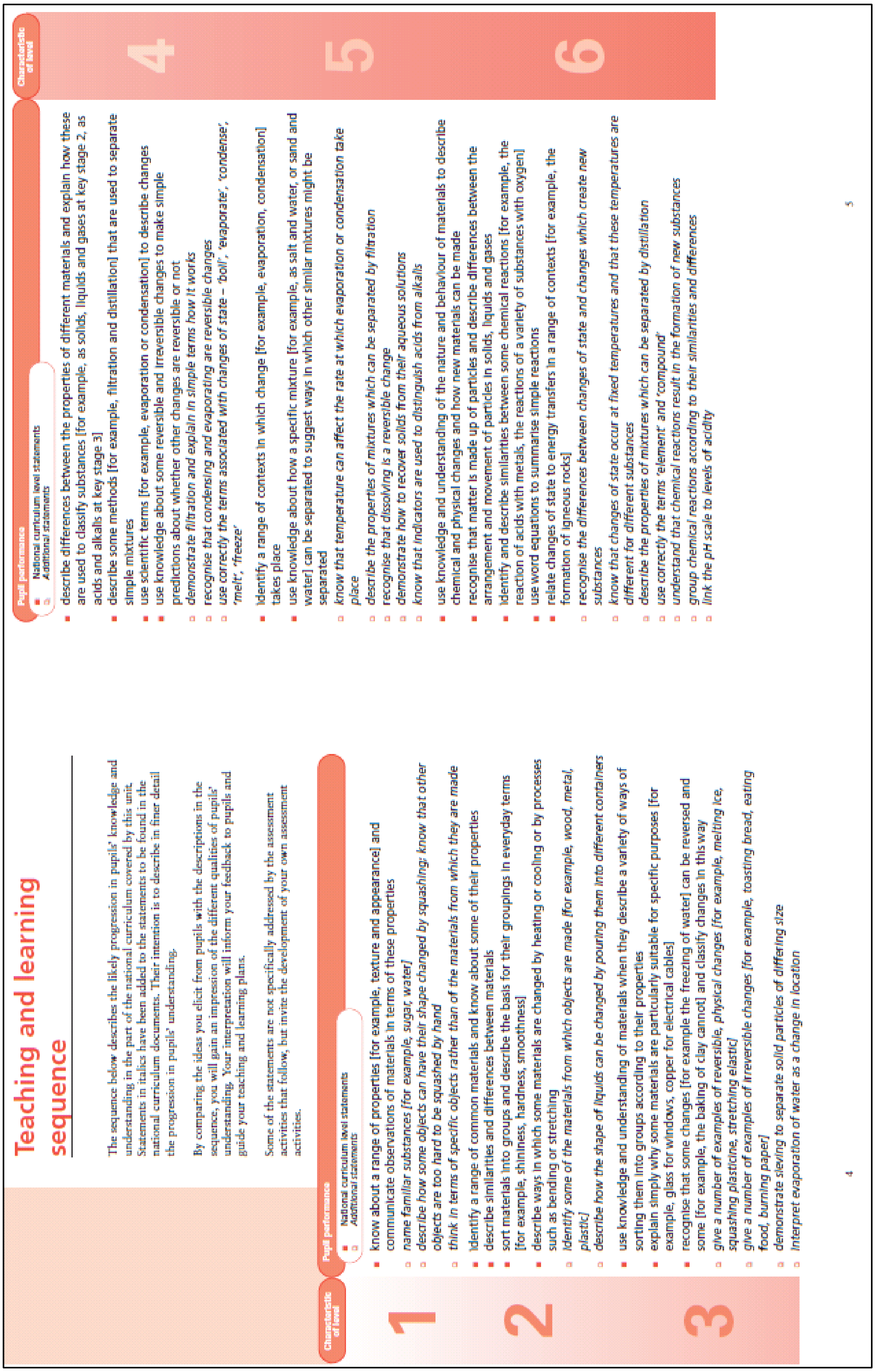
By comparing the ideas you elicit from the pupils with the descriptions in the sequence, you will gain an impression of the different qualities of pupil's understanding. Your interpretation will inform your feedback to pupils and guide your plans for teaching and learning.

Some of the statements are not specifically addressed by the assessment activities that follow, but invite the development of your own assessment units.

Characteristic of Level	Pupil performance (■ National curriculum level statements; □ *Additional statements*)
1	■ know about a range of properties [for example, texture, appearance] and communicate observations of materials in terms of these properties
2	■ identify a range of common materials and know about some of these properties ■ describe similarities and differences between materials ■ sort materials into groups and describe the basis for yheir groupings in everyday terms [for example, shininess, hardness, smoothness] ■ use their knowledge and understanding of material when they describe a variety of ways of sorting them into groups according to their properties
3	□ *define solids as hard, heavy substances, which keep their shape and size* □ *define liquids as wet substances, which can be poured and take on the shape of their containers* □ *use these definitions to sort simple examples [such as wooden ruler, ceramic tile, tap water, cooking oil] into solids and liquids*
4	■ describe differences between the properties of different materials and explain how these differences are used to classify substances [for example, as solids, liquids, gases at key stage 2, and acids and alkalis at key stage 3] □ *define solids as rigid substances which keep their size and shape [unless squashed or stretched] and flow only if in very small pieces* □ *define liquids as substances which flow, take on the shape of their containers but maintain their volume* □ *define gases as very light substances that flow easily and fill any container into which they are put* □ *use these definitions to sort many substances into solids, liquids and gases* □ *refer to solid, liquid and gas as the three 'states of matter'*
5	□ *demonstrate an increasing knowledge of materials and their properties* □ *recognise many everyday substances as mixtures of solids, liquids and/or gases* □ *know that gases can easily be compressed and that solids cannot* □ *recognise that for a substance to be detected by smell some of it must be in the gas state* □ *recognise that substances are made up of particles*
6	■ recognise that matter is made up of particles, and describe differences between the arrangement and movement of particles in solids, liquids and gases □ *know that liquids cannot easily be compressed* □ *use a moving particle theory of matter to explain the differences in density and compressibility of the three states* □ *know that the total mass is conserved in physical and chemical changes*
7	■ use knowledge and understanding drawn from the key stage 3 programme of study to make links between the nature and behaviour of materials and the particles of which they are composed ■ use the particle model of matter in explanations of phenomena [for example, changes of state] ■ explain differences between elements, compounds and mixtures in terms of their constituent particles □ *understand and describe the role of energy in particle behaviour* □ *understand how the movement of particles relates to temperature* □ *use the particle model to explain diffusion and gas pressure* □ *use the knowledge that particles are different in different substances to explain their different properties* □ *use the terms 'atom' and 'molecule' to describe particles.*

4 5

National curriculum level statements and additional statements for Materials (Russell & McGuigan 2003a)



Teaching and learning sequence

The sequence below describes the likely progression in pupils' knowledge and understanding in the part of the national curriculum covered by this unit. Statements in italics have been added to the statements to be found in the national curriculum documents. Their intention is to describe in finer detail the progression in pupils' understanding.

By comparing the ideas you elicit from pupils with the descriptions in the sequence, you will gain an impression of the different qualities of pupils' understanding. Your interpretation will inform your feedback to pupils and guide your teaching and learning plans.

Some of the statements are not specifically addressed by the assessment activities that follow, but invite the development of your own assessment activities.

