FINAL REPORT FOR OFQUAL

ALTERNATIVES TO TEST ACCOMMODATION: NEW MODELS FOR PUPILS WITH SPEECH AND LANGUAGE DIFFICULTIES

THE SUPPORT MODEL

Ayesha Ahmed and Alastair Pollitt

CamExam

June 2010 Ofqual/10/4782

Ofqual This report has been commissioned by the Office of Qualifications and Examinations Regulation.

Executive Summary

- The aim of this study is to test the suitability of the Support Model methodology for use with pupils with Speech and Language Difficulties (SLDs).

- The Support Model assesses a pupil's ability based on the amount of support they need to complete a task. Support is given in the form of systematically delivered prompts, and all pupils complete the task, and so end with a positive sense of achievement.

- Prompts are categorised as: Reading prompts, Understanding prompts, Activation prompts, Writing prompts and Affective prompts. The prompts allow us to use the Support Model to assess the pupils' understanding of the important concepts in a subject, without their Speech and Language Difficulties getting in the way. The scoring system can be altered to include different categories of demand as 'construct relevant'.

- In general, there is threat to validity when we put exam tasks into words to communicate to pupils. The Support Model deals with this by allowing an interaction, so that pupils understand fully what the task is and can show their understanding, even when the task is not clearly communicated via the question. Pupils with SLDs are likely to be more affected than other pupils by construct irrelevant language demands.

- The prompts are standardised, so that every pupil receives, or at least may receive, exactly the same science-related prompts. The assessor gives each pupil every prompt they might need, and they are therefore assessed in a standardised way.

- Participants were 7 pupils aged from 10 to 11 years, in Years 5 and 6 of a primary school in Cambridgeshire with a Speech and Language Centre attached to it. Each pupil had a different profile of SLDs; 3 were classified as 'mild' and 4 as 'moderate'.

- The materials used were three questions from past papers in Key Stage 2 science.

- The pupils showed evidence of understanding the science concepts and being able to carry out the tasks, but had difficulty reading and interpreting the tasks, and putting their answers into appropriate words.

- Three of the pupils with moderate SLDs scored higher on science than on communication for all three questions. This indicates that their science understanding was good but that they would have found it difficult to show this in a traditional exam setting without communication support. Two other pupils, one 'moderate' and one 'mild', showed this pattern on two out of the three questions.

- The results show that these pupils were able to use the prompting system successfully and that it allowed them to show science knowledge and understanding that otherwise may have been masked by their Speech and Language Difficulties.

- Pupils using the Support Model are likely to get higher scores than they would have done without it. In effect there is a bias against all pupils who are not given this kind of language support. The solution is to find a method of establishing a fair equating so that all pupils are assigned the correct level, whether or not they had this language support.

- The present system will need a human interpreter for each pupil, and this assessor will need to have been trained briefly on how to use the prompts. In order to remove the need for a human interpreter, we hope to develop an automated system which will evaluate the pupils' responses and select the appropriate prompt.

- The Support Model method gives us a more valid assessment of the science knowledge and understanding of pupils with SLDs because language related construct irrelevant variance is removed. It is also a better tool for overcoming reading difficulty than Readability Formulae.

Contents

Table of Contents

Executive Summary 2

Contents 4

Introduction 5

The Support Model 5

The 'explain' question 6

Speech and Language Difficulties 6

A Model of the Question Answering Process 7

Validity 8

Methodology 8

Previous research 8

Participants 9

Mild SLDs 9

Moderate SLDs 9

Materials 10

Procedure 10

The Prompts 13

Data Capture and analysis 13

Coding 14

Ethics 14

Results 14

Scoring 14

Interpretation of scores 15

Particular words that caused difficulty 16

Comments and quotes from transcripts 16

Explain questions 16

Understanding the words of the question 16

Activation of the scientific concepts 17

Putting an answer into words 18

Conclusions 18

Scoring 18

Reliability 19

Validity and bias 19

Pupils' evaluation of the prompt system 20

Confidence 20

Manageability 21

Teacher Assessment and Formative Application 21

Summary 22

References 23

Appendix 1 – The questions and mark schemes 25

Appendix 2 – The prompts **34**
Appendix 3 - A sample transcript **38**

Introduction

How can we assess children with Speech and Language Difficulties fairly? Should we adapt existing assessments or should we design something new, specifically targeted to these children? The Support Model offers a third approach: it is a new and innovative method which can be used with existing tasks or completely new ones, to enable every child to perform to their full potential. It allows us to assess children's ability without 'the question getting in the way'.

The standard approach to assessing pupils with SLDs is to modify the paper in a way that is designed to help them to understand the task they are being asked to complete and to communicate their response to the assessors. This leads to concerns about the reliability of the modifications, as assessors try to anticipate the difficulties that particular pupils will face. The Support Model provides a standardised form of prompting pupils to complete a task, and it does so in an interactive and reactive way, tailored to each pupil's needs.

Our aim in this study is to discover the suitability of the Support Model for assessing students with SLDs. We wish to find out whether the Support Model can compensate for the pupils' communication difficulties and so provide a fairer assessment of their science ability. This study is designed to be exploratory, using a small sample of pupils, with the emphasis being on qualitative data collection.

The Support Model

Traditional assessments can be classified as following either a Difficulty Model or a Quality Model (Pollitt, 1990). The Difficulty Model involves a student getting over hurdles, each more difficult than the one before it, until they can no longer clear the hurdle. The point at which they fail gives their level - rather like a high jump. The Quality Model involves the student producing a piece of work that is judged on its quality: students are assessed based on *how well* they performed.

The Support Model differs from the two traditional models as it does not assess how difficult a task a child can complete, or how well a child can complete a given task. Instead, the assessment is based on the amount of support the child needed to complete the task. Support is given in the form of prompts, and assessment is based on the number and nature of prompts needed. With this model all students 'pass' and complete the task, and so end with a positive sense of achievement.

The great advantage of the Support Model is that we can choose to assess children on certain kinds of prompt they might need while ignoring other kinds. For example, when assessing science we are able to judge the pupils on the number of prompts they needed to help them with the understanding and processing of the scientific concepts

in the question, whilst not penalising them for prompts which simply helped them to read and understand the task.

This is an advantage for all pupils as it allows us to assess what is important in a subject area (Pollitt & Ahmed, 2009); if we do not want to assess their reading or even their writing, we can assign a level based only on what we do want to assess. For pupils with Speech and Language Difficulties this has even more obvious advantages. We can use the method to assess the pupils' understanding of the important concepts in a subject, without their Speech and Language Difficulties getting in the way.

The 'explain' question

From our recent studies of questions and mark schemes at GCSE level (Pollitt, Ahmed, Baird, Tognolini & Davidson, 2008), as well as previous work on comprehension failures (Pollitt & Ahmed, 2000), question wording (Pollitt & Ahmed, 2001) and Sources of Difficulties in exam questions (Ahmed & Pollitt, 1999), we have compelling evidence of the effects of the wording of questions on the suitability of a test to assess what it is intended to assess: that is, the threat to validity which comes about when trying to put exam tasks into words to communicate them to the pupils.

Many of these communication issues can be dealt with by ensuring that question writers are aware of the way in which pupils read in an exam, and of how to ensure that the language of a question is as clear as possible. However, we cannot realistically eliminate *all* construct irrelevant language demands from an exam. Furthermore, there are some types of task that are very difficult to communicate clearly in a written paper for any pupil and in particular for those with SLDs. The most challenging of these in our experience is the 'explain question' or the 'why question', which led us to conceive of a new approach to assessment: the Support Model.

We originally designed the Support Model to address the issue of 'the explain question'(Ahmed & Pollitt, 2010) as very often pupils will write an answer to an 'explain' question and examiners reading their script wish they could probe further by asking them 'Yes but why?'. The examiners are looking for a deeper explanation, or the next level of explanation, and often the pupils may know the deeper explanation but not have realised that was what was expected. The Support Model deals with this by allowing an interaction, so that pupils understand fully what the task is and can show their understanding, even when the task has not been clearly communicated by the question.