Characteristic of level	Pupil performance (■ National curriculum level statements; □ *Additional statements*)
1	■ know about a range of properties [for example, texture and appearance] and communicate observations of materials in terms of these properties □ *name familiar substances [for example, sugar, water]* □ *describe how some objects can have their shape changed by squashing; know that other objects are too hard to be squashed by hand* □ *think in terms of specific objects rather than of the materials from which they are made*
2	■ identify a range of common materials and know about some of their properties ■ describe similarities and differences between materials ■ sort materials into groups and describe the basis for their groupings in everyday terms [for example, shininess, hardness, smoothness] ■ describe ways in which some materials are changed by heating or cooling or by processes such as bending or stretching □ *identify some of the materials from which objects are made [for example, wood, metal, plastic]* □ *describe how the shape of liquids can be changed by pouring them into different containers*
3	■ use knowledge and understanding of materials when they describe a variety of ways of sorting them into groups according to their properties ■ explain simply why some materials are particularly suitable for specific purposes [for example, glass for windows, copper for electrical cables] ■ recognise that some changes [for example the freezing of water] can be reversed and some [for example, the baking of clay cannot] and classify changes in this way □ *give a number of examples of reversible, physical changes [for example, melting ice, squashing plasticine, stretching elastic]* □ *give a number of examples of irreversible changes [for example, toasting bread, eating food, burning paper]* □ *demonstrate sieving to separate solid particles of differing size* □ *interpret evaporation of water as a change in location*
4	■ describe differences between the properties of different materials and explain how these are used to classify substances [for example, as solids, liquids and gases at key stage 2, as acids and alkalis at key stage 3] ■ describe some methods [for example, filtration and distillation] that are used to separate simple mixtures ■ use scientific terms [for example, evaporation or condensation] to describe changes ■ use knowledge about some reversible and irreversible changes to make simple predictions about whether other changes are reversible or not □ *demonstrate filtration and explain in simple terms how it works* □ *recognise that condensing and evaporating are reversible changes* □ *use correctly the terms associated with changes of state – 'boil', 'evaporate', 'condense', 'melt', 'freeze'*
5	■ identify a range of contexts in which change [for example, evaporation, condensation] takes place ■ use knowledge about how a specific mixture [for example, as salt and water, or sand and water] can be separated to suggest ways in which other similar mixtures might be separated □ *know that temperature can affect the rate at which evaporation or condensation take place* □ *describe the properties of mixtures which can be separated by filtration* □ *recognise that dissolving is a reversible change* □ *demonstrate how to recover solids from their aqueous solutions* □ *know that indicators are used to distinguish acids from alkalis*
6	■ use knowledge and understanding of the nature and behaviour of materials to describe chemical and physical changes and how new materials can be made ■ recognise that matter is made up of particles and describe differences between the arrangement and movement of particles in solids, liquids and gases ■ identify and describe similarities between some chemical reactions [for example, the reaction of acids with metals, the reactions of a variety of substances with oxygen] ■ use word equations to summarise simple reactions ■ relate changes of state to energy transfers in a range of contexts [for example, the formation of igneous rocks] □ *recognise the differences between changes of state and changes which create new substances* □ *know that changes of state occur at fixed temperatures and that these temperatures are different for different substances* □ *describe the properties of mixtures which can be separated by distillation* □ *use correctly the terms 'element' and 'compound'* □ *understand that chemical reactions result in the formation of new substances* □ *group chemical reactions according to their similarities and differences* □ *link the pH scale to levels of acidity*

4 5

National curriculum statements and additional statements for solids, liquids and gases. (Russell & McGuigan, 2003b)

Solution Chemistry Story

Section 1

Substances dissolve or do not dissolve. Dissolving can happen in water or other solvents. Solutions are clear or transparent: they are not cloudy or have solute which will not mix. A misconception is that undissolved solid at the bottom of a container of solvent is insoluble. Given time it will dissolve if it is soluble.

Section 2

When substances dissolve their particles separate and mix with the solvent. If they do not dissolve then they do not mix with the solvent. The solute particles are so small they cannot be seen but they can be detected by colour or taste. This means that solute cannot be filtered out of solution. Because dissolving is a physical change and not a chemical one solute can be retrieved by evaporation of the solvent.

Section 3

Stirring is not required for dissolving. Dissolving will occur because the solvent particles have kinetic energy and move. They physically interact with the solute particles on its surface. Dissolved solute particles will diffuse within the solvent. Stirring will speed up dissolving because it removes the concentration gradient around the solute.

Section 4

Mass is conserved during dissolving but not volume. This is because solute particles can occupy the ‘spaces’ in between solvent particles. Even though the solute particles are separated in this way they have not melted because they have not been given the energy to change state.

Section 5

Some substances will not dissolve in water but there are other solvents these will dissolve in (organic solvents). We know everyday examples of these like paint thinners and nail varnish remover. These solvents will not dissolve the solutes which dissolve in water. The dissolving process in these substances is similar to that for a solute which is soluble in water.

Some liquids such as alcohol will mix with water and act as organic substances. There is a special reaction which tests for fats which uses this property.

Section 6

The surface area of a solid will affect how quickly it will dissolve. This is because a large surface area allows more interaction between the particles of solute and solvent. The relation ship between total surface area and the time taken to dissolve is best shown on a line graph. This type of graph can be used to obtain other information or make predictions about the relationship in different conditions.

Section 7

A solvent will only be able to dissolve a certain amount of solute. The extent at which no more solute will dissolve in a solvent is called the saturation point. Saturation is linked to solubility which is the mass of solvent in a given volume of solvent. Saturation and solubility depend on solute and temperature but not volume of solvent. Solubility for most solids increases as the solvent becomes warmer. The solubility of a solid at different temperatures is best shown on a line graph. Again, the graph can be used to obtain more information about solubility. Solubility, solute mass and solvent volume are linked and can be used to calculate the third quantity when the other two are known.

Section 8

Gases will dissolve. The solubility of a gas is the inverse of that for solids in that it becomes less soluble at higher solvent temperatures. This has implications for aquatic animals which depend on dissolved oxygen to live and what temperature drinks are best served at. The solubility of a gas is also affect by pressure. When pressure is lowered dissolved gas will nucleate and come out of solution. Nucleation happens more easily at an interface. The cola-mento geyser is a dramatic example of this happening

Section	Learning intention Students will have been taught:	Misconceptions	Johnson & Tymms (2011) description and score of difficulty	National Curriculum Pupils should be taught:
1	solids are soluble or insoluble	use of words such as melt disappear	21 distinguish between dissolving and not dissolving	
	solutions are transparent and possible coloured		31 recognise dissolving for coloured solutions	
	dissolving can occur if some solute is left	solute is located at the bottom and not evenly distributed throughout the mixture		
	dissolved solute can be detected by colour change or taste	solute does not disappear in water	31 recognise dissolving for coloured solutions	
2	dissolving happens because solute particles separate and mix with solvent particles	solute depicted a bits drawn in and distributed without water particles	33-44 Recognise solutions as mixtures	level 5 to recognise that dissolving is a reversible change
		dissolved sugar will not pass through a filter paper during filtration	54 filtering will separate a suspension not a solution	level 5 to describe properties of mixtures which can be separated by filtration
	solution can be retrieved by evaporation		34 salt crystal residue on evaporating salt solution	level 5 to demonstrate how to recover solids from their aqueous solutions
	dissolving and melting are not the same physical process	solute does not melt in water	28 distinguish between melting and dissolving	
3	solute particles mix and diffuse	dissolving requires stirring stirring makes dissolving faster as particles have more energy	31 intrinsic motion in liquid state	level 7 use particle model to explain diffusion

4	mass is conserved but not volume	mass is lost during dissolving water absorbs solute	45 mass does not change with dissolving	level 6 when a physical change takes place, mass is conserved
5	there are solvents other than water			solids can dissolve in liquids other than water there are differences in solubility in different solvents
6	a powdered substance will dissolve faster that lumps		60-65 a substance's particles can explain dissolving	level 6 to choose scales for graphs to show data effectively level 7 to interpret patterns in data using scientific knowledge
7.	solutions become saturated solubility depends on solute and temperature			formation of saturated solutions variations of solubility with temperature level 7 to interpret patterns in data using scientific knowledge
8.	gases can dissolve		62-66 air and water can form a mixture	level 7 to use the particle model to explain gas pressure

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items Recognising dissolving Question numbers [1]
• solids are soluble or insoluble • solutions are transparent not cloudy or with solute not mixed in	• identify soluble substances from a range of candidates • give correct reasons for selecting test substance as soluble or insoluble	• 1 • 2a • 2b
• there is a misconception that dissolving has not taken place if there is any solute left	• give suitable reasons to explain why solute may not fully dissolve in water	• 3
• although the solution is transparent the solution can still be detected by colour change or taste	• select a suitable means of detecting dissolved solute in a given solution	• 4

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items Investigating solutions Question numbers
• Dissolving happens because solvent particles interact with solute particles and cause them to separate and mix	• explain why salt solution should pass through filter paper but chalk suspension should not • correctly predict that reducing the size of holes in the filter paper would not stop the solution passing through • correctly account for differences between microscope slides of salt solution and chalk	• 5 • 6 • 7 • 8 • 9
• solute can be retrieved by evaporation of solute	• recognise that salt is left when water is evaporated • illustrate the process of evaporating filtrate	• 10 • 12
• dissolving and melting are not the same physical process	• give reasons why evaporation of filtrate is similar to and different to freezing of a liquid to a solid	• 11a • 11b

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items Will the crystal dissolve? Question numbers
• Solute particles mix and diffuse throughout the solvent	• Account for the order of a series of slides showing the dissolution of a solid in solvent • Predict correctly and explain why a solute crystal will slowly dissolve over time without stirring	• 13 • 14a • 14b
• concentration of dissolved solute is highest nearest around solid	• base an explanation of a crystal dissolving on a diagram representing concentration gradients around undissolved solute	• 15
• stirring removes dissolved solute from around solid allowing the water to gain more access	• develop the idea above to account for stirring speeding up dissolving by allowing more water to come into contact with undissolved solute	• 16