The application of the Support Model for pupils with SLDs was an obvious next step. These pupils are likely to be more affected than other pupils by construct irrelevant language demands. We chose to use questions from KS2 Science in this project as there is a clear distinction between the science content of the question and the communication of that content.

Speech and Language Difficulties

Speech and Language Difficulties are considered to affect a pupil's ability to understand language and/or to produce language.

In Ofqual's 2008 report (Lampreia-Carvalho, 2008) states that:
"Candidates with delayed language development may:

- omit or substitute sounds
- experience articulation difficulties
- not understand, learn or remember the meaning of words
- fail to understand and use abstract concepts (time, space, quantity, emotions)
- encounter difficulty in finding appropriate words
- interpret language literally
- struggle with verbal reasoning involving cause and effect, deduction, prediction and inference
- find it hard to order events in sequence."

And:
"Contrary to cases of delayed development, speech and language 'disorder' encompasses cases of severe delay and abnormality in the development of language comprehension and/or use. It is very hard for young people with this type of SLD to develop complex language skills."

Currently pupils with these difficulties could be offered a variety of aids such as a reader, an amanuensis, a rest break or an Oral Language Modifier.

The pupils involved in the current study had a range of Speech and Language Difficulties. All attended the Speech and Language Centre in the mornings and the mainstream school in the afternoons. More detail on the specific SLDs of each of the pupils can be found in the Methodology section below.

A Model of the Question Answering Process

The prompting system used for the Support Model was designed using our Model of the Question Answering Process (Pollitt & Ahmed, 1999). This model was developed in order to help assessors to understand the psychological processes involved when a pupil meets an exam question, and so to improve the wording and presentation of questions and mark schemes. The model has six phases:

Phase 0	Learning		
Phase 1	Reading		
Phase 2	Searching	}	
Phase 3	Matching	}	Activation
Phase 4	Generating	}	
Phase 5	Writing		

The learning phase happens before the exam and is of course what we are trying to measure. Pupils then read a question and form an understanding of the task. This phase is prone to many possible pitfalls as the pupil interprets the examiner's

intentions. The following three phases, searching, matching and generating, are processes of activation of concepts in the pupil's mind, which result in an idea of an answer which is then usually put into words in the writing phase. For the purposes of this study we can group Phases 2-4 and call them Activation.

We used the phases of the Question Answering Process to guide the writing of prompts for each question. Thus the prompts needed to help a pupil answer a question can be put into three categories: Reading prompts, Activation prompts and Writing prompts. There is also a fourth category of prompts which has proved critical in all of our studies using the Support Model, and this is Affective prompts to give praise and encouragement to the pupils.

In the current study we adapted the prompt categorisation further and this is described below. (See Methodology – The Prompts).

Validity

How much ability does a pupil have? How well have they learned the important concepts in the subject? These are the questions a valid assessment must answer. The key to achieving this is in what is happening in the pupils' minds while they are answering an exam question. Are their minds doing *the things we want them to show us they can do?*. As assessors we have to be in control of what is happening in the pupils' minds, so that we get the evidence we need in order to make judgements and give credit (Ahmed & Pollitt, 2008).

Validity must be built into any assessment *from the start*. At every step of the assessment process some of this validity can be lost, and this often occurs when there is a breakdown in communication between question writer and pupil or pupil and marker (Pollitt & Ahmed, 2009). Our aim is to reduce any threats to validity from the start, and throughout the process of assessment. The Support Model is a post-hoc approach to circumvent the problems in communicating the task to pupils. Its interactive potential allows pupils to form a clear understanding of any task, even one that has proved difficult to word clearly. Judgements can then be based solely on the amount and nature of support the pupil needed to complete the task, and not on support needed to understand the question or to formulate a response. In this way validity is maintained in the assessment.

Methodology

Previous research

In our previous studies using the Support Model (Ahmed & Pollitt, 2010) we have shown that:

1. the prompting methodology is successful when used with low ability 15 year olds, who are able to respond to the prompts and complete the tasks.

2. the scores generated are valid – these pupils are correctly rank ordered.

3. the prompting system can be implemented on computer, with a human assessor, and these pupils are able to use the interface successfully to complete the tasks.

4. the methodology can be used successfully with 11 year old pupils, who are able to respond to the prompts and complete the tasks (Ahmed, 2009).

A senior computational linguist advised us that full automation of the computer-based procedure would be feasible, given about one person-year of further development.

In this study the aim is to test the suitability of the Support Model methodology for use with pupils with Speech and Language Difficulties. The intention was to use the computerised format with these pupils, but this proved to be inappropriate because of the extent of their reading difficulties. A version with voice-over for the questions and prompts would have overcome this problem, but could not be implemented within the timescale of the current project. There is no difficulty in principle with adding voice-over to the prompts delivered on computer, but it is time-consuming to set this up. Since we have already shown that the computerised version is equivalent to the human version of systematic prompt delivery (Ahmed & Pollitt, 2010), there was no need to do this to establish the efficacy of the method.

Participants

The participants were 7 pupils aged from 10 to 11 years old, in Years 5 and 6 of a primary school in Cambridgeshire with a Speech and Language Centre attached to it. There were six boys and one girl.

Each pupil had a different profile of SLDs and we will describe each pupil, using only an initial to identify them. These descriptions are based on information from the Language Support Advisor.

Mild SLDs

E has good understanding of language but poor speech. Her expressive language is impaired and she often makes errors with pronouns, and has difficulty pronouncing some consonants. She is due to re-enter mainstream school in September.

B has a mild autistic diagnosis. His understanding is good when he is listening but he tends to tune out. He finds it difficult to respond to ‘why’ questions.

N has a speech disorder. His understanding is good and he has recently improved a lot in expressive language. He gets self-conscious and anxious about his speech.

Moderate SLDs

L’s speech is good but he has difficulty understanding language, and his expressive language is disordered. He has problems finding the words he needs to express himself.

T has clear speech and his understanding is not bad. His expressive language is limited and he can get distressed by this.

P has good vocabulary retention but his ability to remember concepts and ideas, and his expressive language are limited. His speech is fine. His difficulties are not specific to language – he also performs poorly on non-verbal tasks.

J has a good understanding of language but has problems with word finding and expressive language. He finds it hard to give explanations and to talk about things that he has done. He can get upset if he is unable to express himself.

Materials

The materials used were three questions from past papers in Key Stage 2 science: Q6 Straw Sounds, Q4 Fish Tank and Q9 Periwinkles. See Appendix 1 for the original questions and mark schemes.

Procedure

The researcher worked individually with each pupil, first talking generally about school and science lessons in order to put the pupil at ease and ensure that he or she was comfortable talking to the researcher. The pupil was then shown the first question and asked to read each part out loud. Help was given when necessary with the reading, and prompts were given towards the completion of each part of each question. The whole interaction was recorded on audio tape to be analysed afterwards. Each session lasted 20-25 minutes. Six out of the seven pupils completed all three questions, and one pupil completed two of the questions and was too tired to continue.

In our previous study using the Support Model (Ahmed & Pollitt, 2010) GCSE questions were presented on a laptop screen and pupils answered by typing their answer in a response window on the screen. Pupils were sent prompts from the assessor's laptop, which then appeared on their screen. There followed a 'dialogue' of responses and prompts which appeared in a messenger chat screen with the full dialogue history available. This method worked very well with the pupils in that study who were aged 15 with no Speech and Language Difficulties.

Figure 1 below shows the pupil's screen from the 2010 study.

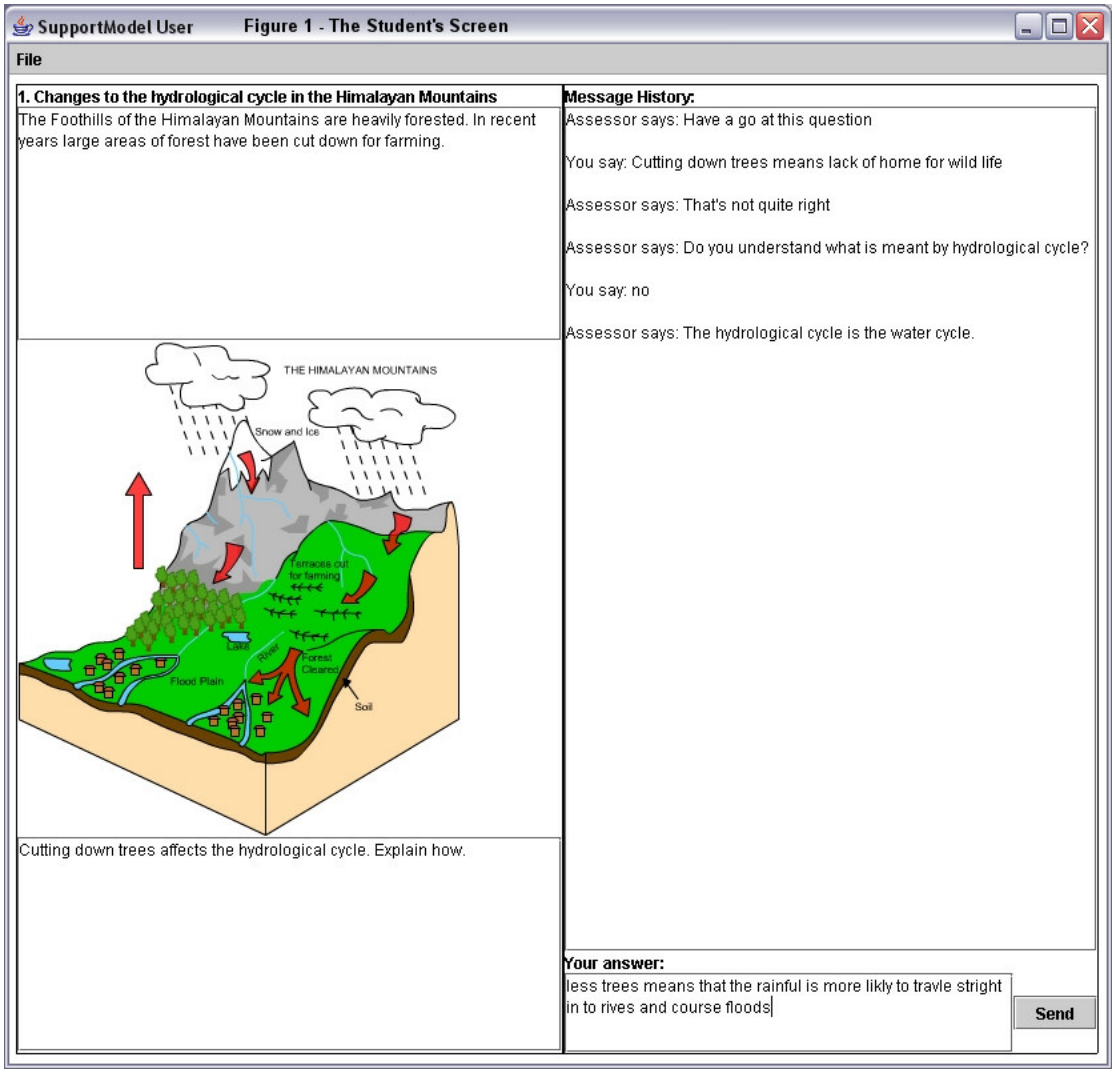
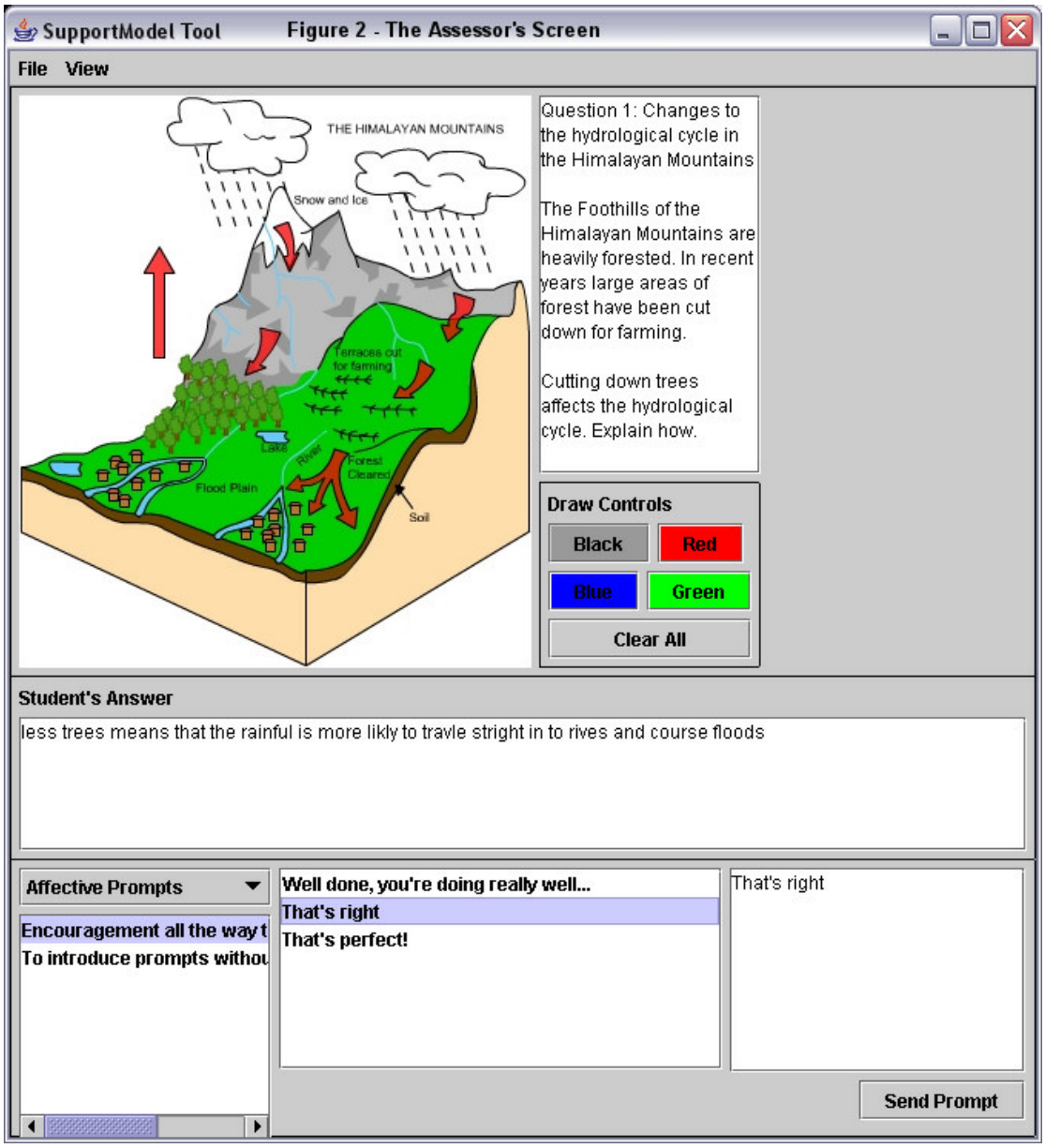


Figure 2 below shows the assessor's screen.



However, on talking to our Language Support Advisor, we decided that computer delivery would not be appropriate for the pupils in the current study. These pupils would need help with the reading of the questions and would not be able to manage prompts presented on screen, but instead would need oral prompts.

It was therefore decided that this study should follow the method used in Phase 2 of our previous study (Ahmed & Pollitt, 2010) in which the assessor gives a series of systematically derived prompts in an oral interview situation. The assessor simulates a computer, delivering a fixed set of prompts in a fixed order. This method was successful in our previous research and proved equivalent to the computer delivered method. In any subsequent study we would deliver prompts to pupils with speech and language difficulties using the screens shown above plus a voice-over. The pupils

themselves could choose whether to use the voice-over, and we would have a record of which pupils clicked for the voice-over and when.

Also, in the current study, pupils were not required to write answers to the questions, but just to answer orally. In any future study pupils would be given the opportunity to type their responses on screen, with an option to answer orally into a voice recorder if needed.

All of this, although time-consuming to set up, is technically quite trivial to add and would not change the conclusions from this study about the feasibility of the method for use with these pupils.