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items Where does the solute go? Question numbers
• mass is conserved but not volume because solute particles can occupy the spaces between the solvent particles	• correctly predict the changes in mass and volume when salt is added to a known volume of water • explain why the total volume of salt and water is less when salt dissolves	• 17a • 17b • 18
	• select a model to account for the loss of total volume	• 19

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items Is water the only solute? Question numbers
• there are solvents other than water which will dissolve substances which are insoluble in water	• identify that a substance such as paint is soluble in the correct solvent. • show an understanding that salt particles will not interact with organic solvents	• 20a • 20b
• particles of miscible liquids can mix together and this is similar to dissolving	• represent a series of different mixtures liquids to show the outcome of mixing liquids in different combinations	• 21

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items One lump or two? Question numbers
• the surface area of cube has an increased surface area when it is cut into smaller and smaller pieces • the same mass of a powdered substance will dissolve faster than lumps of the substance	• link the increasing surface area of a solid with an increase in dissolution rate • explain faster dissolving through the greater interaction between solute and solvent particles	• 22
• a linegraph is a suitable means of displaying the relationship between two variables	• represent the rate of dissolution of a solid which has been cut into smaller and smaller pieces • use a given linegraph to draw conclusions about dissolving speed under a range of different conditions	• 23 • 24

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items How much will dissolve? Question numbers
• solutions can become saturated this determines solubility • solubility depends on the type of solute and the temperature of the solvent	• organise information on masses of solute dissolving at different temperatures to construct linegraphs • interpret the information on linegraphs to describe different trends in solubility	• 25 • 26 • 27
• solubility can be calculated using the procedure mass/volume which can be rearranged	• calculate solubility, mass or volume in a series of simple problems	• 28

Learning intentions Students will have been taught	Intended learning outcomes Students who know and understand this will be able to	Evidence of learning items Do gases dissolve Question numbers
• gases can dissolve • the solubility of gases is lower at higher temperatures	• Explain why bubbles form in a beaker of water as it is slowly heated • Use data to describe the relationship between gas solubility and temperature	• 29 • 30
• solubility of gases is higher at higher pressures	• describe and explain what happens in the cola-mentos geyser demonstration	• 31

Recognising Dissolving

Equipment

6 test tubes test tube rack plastic rod spatula
10ml measuring cylinder

Method

1. Measure 5ml of water into the measuring cylinder and pour it into a test tube.
2. Add a spatula or a few drops from a pipette of the test material to the water.
3. Use a plastic rod to stir the mixture.
4. Write down your observations in the table below.

Name of substance	Observations	Does it dissolve?
table salt		
copper sulphate		
chalk powder		
glycerine		
green food colouring		
wax		

1. How do you know which substances dissolved? What made you decide?

____

____

____

Recognising Dissolving

William added a small amount of 5 different substances to 5 different test tubes containing water. He stirred them using a plastic rod. He observed the results and wrote them in a table.

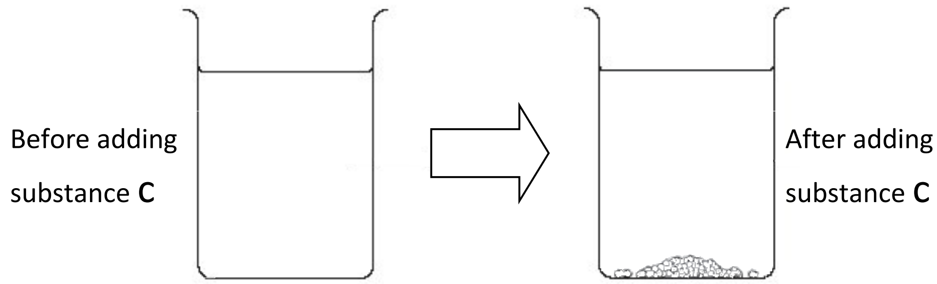
The table is shown below.

Substance	Observation
a	went cloudy
b	went clear blue colour
c	went transparent
d	floated on top in white clumps
e	sank to bottom

2. a) Which of the substances do think dissolved in the water.

2. b) Give a reason why you think this.

Substance c was sugar. William added a larger amount of substance c (sugar) to a beaker of water and stirred it. A picture of the beaker before and after is shown below.



3. Which of the following statements do you think explains what has happened:

   a. None of the sugar has dissolved; it's just sunk to the bottom of the beaker.
   b. Some of the sugar has dissolved because the water is still clear.
   c. The sugar did not dissolve because the pupil did not stir it enough.
   d. William added too much for it all to dissolve.

Write below the letter of the statements you agree with and give your reasons for your choice.

4. William wanted to check if any of the sugar had dissolved. Which of the following could he do to see if there was any sugar dissolved in the water? *Circle any answers that you think are correct.*

taste the water   filter the water   evaporate the water

look at the water under a microscope   look at the colour of the water

Investigating solutions

Equipment

2 test tubes 2 filter papers 2 funnels 2 white tiles

Method

1. Fold the filter paper to make a cone.
2. Wet the filter paper slightly and place it into the funnel. Place the funnel into one of the clean test tube.
3. Pour the mixture of table salt and water into the filter paper.
4. When mixture has gone through the filter paper remove it open it and place it onto a white tile.
5. Now repeat this process for the chalk mixture. (Make sure you shake up the mixture first).

Results

Draw what you can see on the diagram below.

Table salt	Chalk
Appearance of filter paper	Appearance of filter paper

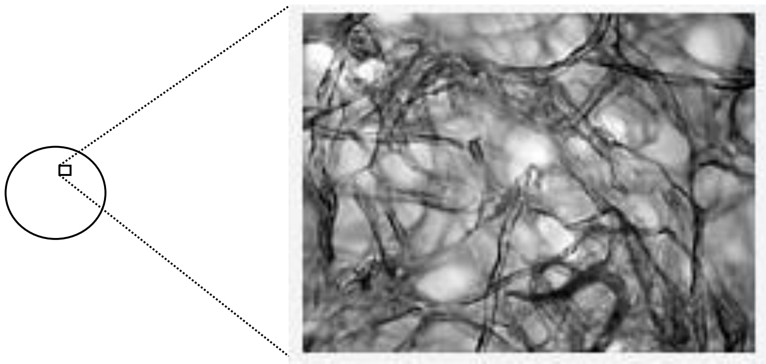
5. What explanation can you give for these results?

Investigating Solutions

filter paper
magnified x 200

The picture shows

the filter close up.

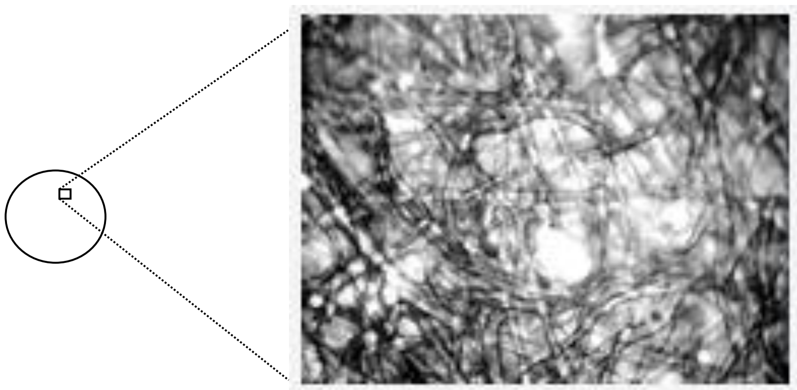


The photograph above shows a magnified image of the filter paper that was used to filter the chalk and water mixture and the salt solution.

6. Describe what happened to the chalk and the salt when they were passed through this filter paper.

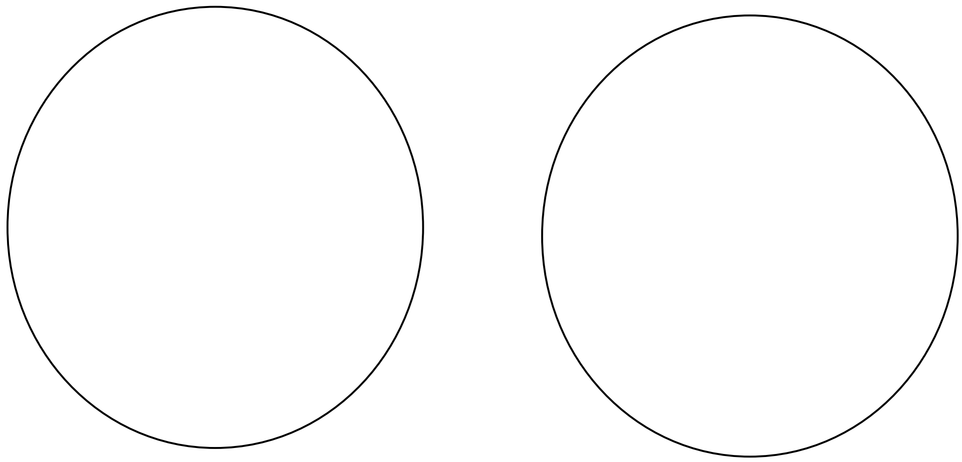
The photograph below shows an image of a filter paper with much smaller holes.

filter paper
magnified x 200



7. If this filter paper was used, do think the results of filtering the chalk or salt would be different? Explain why you think this.

Investigating Solutions
Making slides



salt chalk

Show on the circles above what you saw when you viewed them using the microscope.

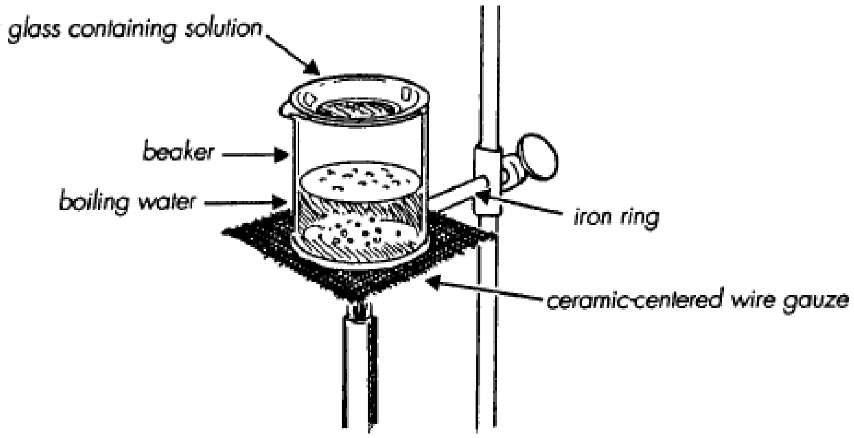
8. Which of the following best describes what you saw on the chalk slide:

a) small pieces of undissolved chalk
b) individual chalk particles
c) individual chalk molecules
d) chalk and water particles reacting

9. How can you explain the observations you made on the salt slide:

a) the salt particles have disappeared into the water
b) the salt particles are hidden by the water
c) the salt particles are too small to see
d) the salt cannot be seen as it is a liquid

Evaporating the water

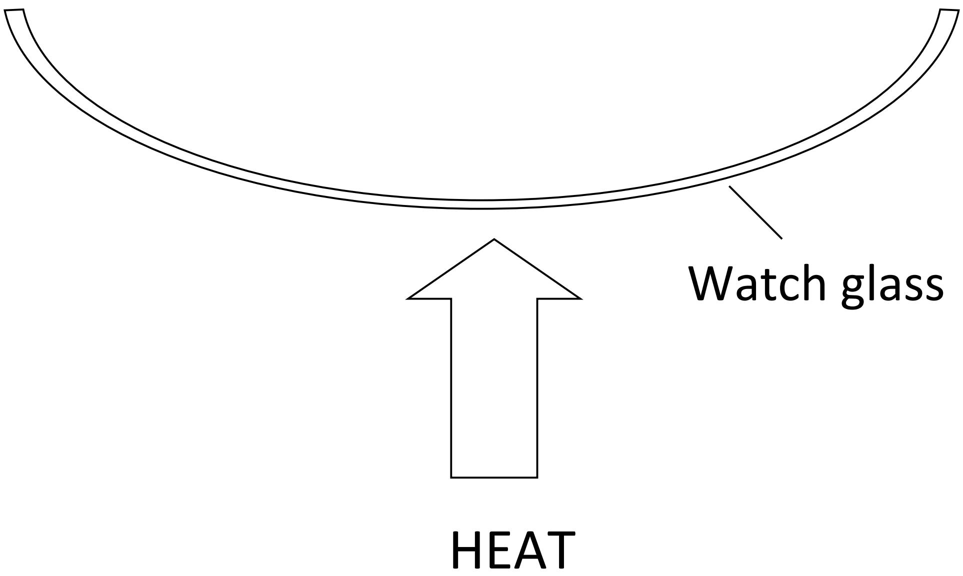


10. What did you see on the watch glass when the water was evaporated?

11a. In what way is this process *similar* to freezing?

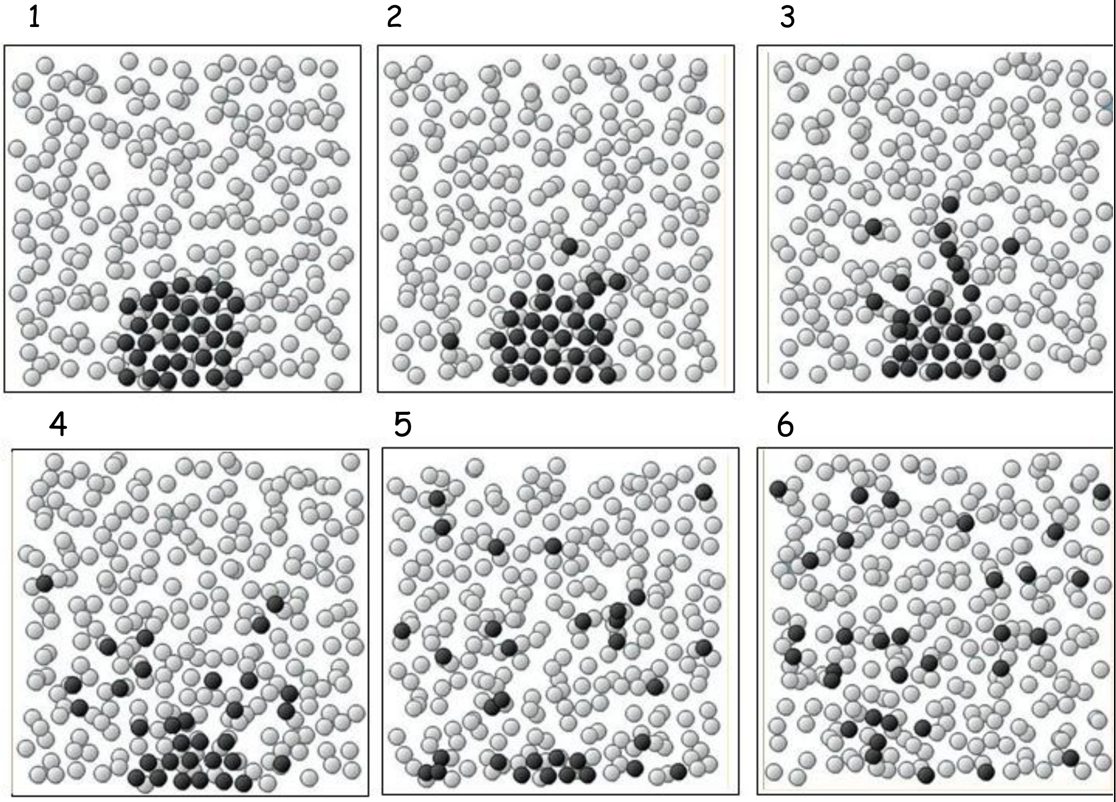
11b In what way is this process *different* to freezing?

12. In the space below draw the particles of water and salt as they are heated in the watch glass.



Explaining Dissolving

Look at the following pictures. They show a sequence of a solid dissolving in water.



The white dots represent water and the black dots represent salt.

13a. In what ways is this sequence of pictures a good model for explaining dissolving?

____

____

____

13b. In what ways are these pictures not a good model for explaining dissolving?

____

____

____

Will the crystal dissolve?

The diagram shows two identical beakers of water with a crystal of copper sulphate at the bottom. In one beaker the water has not been stirred, the other beaker will be placed on a mechanical stirrer. In the space beside the diagram, draw what you think the beakers of water will look like in a week's time.

14a

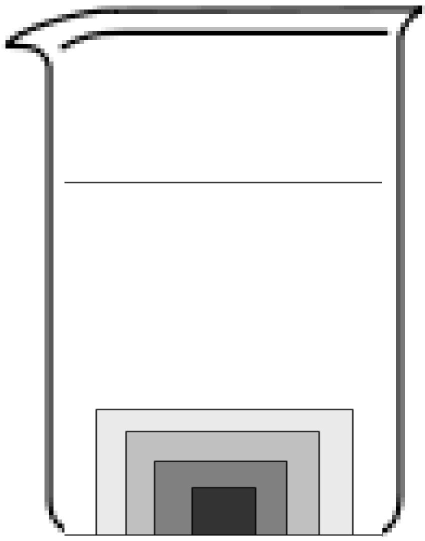
 *from tutorvista.com*	
Beaker of water with crystal now.	Beaker after one week **without** stirring

 *from tutorvista.com*	
Beaker of water with crystal now.	Beaker after one week with stirring

14b. Will stirring change what happens to the crystal in the water? Why do you think this?