The Prompts

The prompts were based on our psychological model of the question answering process (Pollitt & Ahmed, 1999). For a full list of prompts for each question see Appendix 2.

In the original Support Model studies (Ahmed & Pollitt, 2010) the prompts were categorised as Reading, Activation, Writing or Affective prompts. This categorisation allowed us to develop prompts systematically for each question. Reading prompts help the pupil to understand the task by interpreting the question. Activation prompts help with the concepts in the questions: these are critical in terms of the assessment of the pupils' ability. Writing prompts help the pupil to put their answer into appropriate words.
The prompts are ordered in a strict sequence; for each one, the assessor (or the system if it is automated) decides, on the evidence of responses so far and with reference to the mark scheme, whether or not the pupil needs to be given that prompt.

In the current study we found that we needed an extra category of prompts. These pupils needed prompts to help them to read the words in the question as well as prompts to help them to understand the task by interpreting the words in the question. The Reading prompts category was therefore split into Reading prompts (reading the words) and Understanding prompts (interpreting the words).

Some pupils reached an answer without the need for any prompts, but in other cases the pupil could not reach an answer and in these cases the answer was given, as a final prompt. Affective and motivational prompts were also critical in this study, perhaps even more than in our earlier work.

Data Capture and analysis

Data were captured by audio-recording the session with each pupil and transcribing this interaction. A qualitative data analysis coding procedure was then followed in order to identify how many of each type of prompt were needed by each pupil for each question. The coding system is described below.

Coding

The transcripts were coded according to which prompts were needed by each pupil for each question. The codes were as follows:

- Prompt needed to Read words in question R
- Prompt needed for Understanding of the question U
- Activation prompt needed for scientific concepts A
- Prompt needed to form the Words for an answer W
- Prompt Giving the answer G
- No prompt needed N

Ethics

Approval was gained from the Headteacher of the school involved, and from the head of the Speech and Language Centre at the school. Parental permission was also obtained for all the children who took part. Each child was told that they could choose to stop working through the questions if they felt tired or had simply had enough. One child chose to stop after the second question, but the others were happy to continue to the end of the third question.

Results

Scoring

Each pupil was given a science score for each question. This was calculated by counting the number of A prompts, G prompts and N prompts from all parts of the question and applying the formula:

Raw science score = - A - G + N

As this gave a negative score it was then subtracted from a nominal 40 marks for each question, to give each pupil a positive mark.

Final science mark = 40 + raw science score.

Each pupil was also given a communication score. This was calculated by counting the number of R, U and W prompts:

Raw communication score = - R - U - W

Again a final communication mark was derived by subtracting this from a nominal 40 marks.

Final communication mark = 40 + raw communication score

The table below shows each pupil's final science mark and final communication mark for each question. The consequence of turning the raw score into a positive mark is

that the numbers in the table are not meaningful in themselves, but only relative to each other as indicators of ability.

	Q6 Sci	Q6 Comm	Q4 Sci	Q4 Comm	Q9 Sci	Q9 Comm	Total Sci	Total Comm	Sci minus Comm
Pupil									
E (mild)	32	37	31	34	38	31	101	102	-1
B (mild)	29	39	36	36	31	38	96	113	-17
N (mild)	36	35	34	36	40	33	110	104	6
L (mod)	29	34	42	25	36	33	107	92	15
T (mod)	36	32	38	22	33	25	107	79	28
P (mod)	31	25	32	17	31	14	94	56	38
J (mod)	28	20	32	22	----	----	60	42	(18)
Average	31.6	35.7	35.0	27.4	34.8	29.0			

Interpretation of scores

The scoring system has not been perfected but seems to work well enough. It is a development of the one used in our paper (Ahmed & Pollitt, 2010) which was validated by a senior examiner but has been extended to highlight the communication versus subject knowledge distinction. We would hope to improve the scoring with some further empirical exploration.

In questions 4 and 9 the communication marks are generally lower than the science marks, indicating that these were difficult questions to read and understand, but that the pupils did well on the science.

This is a clear indication of the success of the Support Model approach, as the pupils showed evidence of understanding the science concepts and being able to carry out the task, but had difficulty reading and interpreting the task. Without the Support Model approach we would not be able to assess their ability in science fairly since we would simply not have seen the evidence of their understanding.

Question 6 presents a less clear picture, and it appears that the science itself was also difficult in this question.

Three of the pupils, T, P and J, scored higher on the science than on communication for all three questions. This indicates that their science understanding was good but that they would have found it difficult to show this in a traditional exam setting without communication support. It is significant that all three were in the group with *moderate* language difficulties. The other 'moderate' pupil – L – and one of the 'mild' group – N – showed the same pattern on two out of the three questions.

Looking at the total scores for the pupils with moderate SLDs, there is a clear trend visible with science scores greater than communication scores. The communication difficulties of the pupils classified as having moderate SLDs are clearly shown by their large positive 'Science minus Communication' scores in the final column.

Particular words that caused difficulty

There were some words in the questions that many of the pupils had difficulty reading. Six out of the seven pupils could not read 'caused' and 'identical' in question 6, and 'algae' in question 4. Four out of the seven could not read 'Neil' and 'scrapes' in Question 4. Six of the pupils completed Question 9, and four out of the six could not read 'periwinkles' or 'seashore'.

Comments and quotes from transcripts

The following examples from the transcripts illustrate how the pupils responded to the prompts.

Explain questions

There were some examples of 'explain' questions in which pupils gave a first level of explanation without going as deep as the assessor wanted. One example is Question 6 d (ii). One pupil responded as follows:

> Cos one straw she might be blowing not so hard and the other one she might blow really hard.
>
> *Exactly right so why does that mean it's not a fair test?*
>
> Because if you blow hard and not so hard it will make a different sound.
>
> *Yes and why would that mean it was not a fair test?*

Similarly in Question 4 c:

> Because one of them will have a S on and one of them will have an N on
>
> *Thats right yeah ok and so what happens?*
>
> When the one with S on goes closer it will connect
>
> *Thats right it will connect do you know what the word is for that in science?*
>
> No
>
> *Attraction – they are attracted to each other.*
> *So you're absolutely right they connect together well done.*

Understanding the words of the question

Question 9 b

Do you know what a function is?

No

It's something the shell is used for. Can you describe a function of the shell?

Silence

What do you think the Periwinkle's shell is for?

To protect the animals

Well done that's absolutely right well done.

Activation of the scientific concepts

Question 9 a

Periwinkle B is the big one and Periwinkle A is the little one

That's right but can you think of something that isn't to do with size?

The other one's taller?

That's also to do with size isn't it?
Anything else that looks different?

One is white one is grey

That's right the colour is different. Well done.

……

What do you think might vibrate to cause the sound?

The straw

The straw well done what else do you think might vibrate?

Umm umm she's blowing

Yes what happens when she blows?

You go drrrrrrrrrrrr

That's right it does and what comes out of her mouth when she blows?

Air

That's right and that's what they want you to say, fantastic well done.

Putting an answer into words

Question 4 c

Puts fists together and moves them upwards and makes a scraping sound.

Right you've shown me. Can you tell me?

Cos there was a glass and magnetic can go through glass. And if there was like a magnetic um um metal tank then magnetic B would go but magnetic A sticks. They stick to each other

Do you know what it's called when one magnet sticks to another one?

Um they stay together?

What is the word for that in science?

Silence

They are attracted to each other.

The results show that these pupils were able to use the prompting system successfully and that it allowed many of them to show their science knowledge and understanding that otherwise may have been masked by their Speech and Language Difficulties.

Conclusions

Scoring

In this study, we used a simple scoring system to measure the amount of support with science concepts each pupil needed in order to complete the task successfully. The resulting scores are plausible enough to show that the method can generate valid assessments of the scientific understanding of pupils with SLDs. This needs to be confirmed through a validation study on a larger scale in which subject specialists could confirm the assessment judgements on a greater number of questions for a greater number of pupils, as in our previous study (Ahmed & Pollitt, 2010).