Did the crystal dissolve?

15. Use the diagram to explain how the crystal dissolved without being stirred.



16. Why does the crystal in the stirred water dissolve more quickly?

Where does the solute go?

You will see a demonstration of salt being dissolved in water in a volumetric flask. Salt is added to the flask and left to settle. All air is removed from the flask. The volume of water is adjusted to 500ml. Then the flask is vigorously shaken.

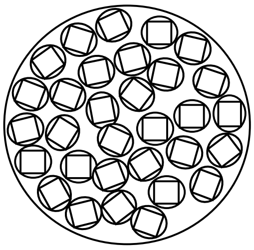
17a. How will the mass of the salt and water change during the experiment?

17b. How will the total volume of the salt and water change during the experiment?

18. Do you have any explanation for the change in total volume in the flask after the salt and water are shaken together?

Where does the solute go?

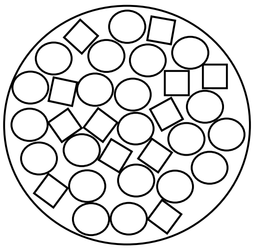
19. Choose one of the following models and say why it is the best one to explain why volume but not mass changed in the demonstrations? Which models are not useful?



Key salt particle ☐ water particle ◯

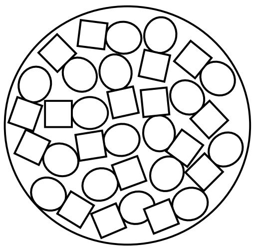
Model 1

The salt particles are taken up into the water particles.



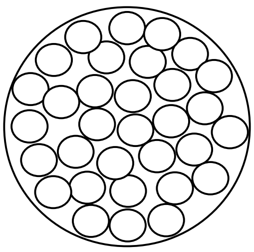
Model 2

The salt particles fit into spaces between the water particles.



Model 3

The salt and water particles combine in some way that makes the salt particles take up less space.



Model 4

The salt just melts away and leaves the water particles on their own.

Is water the only solute?

Equipment

4 small test tubes pipette

samples of dried solute

Method

1. Measure a several millilitres of solvent into a small test tube.
2. Hold the lolly-stick with the dried solute into the solvent.
3. Observe what happens. Put a tick if it dissolves or a cross if it does not.
4. Repeat for all solvent and solutes.

Results

	Does it dissolve in solvent			
solid	water	white spirit	acetone	ethanol
glass paint (white)				
nail varnish (red)				
model paint (black)				
board marker (blue)				
Hammerite© (green)				

What are your conclusions from this experiment?

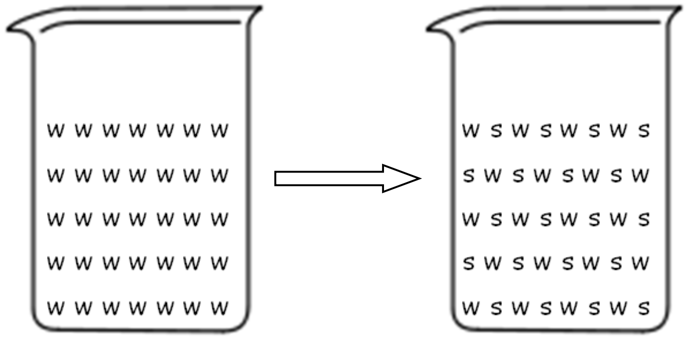
20. Read the following statements. Decide whether they are **true** or **false** and then select the reason why you think this.

a) Paint is insoluble. This statement is **true/false**
because:
   i. paint will not dissolve in water
   ii. paint will mix with water but only slightly
   iii. paint dissolves in white spirit
   iv. too much paint was used the experiment

b) Ethanol is a suitable solvent for nail varnish. This statement is **true/false**
because:
   i. nail varnish is too heavy to dissolve in ethanol
   ii. nail varnish particles will not mix with ethanol particles
   iii. nail varnish has a much higher boiling point than ethanol
   iv. ethanol evaporates to leave the nail varnish behind

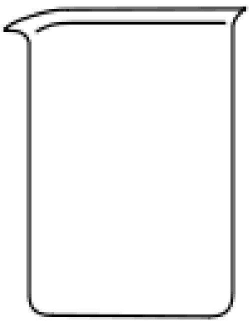
Is water the only solvent?

Amy added salt to water. She represented her beaker of water particles using the letter W. She represented her beaker of salt solution using the letters S and W.



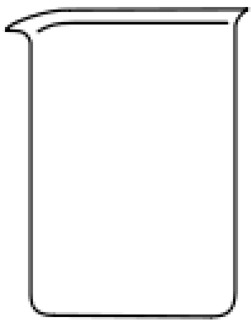
21. Use this model to represent what you would see saw in the mixture below:

a)



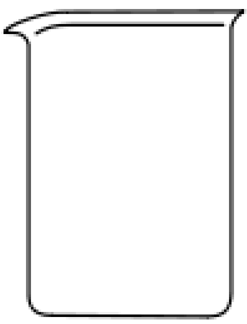
water (W) and olive oil (O)

b)



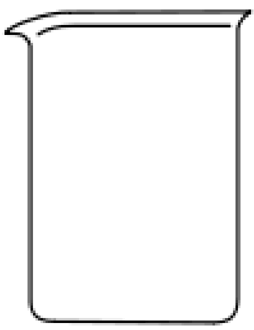
water (W) and ethanol (E)

c)



olive oil (O) and ethanol(E)

d)



water (W), olive oil (O) and ethanol (E)

One lump or two?

We know that sugar lumps take longer to dissolve than granulated sugar. But how could we investigate the relationship between the grain size of a solute and how fast it dissolves?

$8cm^3$ cube of jelly cut into pieces	time taken for jelly to dissolve at 50°C (s)	mass of jelly cube (g)	surface area of jelly cube (cm^2)
1	635	10	24
2	345	10	32
4	254	10	40
8	111	10	48

How does the speed of dissolving change as the cube is cut into more pieces?

22. Why does it take less time to dissolve the same mass of solid if it is in smaller pieces? Use a diagram if it would help you to explain.

23. How can we present this data using a graph?

24. Use graph to obtain information about the rate of dissolving?

a. How fast would the jelly dissolve if it had a surface are of $35cm^2$?

b. What surface area would take 5 minutes to dissolve?

c. What increase in surface area is needed to halve the dissolving time?

volume of water 100ml	volume of water 100ml	volume of water 100ml	volume of water 100ml
temperature of water 10°C	temperature of water 20°C	temperature of water 30°C	temperature of water 40°C
			
name of salt potassium chloride	name of salt potassium chloride	name of salt potassium chloride	name of salt potassium chloride
mass which dissolves 30.4g	mass which dissolves 33.5g	mass which dissolves 35.6g	mass which dissolves 39.7g

volume of water 100ml	volume of water 100ml	volume of water 100ml	volume of water 100ml
temperature of water 50°C	temperature of water 60°C	temperature of water 70°C	temperature of water 80°C
			
name of salt potassium chloride	name of salt potassium chloride	name of salt potassium chloride	name of salt potassium chloride
mass which dissolves 42.3g	mass which dissolves 44.5g	mass which dissolves 48.6g	mass which dissolves 51.9g

volume of water 100ml	volume of water 100ml	volume of water 100ml	volume of water 100ml
temperature of water 10°C	temperature of water 20°C	temperature of water 30°C	temperature of water 40°C
			
name of salt sodium chloride	name of salt sodium chloride	name of salt sodium chloride	name of salt sodium chloride
mass which dissolves 35.5g	mass which dissolves 35.1g	mass which dissolves 35.5g	mass which dissolves 36.7g

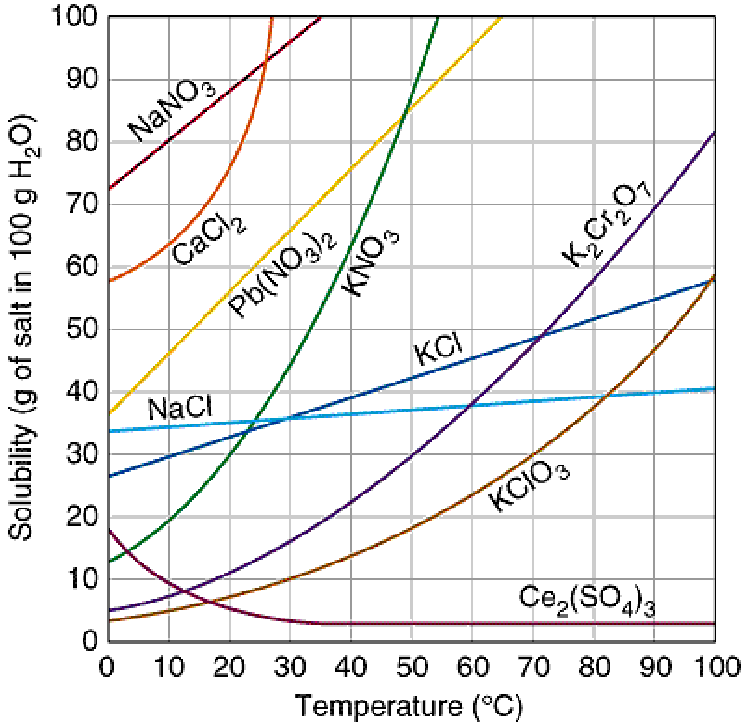
volume of water 100ml	volume of water 100ml	volume of water 100ml	volume of water 100ml
temperature of water 50°C	temperature of water 60°C	temperature of water 70°C	temperature of water 80°C
			
name of salt sodium chloride	name of salt sodium chloride	name of salt sodium chloride	name of salt sodium chloride
mass which dissolves 37.4g	mass which dissolves 38.1g	mass which dissolves 39.2g	mass which dissolves 39.9g

How much will dissolve?

25. Use the cards to show how the solubility of potassium chloride and sodium chloride changes for a range of temperatures using a suitable graph.

26. Describe the similarities and differences between the solubility of potassium and sodium chloride.

27. Use the graphs below to answer the following questions



a) what is the solubility of $CaCl_2$ at 20°C?

b) which substance has a solubility of 30g/100ml at 70°C?

c) at what temperature is the solubility of $K_2Cr_2O_7$ and NaCl equal?

d) Which substance has the lowest solubility at 10°C?

e) Which substance has the highest solubility at 10°C?

28. Use the following procedure to complete the following calculations:

$$\text{solubility (g/100ml)} = \frac{\text{mass (g)}}{\text{volume of solvent (ml)}} \times 100$$

Calculate the solubility of following solutes:

a) 37g of solute A will dissolve in 100ml of water at 25°C.

b) 48g of solute B will dissolve in 200ml of water at 25°C.

c) 15g of solute C will dissolve in 50ml of water at 25°C.

d) 19.5 g of solute D will dissolve in 30ml of water at 25°C.

e) 77g of solute E will dissolve in a litre of water at 25°C

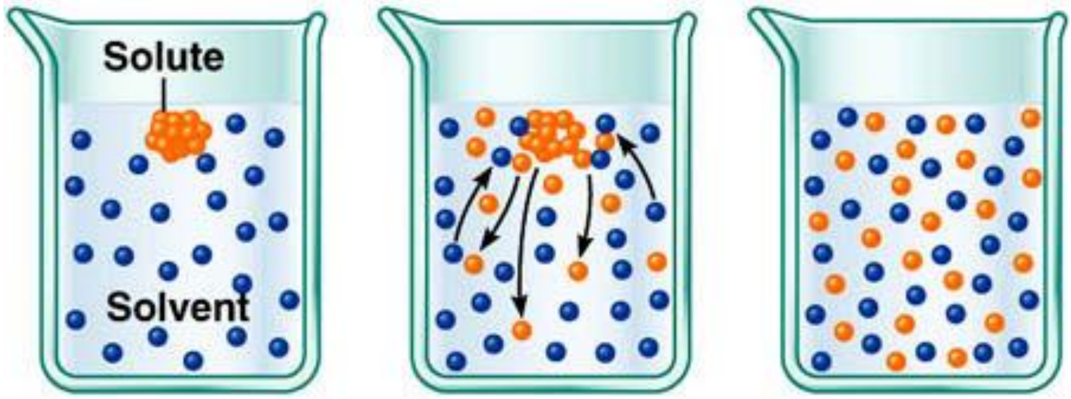
Calculate the mass of solute which will dissolve:

a) Solute A in 75ml of water at 25°C.

b) Solute B in 550ml of water at 25°C.

c) Solute C in 30ml of water at 25°C.

Why do solutions become saturated?



28s Use the idea of forces of attraction between particles to explain the following:

Solids dissolve in liquids.

Solutes will dissolve faster in hotter solvents

Adding more solute eventually produces a saturated solution.

Do gases dissolve?

Equipment

250 ml beaker Bunsen burner tripod gauze
heat proof mat thermometer clamp stand

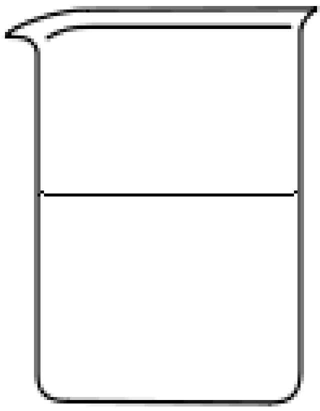
Method

1. Clamp a thermometer into a beaker containing about 150ml of water.
2. Gently heat the water and watch carefully noting any observations

Results

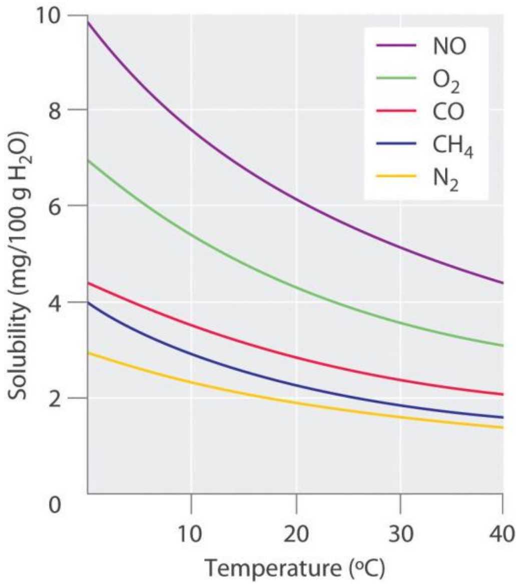
Temperature (°C)	Observations
20	
40	
50	
60	
70	

In the space below draw what you saw in the water.



29. Explain why bubbles of gas formed when the water was gently heated.

Do gases dissolve?

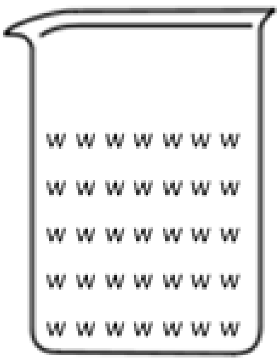


30. Use the graph to answer the following questions:
    a) How does the solubility of the gases change as temperature increases?

    b) Why would a fish find it difficult to survive in water at 30°C

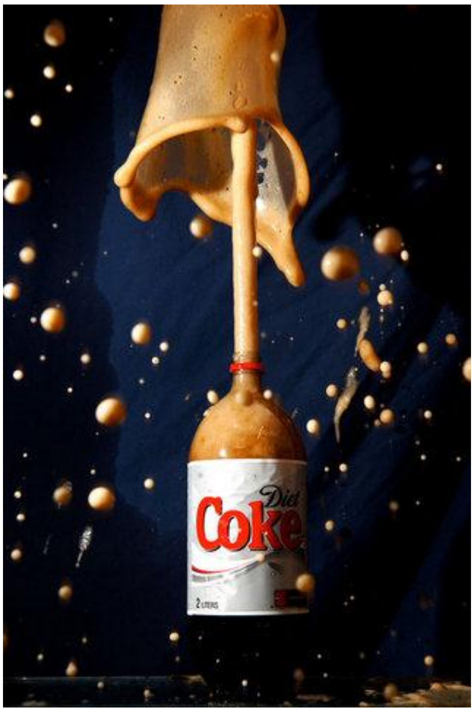
    c) Why are fizzy drinks better when they are served from the fridge?

    d) In the beaker below how could you represent NO dissolved in water?

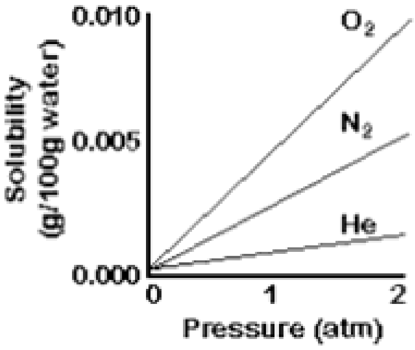


Do gases dissolve?