One issue which would need discussion with a subject specialist would be when it might be appropriate to penalise for reading, for example of technical terms, and when it is not appropriate because it is part of normal reading. The point is that in most exam questions the reading demands are not construct relevant demands: we do not want pupils to lose marks if they cannot read the question. The role of the question is simply to communicate the task (Pollitt & Ahmed, 2009). However, there

are a few instances (apart from the obvious reading tests) in which an examiner may consider the reading of a technical term to be a construct relevant demand. It is not always easy for a non-subject specialist to distinguish these two, for example the technical use of an everyday word or phrase, such as 'reaction' in Chemistry or Physics, or 'power' in Physics or Politics.

We specifically set out to assess science understanding, and to exclude the ability to communicate that understanding in writing or orally. The Support Model allows us to choose which demands we consider relevant, or important, for any particular assessment: if the ability to communicate is to be included in the assessment then we would simply modify the scoring scheme so that pupils are penalised for needing 'writing' prompts.

Reliability

There are two key 'threats' to reliability that need to be considered. As in any assessment, the internal consistency of the questions needs to be high, and for the Support Model we need to be sure that the prompting system does not affect this. Because the prompts are generated from a model for ensuring the validity of the assessment process, they are expected to maintain high construct relevance – higher in fact than might be expected without prompting. There is no reason therefore to expect internal consistency to be any lower for a Support Model system than for a standard system.

The second issue is the role of the assessor. It is important to understand that the prompts are standardised, so that every pupils receives, or at least may receive, exactly the same science-related prompts. The assessor gives each pupil every prompt they might need, and they are therefore penalised in a standardised way. The assessor is free to give affective and motivational prompts whenever they seem appropriate, and these should not affect scores in a way that would threaten reliability.

The modification of the language in the Support Model method is standardised, and theoretically driven. A clear distinction is drawn between instances where this change affects the pupil's understanding of the science, where a penalty is applied, and those where only communication is affected, where no penalty is applied. The assessor is not free to make any other modifications than those specified, and so the Support Model does not suffer from the same kinds of threat to reliability that other marking methods may.

Validity and bias

The Support Model method gives us a more valid assessment of the science knowledge and understanding of pupils with SLDs because language related construct irrelevant variance is removed. It is a better tool for overcoming reading difficulty than Readability Formulae, since these are only valid for use with extended natural texts and are therefore difficult or even impossible to apply to questions with a small amount of text, with subject-specific terms, and with diagrams and graphs (Allan, McGhee and van Krieken, 2005). In principle, there is no reason to try to set the

readability level of an exam to suit the age or level of the pupils, but there is every reason to make the language as clear as possible for every pupil in every case. The only purpose of an exam question is to communicate the task as clearly as possible to the pupils (Pollitt & Ahmed, 2009).

Pupils using the Support Model are likely to get higher scores than they would have done without it. If they are now to be compared to pupils in general at this age, are we introducing a bias in favour of pupils with SLDs?

Pupils who do not have specific SLDs but have difficulty understanding the questions and therefore the tasks being set will score lower than their science understanding suggests they should. This was established in our previous study (Ahmed & Pollitt, 2010). In effect there is a bias against all pupils who are not given this kind of language support.

If we were to extend this method to assess all pupils using the Support Model, we would remove the issue of bias. However, this may not be feasible in practice. The alternative solution is to find a method of establishing a fair equating so that all pupils are assigned the correct level, whether or not they had language support. There are significant difficulties with equating in this context, and it may be that simply extending the method for all pupils is the preferred solution.

Pupils' evaluation of the prompt system

The Support Model method was successful in assessing the 11 year old pupils with SLDs. The pupils were able to respond to the prompts and use them as support towards completing the tasks. They were able to respond to activation prompts that helped with the science in the questions, and also to prompts that helped them to read and understand the questions and turn their answers into suitable words.

The pupils reported enjoying the assessment process. One pupil commented 'I like these I do' halfway through the session.

Confidence

The Support Model method can also help pupils who lack confidence: it encourages them to make a first attempt at a task that seems too difficult, as they know that support is available if they get stuck. In a traditional setting they might have simply left that question blank and scored zero. This again raises the issue of bias and equating as discussed above.

The issue of confidence can be illustrated by an example:

Periwinkles Q 9e *Describe what happened to the water in the rockpool during the five hours.*

> Well it's got a bit smaller
>
> *What has happened to the water?*

I don't know.

Where has some of the water gone?

It might have evaporated.

Perfect that's exactly right. Well done.

This pupil knew the answer but started by saying he didn't know, then gave a perfect response after a prompt.

Manageability

The Support Model procedure will be easy to implement on a computer screen with a voice-over for pupils with reading difficulties. Those with difficulties in expressing their answers are supported by the prompts which help them to formulate their response. Some may be able to respond by typing; for others the human assessor will need to interpret oral responses to select the next appropriate prompt.

In all cases, the present system will need a human interpreter for each pupil, and this assessor will need to have been trained briefly on how to use the prompts. In order to remove the need for a human interpreter, we hope to develop an automated system which will evaluate the pupils' responses and select the appropriate prompt. This requires the use of language processing software similar to the automarking software developed by Sukkarieh, Pulman & Raikes (2003) or a dialogue system which could be trained and tested using transcripts from the current approach. A dialogue system would identify patterns in the pupils' responses which could trigger the selection of appropriate prompts. More detail on these systems is available in our paper (Ahmed & Pollitt, 2010).

In our recent meeting with Ofqual and Synapse (May 2010) we discussed the possibility of using a Computer Adaptive Testing (CAT) system to implement the Support Model. The system would be designed to decide on the basis of previous responses, whether or not the next prompt is needed. The Support Model is significantly different from a standard CAT system in that it would involve qualitative analysis of responses rather than declaring them right or wrong. In the interim a human assessor would be needed to monitor the effectiveness of the system.

Teacher Assessment and Formative Application

In order to use the Support Model in the classroom we would not need to wait for the development of an automated system. Teachers could use the Support Model to inform their judgements about pupils' current level. This would involve the teacher working with the pupil on a computerised system in which the teacher decides which prompts are needed and which are not needed, thereby gaining a clearer understanding of the pupils' strengths and weaknesses. In particular the teacher can more clearly distinguish between difficulties caused by communication problems and those caused by lack of subject knowledge.

Once we have an early and imperfect automated system it can be used formatively in the classroom. The prompts would be delivered automatically, providing support for the pupil, and leading always to an appropriate answer to the question. If, sometimes, the system delivers an unnecessary prompt there will be no serious consequence, and any argument between the pupils and the system may even be rewarding.

In both of these applications the low stakes nature means that the same few questions with their system of prompts can be used many times.

Summary

There are clear advantages in using the Support Model to assess pupils with SLDs. We can assess their subject knowledge and understanding without the question getting in the way. In this study we were able to assess the pupils, giving each one a science mark and a separate communication mark. All pupils were able to complete the tasks, some with more support than others on the science, and some with more support with the language. The pupils ended their sessions with a positive feeling of achievement, and reported enjoying the experience of assessment.

The method can be extended to assess other subjects: a series of prompts for mathematics questions would be easy to produce and there are existing interactive assessments for mathematics that use devices similar to our prompting system, although they are based on components of the task rather than a theoretical model of the question answering process (e.g. McAlpine & Ware, 2003). The application of the method to the assessment of English is considerably more complex and would require exploring the issue of what would constitute valid language support. For example, questions should still not get in the way of pupils' ability to respond, even in a language exam.

The method can be standardised further, by delivering the prompts on-screen, adding a voice-over facility for pupils with SLDs: this would not require very much additional work. Full automation, using language processing software to evaluate responses and select prompts, is feasible but would need a separate study to build an appropriate system. Scoping work carried out within UCLES a few years ago suggested that this might require about one year's development by a post-doctoral researcher. The method could then be used in a much wider range of ways to support learning and assessment.

We have shown that the Support Model is suitable for assessing pupils with SLDs. In principle it could be used with every pupil and many would benefit from the increased clarification of the task, in particular with 'explain' questions.
In any valid assessment no pupil should find that they cannot understand what they are expected to do, and the readability formula approach does not address this appropriately. Awarding bodies should make every effort to ensure that language is comprehensible to all pupils.