Coke-Mentos Geyser



31. Use the information in the graph to help you explain how the Coke-Mento geyser works?



References

Adey, P., & Shayer, M. (1993). An exploration of long-term far-transfer effects following an extended intervention program in the high school science curriculum. *Cognition and Instruction, 11*(1), 1-29.

Alonzo, A.C., Steedle, J.T., (2009). Developing and assessing a force and motion learning progression. *Science Education,93*(3), 389-421.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

ARG. (1999). *Assessment for learning: beyond the black box.* Cambridge: University of Cambridge Schools of Education.

ARG. (2002). *Assessment for learning: 10 principles.* University of Cambridge: Assessment Reform Group

Azar, A. (2005). Analysis of Turkish high-school physics-examination questions and university entrance exams according to Bloom's taxonomy. *Journal of Turkish Science Education, 2*(2), 144-150.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., &Morgan, M.T. (1991).The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238.

Bennett, J. (2003a). *Teaching and learning science.* London: Continuum.

Bennett, J. (2003b). *Evaluation methods in research*. London: Continuum.

Bennett, R.E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5-25.

Black, P. (1997). Whatever happened to TGAT? In C. Cullingford (Ed.), *Assessment versus evaluation,* (pp. 24-50).London: Cassell.

Black, P. (2003.) The nature and value of formative assessment for learning. *Improving Schools, 6*(3), 7-22.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning putting it into practice.* Maidenhead: Open University Press.

Black, P., & Wiliam, D. (1998a). *Inside the black box: raising standards through classroom assessment.* London: King's College.

Black, P., & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Bliss, J. (2002). Learning science: Piaget and after. In S. Amos, & R. Boohan (Eds.), *Teaching science in secondary schools: A reader* (pp. 154-163). London: RoutledgeFalmer.

Blosser, P. (2000). *How to ask the right questions.* Arlington, VA : National Science Teachers Association.

Briggs, D.C., Alonzo, A.C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment 11*(1), 33-63.

Brill, F. and Twist, L. (2013). Where Have All the Levels Gone? The Importance of a Shared Understanding of Assessment at a Time of Major Policy Change. *NFER Thinks: What the Evidence Tells Us.* Slough: NFER.

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comment, grades of intrinsic motivation and performance. *Journal of Educational Psychology, 78*(3), 210-216.

Calyk, M., Ayas, A., & Ebenezer, J. V. (2005). A review of solution chemistry studies: insights into students conceptions. *Journal of Science Education and Technology 14*(1), 29-50.

Ciofalo, J. F., & Wylie, E. C. (2006). Using diagnostic classroom assessment: one question at a time. *Teachers College Record (ID Number 12285).* Retrieved September 26, 2014, from http://www.tcrecord.org/Home.asp

Cizek, G. J. (2010). An introduction to formative assessment: history, characteristics and challenges. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 3-17). Abingdon: Routledge.

Coe, R. (2013) *Improving Education: A triumph of hope over experience*. University of Durham: Centre for Evaluation & Monitoring. Retrieved July 14, 2014, from http://www.cem.org/attachments/publications/ImprovingEducation2013.pdf

Cohen, L., Manion, L., & Morrison, K. (Eds.). (2011). *Research methods in Education.* (7th ed.) London: Routledge.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of research in science teaching, 48*(10), 1109-1136.

Coffey, T. S., (2008). Diet Coke and Mentos: What is really behind this physical reaction? *American Journal of Physics76*(6), 551-

Cowie, B., & Bell, B. (1999). A model of formative assessment in science. *Assessment in Education: Principles, Policy & Practice*, *6*(1), 101-116.

Cronbach, L. (1987). Issues in planning evaluation. In R. Murphy, & H. Torrance (Eds.), *Evaluating education: Issues and methods* (pp. 4-35). London: Harper and Row Ltd.

Daugherty, R. (1995). *National curriculum assessment. A review of policy 1987-1994.* London: The Falmer Press.

DCSF. (2007). *Going for gold: securing attainment.* Nottingham DCSF Publications.

DCSF. (2008). *The assessment for learning strategy.* Nottingham: DCSF Publications.

De Jong, O., & Taber, K. S. (2007). Teaching and learning: the many faces of chemistry. In S. K. Abell, & N. G. Ledermann (Eds.). *Handbook of research on Science Education.* Mahwah, NJ: Lawerence Erlbaum Associates, Inc.

Devetak, I., Vogrinc, J., & Glazar, S. A. (2009). Assessing 16-year-olds students' understanding of aqueous solution at submicroscopic level. *Research in Science Education, 39*(2), 157-179.

DfE. (2011) *The framework for the National Curriculum. A report by the Expert Panel for the National Curriculum review.* London: Department for Education.

DfE. (2013).Science programmes of study: key stage 3. Retrieved December 5, 2014, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335174/SECONDARY_national_curriculum_-_Science_220714.pdf.

DfEE/QCA. (1999). *The national curriculum for England: Science.* London: Department for Education and Employment/Qualifications and Curriculum Authority.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: research into children's ideas.* London: Routledge.

Duncan, R. G., Rogat, A. D., & Yarden, A. (2009) A learning progression for deepening students' understandings of the modern genetics across 5^{th}-10^{th} grades. *Journal of research in Science Teaching, 46*(6), 655-674.

Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: a review and analysis. *Studies in Science Education, 47*(2), 123-182.

Ebenezer, J. V., & Erickson, G. L. (1996). Chemistry students' conceptions of solubility: a phenomenography. *Science Education, 80*(2), 181-201.

Ebenezer, J. V., & Fraser, D. M. (2001).First year chemical engineering students' conceptions of energy in solution processes: phenomenographic categories for common knowledge construction. *Science Education, 85*(5), 509-535.

Elliot, J. (1991). *Action research for educational change.* Milton Keynes: Open University Press.

Erickson, G. (2000). Research programmes and the student science learning literature. In. R. Millar, J. Leach, & Osborne, J. (Eds.), *Improving science education* (pp. 271-292). Buckingham: Open University Press.

Fairbrother, R. (2008). The validity of the key stage 3 science tests. *School Science Review, 89*(329), 107-113.

Gall, M. (1984). Synthesis of research on teachers' questioning. *Educational Leadership. 42*(3), 40-47.

Hakim, C. (2000). *Research Design* London: Routledge.

Harlen, W., & Deakin Crick, R. (2003) Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice, 10*(2), 169-207.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365-379.

Harrison, C. (2011), Assessment for learning. In M. Hollins (Ed.), *ASE guide to secondary science education* (pp. 154-160). Hatfield: The Association for Science Education.

Hattie, J. (2009). *Visible learning: a synthesis of meta-analyses relating to achievement.* Abingdon: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Heritage, M. (2008). *Learning progressions: supporting instruction and formative assessment.* Los Angeles: National Centre for Research on Evaluation, University of California. Retrieved October 15, 2014, from https://www.cse.ucla.edu/products/misc/cse_heritage_learning.pdf

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: are we losing and opportunity.* Los Angeles: National Centre for Research on Evaluation, University of California,. Retrieved October 15, 2014, from http://www.ccsso.org/Documents/2010/Formative_Assessment_Next_Generation_2010.pdf

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009) From evidence to action: a seamless process in formative assessment? *Educational measurement: Issues and Practice, 28*(3), 24-31.

Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L. E., & Coe, R. (2013). *The Sutton Trust-Education Endowment Foundation teaching and learning toolkit.* London: Education Endowment Foundation.

Holding, B. (1987) *Investigation of schoolchildren's understanding of the process of dissolving with special reference to the conservation of matter and the development of atomistic ideas.* Unpublished doctoral dissertation, University of Leeds.

Hunt, A. (2011). Five decades of innovation and change. In M. Hollins (Ed.), *ASE guide to secondary science education* (pp. 12-22). Hatfield: The Association for Science Education.

Institute of Physics (2004). Volume change on dissolving salt in water. Retrieved September 1, 2014 from http://www.nuffieldfoundation.org/practical-physics/volume-change-dissolving-salt-water.

Johnson, P. (2011) *Stuff and substance: ten key practicals in chemistry.* London: The Gatsby Science Enhancement Programme.

Johnson, P., & Tymms, P. (2011). The emergence of a learning progression in middle school chemistry. *Journal of research in science teaching, 48*(8), 849-877.

Kapabinar, F., Leach, J., & Scott, P. (2004). The design and evaluation of a teaching-learning sequence addressing the solubility concept with Turkish secondary school students. *International Journal of Science, 26*(5), 635-652.

Keeley, P. D., & Eberle, F. Q. (2008) Using standards and cognitive research to inform the design and use of formative assessment probes. In J. Coffey, R. Douglas & C. Stearns (Eds.), *Assessing science learning: perspectives from research and practice.* Arlington, VA : National Association Teachers Association.