The most competent pupils almost never need help to understand the task, although even they can have problems with 'explain' questions; at the other end there will always be some pupils who need considerable help (e.g. pupils with severe SLDs). The problem of communication can be tackled using language modification (see for example the guidance in the BATOD/NATED Training Materials on Language Modification) which aims to allow as many pupils as possible to understand exam tasks without the need for interactive support. The Support Model tackles the problem from the other end by aiming to provide each pupil with just the support they need, while ensuring that the support is delivered in a completely standardised way. Thus reliability – and validity – can be optimised.

From our previous work we know that the Support Model method can generate scores that are considered reliable and valid by an expert examiner. This relates to reliability *within* the group of pupils assessed using the Support Model, but does not deal with the issue of *equating* these scores to those of a group assessed in the traditional way: this would need to be addressed separately, perhaps involving a separate Level Setting event.

The Support Model is based on the use of standardised prompts which are theoretically grounded and systematically generated. This is a significant improvement on any approach which either pre-determines the wording modification before any interaction with the pupil, or is reactive but non-systematic.

References

Ahmed, A. (2009) Conceptualising an assessment for gifted students – how can we assess creativity? *Paper presented at the International Association for Educational Assessment conference, Brisbane, September 2009.*

Ahmed, A. & Pollitt, A. (2010) The Support Model for Interactive Assessment. *Assessment in Education 17 (2) pp 133-167.*

Ahmed & Pollitt (2008) It's judging the evidence that counts – introducing OSCA. *Paper presented at the International Association for Educational Assessment conference, Cambridge, September 2008.*

Ahmed & Pollitt (1999) Curriculum demands and question difficulty. *Paper presented at the International Association for Educational Assessment conference, Slovenia, May 1999.*

Allan, S. McGhee, M. & van Krieken, R. (2005) Using readability formulae for examination questions. *Report to QCA.*

BATOD/NATED Training Materials on Language Modification
http://www.batod.org.uk/content/articles/resources/training-materials/language-modification/Training%20materials%20for%20language%20modification.pdf

Lampreia - Carvalho, F. (2008) Oral language modifiers focus group. *Ofqual report*.

McAlpine, M. & Ware (2003) Introducing computer-assisted assessment in Scotland: Laying the foundations for an integrated approach. *Paper presented at the International Association for Educational Assessment conference, Manchester, October 2003.*

Pollitt, A. & Ahmed, A. (2009) The importance of being valid. *Paper presented at the Association for Educational Assessment – Europe conference, Malta, November 2009.*

Pollitt & Ahmed (2000) Comprehension failures in educational assessment. *Paper presented at the European Conference on Educational Research, Edinburgh, September 2000.*

Pollitt & Ahmed (2001) Science or reading? How students think when answering TIMSS questions. *Paper presented at the International Association for Educational Assessment Conference, Rio de Janeiro, May 2001.*

Pollitt, A & Ahmed A. (1999). A New Model of the Question Answering Process. *Paper presented at the International Association for Educational Assessment Conference, Slovenia, May 1999.*

Pollitt, A., Ahmed, A., Baird, J-A., Tognolini, J. & Davidson, M. (2008) Improving the quality of GCSE Assessment. *Report to QCA.*

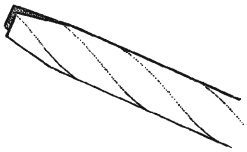
Pollitt, A (1990). Giving students a sporting chance: assessment by counting and by judging. In Alderson, C. & North, B. (Eds) *Language Testing in the 1990s*, pp 46-59. Macmillan: London.

Sukkarieh, JZ, Pulman, SG & Raikes, N (2003). Auto-marking: using computational linguistics to score short, free text responses. *Paper presented at the International Association for Educational Assessment conference, Manchester 2003.*

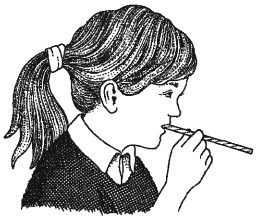
Appendix 1 – The questions and mark schemes

6 **Straw sounds**

(a) Polly has a straw.
She cuts one end of the straw.



She blows into the cut end of the straw.
It makes a sound.



The sound is caused by vibrations.

Name **TWO** things that vibrate to cause this sound.

6a
1 mark

..... and .....

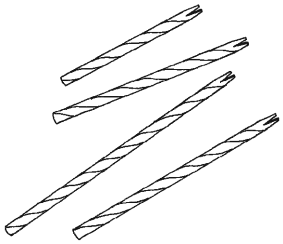
(b) Polly thinks that changing the length of the straw may change how high or low the note is.

What is the scientific name for how high or low a note is?

6b
1 mark

.....

(c) Polly cuts four identical straws into different lengths.



Her friends blow gently into the straws. The note from each straw is different. Some notes are high and some are low.

Describe how the length of a straw affects how high or low the note is.

.....

.....

6c
1 mark

(d) (i) Tania says 'Polly's test is not a fair test because a different person is blowing into each straw.'

Why might Polly's test **not** be a fair test if different people blow into each straw?

.....

.....

6di
1 mark

(ii) Polly says 'It might not be a fair test even if one person blows into each straw.'

Explain why it might **not** be a fair test even if one person blows into each straw.

.....

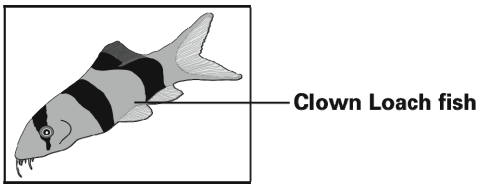
.....

6dii
1 mark

Total out of 5

4 Fish tank

(a) Neil has a tropical fish tank. He has Clown Loach fish in his tank.



Algae also grow in his tank. Algae are small green plants.
Neil moves his fish tank from a dark corner into the sunlight.

Tick **ONE** box to show what will happen to the algae in Neil's tank when he puts it into sunlight.

More sunlight will make the algae (green plants)...

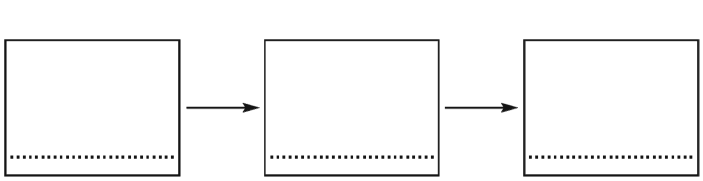
stop growing.	☐	die.	☐
grow more quickly.	☐	turn yellow.	☐

4a ☐ 1 mark

(b) Neil wants to keep his fish tank clean of algae. He knows **water snails** eat **algae**. Neil decides to buy water snails.

He then reads that **Clown Loach fish** eat water snails.

Write the food chain for the Clown Loach fish.



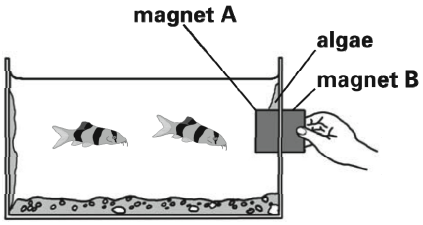
4b ☐ 1 mark

(c) Instead, Neil uses two magnets to clean algae off the sides of the tank.

He puts magnet B on the outside and magnet A on the inside.

Neil moves magnet B on the outside of the tank.

As he moves magnet B, magnet A moves with it.



Magnet A scrapes away the algae inside the tank.

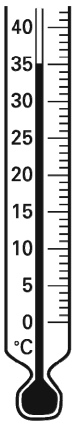
Explain why magnet A moves with magnet B.

.....

4c
1 mark

(d) Neil must keep the temperature of the water at 25°C to keep his fish healthy.

This thermometer shows the temperature of the water in Neil's tank.



(i) Tick **ONE** box to show if the temperature of the water will keep the fish healthy.

yes ☐ no ☐

(ii) Explain your answer.

.....

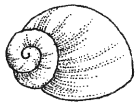
.....

4d
1 mark

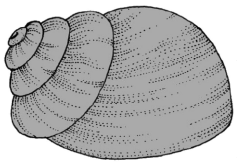
Total out of 4 ☐

9 Periwinkles

(a) Periwinkles are animals with shells that live on rocks at the seashore.



Shell of periwinkle A



Shell of periwinkle B

The shell of periwinkle A is smaller than the shell of periwinkle B.

Describe **ONE other** way the shell of periwinkle A is different from the shell of periwinkle B.

9a

1 mark

.....

(b) Periwinkles do not have bones inside their bodies but they do have a shell. The shell does **not** help the periwinkle to move.

Describe **ONE** function of the shell.

9b

1 mark

.....

(c) A scientist wants to find out if the area the periwinkles live in affects the size of their shells. He measures a sample of 20 periwinkle shells from two different areas of the seashore.

Why does he measure 20 periwinkle shells from each area instead of just one periwinkle?

.....

.....

9c
1 mark

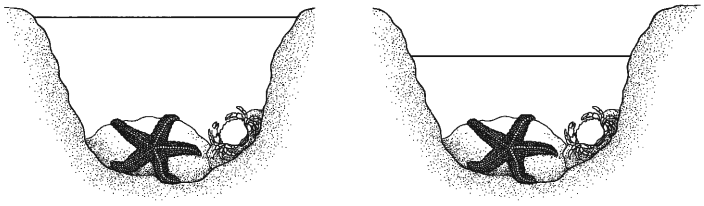
(d) Explain why it is important to return the animals to the same place they were collected from.

.....

.....

9d
1 mark

(e) The scientist measures how much water is in a rockpool. After five hours the water level in the rockpool is lower. No waves splashed into the rockpool during this time. No water could leak out.



Rockpool **Rockpool after 5 hours**

Describe what happened to the water in the rockpool during the five hours.

.....

9e
1 mark

Total out of 5

Test A question 6: *Straw sounds*

Question	Mark	Requirements	Allowable answers	Additional guidance
6a *4/3e,g*	**1m**	Award **ONE** mark for **two** correct responses [given in either order]: ■ the straw/the cut (straw) ■ the air/wind.	**ONE** mark may be awarded for: ■ plastic/paper ■ her mouth/lips.	***Do not*** *give credit for an insufficient response:* ■ *breath/voice* ■ *bits you cut out* ■ *eardrum [does not vibrate to produce the sound]* ■ *sound/air waves* ■ *blowing.* ***Do not*** *give credit for an insufficient response that is a restatement or repetition of the first.*
6b *4/3f*	**1m**	Award **ONE** mark for: ■ pitch. ❖ **Give credit** for a correct response that goes beyond the key stage 2 programme of study: ■ frequency.		***Do not*** *give credit for a response that includes incorrect science:* ■ *volume.* ***Do not*** *give credit for an insufficient response:* ■ *tone [refers to the quality of the sound]* ■ *note [may refer to duration or pitch of sound].*
6c *4/3f*	**1m**	Award **ONE** mark for an indication that a longer straw creates a lower note/sound, and/or a shorter straw creates a higher note/sound: ■ the longer the straw, the lower the note ■ it is higher when the straw is short ■ the longest straw makes the lowest noise.	**ONE** mark may be awarded for: ■ a long straw makes a low sound ■ a short straw makes a high sound.	***Do not*** *give credit for a response that includes incorrect science:* ■ *the longer the straw, the louder/softer the sound* ■ *the longer the straw, the longer the note.*

Test A question 6: *Straw sounds (continued)*

Question	Mark	Requirements	Allowable answers	Additional guidance
6d i *1/2d*	**1m**	Award **ONE** mark for an indication that different people will blow in different ways: ■ they will not know if their results are caused by the length of the straw or by the way each person blew ■ some people might blow harder than others ■ they will not all blow the same ■ they might blow at different speeds.		***Do not*** *give credit for an insufficient response that refers to the independent variable.* ■ *the straws are all different lengths.* ***Do not*** *give credit for an insufficient response referring only to breathing, not to air entering the straw:* ■ *different people have different breathing* ■ *people breathe at different times.* ***Do not*** *give credit for an insufficient response referring to the amount of time that an individual blows in the straw:* ■ *they can blow for different lengths of time.*
6d ii *1/2d*	**1m**	Award **ONE** mark for an indication that one person might blow differently each time: ■ one person will not be able to blow exactly the same each time ■ they might blow harder on some straws than others.		***Do not*** *give credit for an insufficient response that refers to the independent variable:* ■ *the straws are all different lengths.* ***Do not*** *give credit for an insufficient response referring only to breathing, not to the air entering the straw:* ■ *she might run out of breath* ■ *she might breathe differently each time.* ***Do not*** *give credit for an insufficient response referring to the amount of time the individual blows:* ■ *she might not be able to blow for as long into the last straw.*

When applying this mark scheme, please also refer to the *General guidance* given on pages 1 and 2.

Test A question 4: *Fish tank*

Question	Mark	Requirements	Allowable answers	Additional guidance
4a 2/3a	**1m**	Award **ONE** mark for: ☐ ☐ ■ grow more quickly. ☑ ☐		
4b 2/5d	**1m**	Award **ONE** mark for **all three** organisms in the correct order: ■ algae→(water) snails→(Clown Loach) fish.	**ONE** mark may be awarded if (green) plant is used in place of algae in an otherwise correct food chain: ■ (green) plant→snail(s)→fish.	***Do not*** *give credit for an alternative food chain for the Clown Loach fish.*
4c 4/2a	**1m**	Award **ONE** mark for an indication that there are forces of attraction between the magnets **or** that the magnets have opposite poles facing each other: ■ the magnets are attracted to each other ■ because opposite poles attract ■ the North pole of one magnet is facing the South pole of the other magnet.	**ONE** mark may be awarded for: ■ the magnets stick/stay together ■ magnet B pulls magnet A with it. **ONE** mark may be awarded for a response indicating the magnetic force extends through the glass/tank: ■ the force works through the tank.	***Do not*** *give credit for an insufficient response that does not make explicit that the magnets attract:* ■ *there is a force between the magnets* ■ *it moves because they are magnetic* ■ *they connect together.* ***Do not*** *give credit for an insufficient response implying each magnet is a different pole:* ■ *they are opposite poles* ■ *magnet B is South and magnet A is North.*

Test A question 4: *Fish tank (continued)*

Question	Mark	Requirements	Allowable answers	Additional guidance
4d 1/2e 3/2c	**1m**	Award **ONE** mark for: ■ ☐ no ☑ **AND** a response that indicates the temperature is higher than it should be: ■ the water is too hot ■ the water needs to be cooler ■ the water is 10°C higher than it should be ■ it is meant to be 25°C but it is at 35°C.	**ONE** mark may be awarded for: ■ ☐ no ☑ **AND** ■ the tank is too hot/warm ■ there is 10 degrees difference between how the water is and what it needs to be for the fish to be healthy ■ the water/tank is not at 25(°C) ■ the water is at the wrong temperature ■ the water is at 35°C. **ONE** mark may be awarded for: ■ ☐ no ☑ **AND** a response where the thermometer has been misread (providing it is above 25°C) or the temperature difference is incorrect but a higher reading is indicated: ■ the thermometer says the water is 36°C which is too high ■ the temperature of the water is 8°C higher than it should be ■ the water is at 36°C. **ONE** mark may be awarded if neither box is ticked but the creditworthy explanation indicates that the pupil believes the water is too hot for the fish to be healthy.	***Do not*** *give credit for a response where the 'yes' box has been ticked.* ***Do not*** *give credit for a response that includes incorrect science implying that the water temperature needs to be hotter or that the current temperature is too low:* ■ *the water temperature should be higher.* ***Do not*** *give credit for an insufficient response where the 'no' box has been ticked but the explanation does not give a clear indication that the water temperature is incorrect:* ■ *because it will not keep the fish healthy* ■ *the higher the temperature, the more unhealthy the fish.* ***Do not*** *give credit for an insufficient response where the 'no' box has been ticked but is followed by information given in the question:* ■ *25°C keeps the fish healthy* ■ *it needs to be at 25°C for the fish to be healthy.* ***Do not*** *give credit for an insufficient response where* ***neither*** *box has been ticked and the explanation does not make it clear that the fish will not be healthy at this temperature:* ■ *the water is at 35°C.*

When applying this mark scheme, please also refer to the *General guidance* given on pages 1 and 2.