Kemmis, K. (2009). Action research as a practice-based practice. *Educational Action Research, 17*(3), 463-474.

Kingston, N. & Nash, B. (2011). Formative assessment: a meta-analysis and a call for research. *Educational measurement: Issues and Practice, 30*(4), 28-37.

Kluger, A., N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theory into Practice, 41*(4), 212-218.

Leahy, S., Lyon, L., Thompson, M., & Wiliam, D. (2005). Classroom assessment: minute by minute, day by day. *Educational Leadership, 63*(3), 19-24.

Leahy, S., & Wiliam. (2011, April). *Devising learning progressions.* Unpublished paper presented at the annual meeting of the American Educational Research Association, New Orleans,LA.

Lee, C., & Wiliam, D. (2005). Studying changes in practice of two teachers developing assessment for learning. *Teacher Development, 9*(2), 265-283.

Lenton,G., Stevens, B., & Illes, R. (2000) Numeracy in science: pupils' understanding of graphs. *School Science Review, 82*(299), 15-23.

Mercer, N. & Dawes, L. (2008). The value of exploratory talk. In N. Mercer & S Hodgkinson (Eds.), *Exploring talk in school* (pp. 55-71). London: Sage.

Marsh, C.J. (2007). A critical analysis of the use of formative assessment in schools. *Educational Research for Policy and Practice, 8*(1), 25-29.

McNiff, J., & Whitehead, J. (2011). *All you need to know about action research* (2nd ed.). London: Sage.

McSharry, G. & Jones, S. (2000). Role-play in science teaching and learning. *School Science Review 82*(298), 73-82)

Millar, R. & Whitehouse, M. (2012) The York Science Project embedding assessment for learning. *Education in Science Extra (EiSXTRA),* 11, 1-3. Retrieved December 15,2014, from http://www.ase.org.uk/documents/xtra1-the-york-science-project/

Mohan, L., Chen, J., & Plummer, C. W. (2009).Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46*(6), 675-698.

Naah, B. M., & Sanger, J. M. (2013). Investigating students' understanding of the dissolving process. *Journal of Science Education and Technology, 22*(2), 103-112.

National Stem Centre. (2014) Assessment for Learning in Science – training materials – slides. Retrieved December 3, 2014, from http://www.nationalstemcentre.org.uk/elibrary/resource/5304/assessment-for-learning-in-science-training-materials.

Nakleh, M. B. (1992). Why some students don't learn chemistry. *Journal of Chemical Education, 69*(3), 191-196.

Newton, P. E. (2009). The reliability of results from national curriculum testing in England. *Educational Research, 51*(2), 181-212.

Newberry, M., Gilbert, J., & Hardcastle, D. (2005). Visualising progression through the science curriculum in order to raise standards. *School Science Review, 86*(316), 87-96.

Nicol, D. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good practice. *Studies in Higher Education, 31*(2), 199-218.

Ofsted (2008). *Assessment for learning strategy: the impact of National Strategy support.* London: Ofsted

Ofsted. (2013). *Maintaining curiosity: A survey into science education in schools.* Retrieved October 5, 2014, from www.ofsted.gov.uk/resources/130135.

Ofsted. (2014). Note for inspectors: use of assessment information during inspections in 2014/15. Retrieved December 5, 2014, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/379632/Note_20for_20inspectors_20-_20use_20of_20assessment_20information_20during_20inspections_20in_202014-15.pdf

Oversby, J. (2000), Good explanations for dissolving. *Primary Science Review, 63*, 16-19.

Oversby, J. (2002). Assessing conceptual understanding. In S. Amos, & R. Boohan (Eds.), *Aspects of teaching secondary science: Perspectives on practice* (pp. 148160). London: RoutledgeFalmer.

Oversby, J. (Ed.) (2011). *ASE guide to research in science education.*Hatfield: The Association for Science Education.

Parlett, M., & Hamilton, D. (1976). Evaluation as illumination. In D. Tawney (Ed.), *Curriculum evaluation today: trends and implications* (pp. 84-101). London: Macmillan Education Ltd.

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

Prieto, T., Blanco, A., & Rodriguez, A. (1989). The ideas of 11 to 14-year-old students about the nature of solutions. *International Journal of Science Education, 11*(4), 451-463.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science, 28*(1), 4-13.

Reay, D., & Wiliam, D. (1999). 'I'll be a nothing':structure, agency and the construction of identity through assessment [1]. *British Educational Research Journal, 25*(3), 343-354.

Rowe, M. B. (1974). Wait time: slowing down may be a way of speeding up! *Journal of Teacher Education, 37,* 43-50.

Ruiz-Primo, M. A., & Li, M. (2013). Analyzing teachers' feedback practices in response to students' work in science classrooms. *Applied Measurement in Education, 26*(3), 163-175.

Russell, T., & McGuigan, L. (2003a). *Assessing progress in science. Unit 3. Materials.* London: Qualifications and Curriculum Authority.

Russell, T., & McGuigan, L. (2003b). *Assessing progress in science. Unit 7.Solids, liquids and gases.* London: Qualifications and Curriculum Authority.

Sadler, D. R., (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144.

Schneider, M. C., & Gowan, P. (2013). Investigating teachers' skills in interpreting evidence of student learning. *Applied Measurement in Education, 26*(3), 191-204.

Schwarz, C. V., Reister, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Schwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modelling: making scientific modelling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-654.

Science Enhancement Programme. (2005). *Stuff and substance* [CD-Rom]. Retrieved September 12, 2013 from http://11567.stem.org.uk/index.htm.

Scriven, M. (1991). Prose and cons about goal-free evaluation. *American Journal of Evaluation, 12*(1), 55-62.

Shavelson, R. J., & Kurpius, A. (2012). Reflections on learning progressions. In A. C. Alonzo, & A. M. Gotwals (Eds.). *Learning Progressions in Science,* (pp. 13-26). Rotterdam: Sense Publishers.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M, K., & Yin, Y. (2008). On the impact of curriculum embedded formative assessment on learning: a collaboration between curriculum and assessment developers. *Applied Measurement in Education, 21*(4), 295-314.

Shayer, M., & Adey, P. (1981). *Towards a science of science teaching: cognitive demand and curriculum demand.* London: Heinemann Educational Books.

Shepard, L. A. (2008) Formative assessment: caveat emptor. In C. A. Dwyer (Ed.), *The future of assessment* (pp. 279-303). Abingdon: Routledge.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Songer, N. B., Kelcey, B., & Gotwals, A. W. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching, 46*(6), 610-631.

Stewart, W. (2012 July 13). Think you've implemented Assessment for Learning? *Times Educational Supplement.* Retrieved July 14, 2014, from http://www.tes.co.uk/article.aspx?storycode=626184

Sugrue, B. (2002). Problems with Bloom's taxonomy. *Performance Express.* Retrieved September 8, 2014 from http://performancexpress.org/0212/

Sunflower Learning Ltd, (2006). Sunflower for Science (Version 2.01f) [Computer software]. Thames Ditton: Sunflower Learning Ltd.

Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice, 18*(4), 433-449.

Taber, K. (2007). *Classroom-based research and evidence-based practice.* London: Sage

Taber, K. (2009). Testing levels of understanding. *Physics Education, 44*(8), 8-10.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice, 14*(3), 281-294.

Treagust, T., (2006). *Diagnostic assessment in science as a means to improving teaching, learning and retention.* Unpublished paper presented at the UniServe Science Assessment Symposium, Curtin University of Technology, Australia.

Tymms, P. (2004). Are standards rising in English primary schools? *British Educational Research Journal, 30*(4), 477-494.

Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice, 16*(2), 165-184.

Whitehouse, M. (2014). Using a backward design approach to embed assessment in teaching. *School Science Review, 95*(352), 99-104.

Wiggin, G. P., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Wiliam, D. (2001). *Level best? Levels of attainment in national curriculum assessment.* London: Association of Teachers and Lecturers.

Wiliam, D. (2011). *Embedded formative assessment.* Bloomington, IN: Solution Tree Press

Wiliam, D., & Black, P. (1996). Meanings and consequences: a basis for distinguishing formative and summative functions of assessment? *British Education Research Journal, 22*(5), 537-548.

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

Wilson, M., & Sloane, K. (2000). From principles to practice: an embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208.

Wylie, C., & Wiliam, D. (2006, April). *Diagnostic questions: Is there value in just one?* Paper presented at the annual meeting of the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME). San Francisco, CA.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Tomita, M. K., & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education, 21*(4) 335-359.

Yorke, M. (2003). Formative assessment in higher education: moves towards theory and the enhancement of educational practice. *Higher Education, 45*(4), 477-501.