Test B question 9: *Periwinkles*

Question	Mark	Requirements	Allowable answers	Additional guidance
9a *1/2f*	**1m**	Award **ONE** mark for a response giving a difference between the shells of the two periwinkles: ■ A is a lighter/different colour ■ the shape (of shell) ■ the spirals are in different places on the shells.	**ONE** mark may be awarded for a response that refers to shell A as 'it': ■ it has a lighter shell.	***Do not*** *give credit for an insufficient response that does not describe how the shells are different:* ■ *it is a different size [given]* ■ *it is a different type (of shell).* ***Do not*** *give credit for an insufficient response that is correct for shell B but does not explicitly refer to shell B:* ■ *it has a darker shell.*
9b *2/2e* *2/5b*	**1m**	Award **ONE** mark for a response that indicates the shell protects or supports the periwinkle's body: ■ the shell protects the periwinkle ■ it supports the periwinkle.	**ONE** mark may be awarded for: ■ it stops the periwinkle drying out ■ it keeps the periwinkle safe ■ it shelters the periwinkle ■ it helps them to hide (from predators). **ONE** mark may be awarded for indicating why it is useful to have a shell that is hard or difficult to break and implies protection: ■ the shell will not break when a crab tries to crack it open.	***Do not*** *give credit for a response that includes incorrect science:* ■ *it helps the periwinkle to move [given that it does not].* ***Do not*** *give credit for an insufficient response:* ■ *the shell/it is hard/keeps it warm* ■ *it cannot break easily [gives only a property of the shell]* ■ *the shell is the periwinkle's house* ■ *so the periwinkle does not get wet/rained on.*
9c *1/2g*	**1m**	Award **ONE** mark for a response that indicates an average size is needed **or** that the reliability of his results/test/ conclusions will be improved: ■ if he measures only one or two shells, they may be smaller than the average ■ because if he only measured one/a few periwinkle shells it/they might be bigger than the others and then his results will be useless ■ to check the results.	**ONE** mark may be awarded for a response that indicates there is a variation in (the size/type of) the periwinkle shells in each area: ■ to see if they are all different/all the same (size or age) ■ there may be more than one type of periwinkle. **ONE** mark may be awarded for a response indicating that the accuracy of the test would improve: ■ the average will be more accurate.	***Do not*** *give credit for a response that includes incorrect science:* ■ *to make the results different.* ***Do not*** *give credit for an insufficient response indicating that the shells from one place should be the same:* ■ *to make sure they are the same.* ***Do not*** *give credit for an insufficient response:* ■ *to get the right/correct results* ■ *to see if the periwinkles in the different areas are different.*

Test B question 9: *Periwinkles (continued)*

Question	Mark	Requirements	Allowable answers	Additional guidance
9d *2/5a*	**1m**	Award **ONE** mark for an indication that the periwinkles may not be able to survive in a different place: ■ the periwinkles may not be adapted/ suited to living in a different place ■ the periwinkles may not live if they are put somewhere new. ❖ **Give credit** for a correct response that goes beyond the key stage 2 programme of study, implying that an environment will be adversely affected by the introduction of a new species: ■ it may deplete the environment/food chain of existing organisms.	**ONE** mark may be awarded for: ■ so they do not die ■ they will be safer in their own habitat ■ they are used to their habitat/home/ where they live ■ it is their habitat ■ the conditions could be different.	***Do not*** *give credit for an insufficient response:* ■ *it is cruel not to put them back [does not explain why it is cruel]* ■ *it is where they live [given]* ■ *it is their home* ■ *it is where their family is* ■ *so they do not get lost* ■ *because they will die [need to qualify].*
9e *3/2d* *1/2l*	**1m**	Award **ONE** mark for an indication that (some of) the rockpool water has evaporated: ■ some of the water/it evaporated ■ the water has changed to water vapour.		***Do not*** *give credit for an insufficient response indicating that the water has leaked or seeped out of the pool [given that it has not].* ***Do not*** *give credit for an insufficient response:* ■ *there is less water* ■ *the water is lower* ■ *the crab drank it.*

When applying this mark scheme, please also refer to the *General guidance* given on pages 1 and 2.

Appendix 2 – The prompts

(Prompts for reading the words of the question are not listed as they simply involved giving the word)

Prompts for Straw Sounds

6a

Do you know what Vibrations are?	U
Vibrations are small movements that cause a sound.	U
What do you think might vibrate to cause the sound?	A
What else do you think might vibrate?	A
What happens when she blows?	A
What comes out of her mouth when she blows?	A
What moves through the straw when she blows?	A
What is her breath made of?	W
The air vibrates.	G
The straw vibrates.	G

6b

Do you know what is meant by the note?	U
The note is the sound she makes by blowing through the straw.	U
Do you know the word for how high or low a note is?	A
Do you know what is meant by a high note and a low note?	A
Can you sing a high note and a low note?	A
How would you describe the difference between the two notes?	W
Do you know the word for that in science?	W
It is called the pitch.	G

6c

Do you understand what identical means?	U
They are identical before they are cut shorter.	U
Then they are all different lengths.	U
The ends are cut as well.	U
Do you know which straw will produce the highest note?	A
What makes the note high?	A
Which straw will produce the lowest note?	A
What makes the note low?	A
How does the length affect the note?	A
The question.	A
The longest straw will give the lowest note.	G
The shortest straw will give the highest note.	G

6di

Do you know what is meant by a fair test?	U
Why might it not be a fair test?	A
Will everyone blow the same?	A
Do you think some people might blow harder than others?	A
Why isn't it a fair test if they all blow differently?	A
Can they tell what causes the different notes?	A
If they don't all blow the same we can't tell what is causing the different notes so it is not a fair test.	G

6dii

Do you understand what is happening here?	U
Do you understand how it is different from the first part?	U
This time the same person is blowing into all of the straws.	U
Do you think she will always blow the same?	A
Why is that not a fair test?	A
If she doesn't blow the same each time we cant tell what causes the different notes.	G

Prompts for Fish tank

4a

What will happen to the algae with more sunlight?	A
What happens to plants when they get sunlight?	A
What about if they get just the right amount of sunlight?	A
They grow more quickly.	G

4b

Do you know what a food chain is?	U
You need to write in these boxes.	U
You need to write down what is getting eaten by what.	A
The arrow means 'is eaten by'. Can you put them in the right boxes?	A
What gets eaten by the water snails?	A
What gets eaten by the Clown Loach Fish?	A
What eats the algae?	A
What eats the water snails?	A

4c

Can you tell me what is happening in this picture?	U
What is Neil holding?	U
What is he doing with the magnet B?	U
What is happening to Magnet A?	U
What is happening to the Algae?	U
Why does Magnet A move when Magnet B moves?	A
Why do they stick together?	A
What kind of force is it?	A
What is it called when one magnet is pulled towards another one?	W
What is it called when they stay/stick together?	W
What is the word for that in science?	W
Attraction. They attract.	G
They are attracted to each other.	G
The North pole of one magnet is attracted to the South pole of the other and vv.	G

4di

What temperature does it say the water is in Neil's tank?	U
What temperature is shown on the thermometer in the picture?	U
What temperature does it need to be to keep the fish healthy?	A
Is this the same?	A
No.	G

4dii

Is 35 hotter or colder than 25?	A
So why wont the fish keep healthy?	A
Is the water too hot?	A

Prompts for Periwinkles

9a	
What looks different?	U
Can you think of something that isn't to do with size?	A
Anything else apart from the size?	A
Is the colour different?	A
Which one is darker?	A
Is the shape different?	A
Periwinkle B is darker than Periwinkle A	G
They are different shapes	G
9b	
Do you know what a function is?	U
A function is something the shell is used for.	U
It is a way in which the shell helps the periwinkle.	U
What do you think the Periwinkle's shell is for?	A
Why might it need a shell?	A
What do you think it would be like without a shell?	A
So how can the shell help?	A
Why is that useful?	A
Where do periwinkles live? Why might they need a shell?	A
The shell protects the Periwinkle	G
9c	
Do you understand what the scientist is doing?	U
What does he want to find out?	U
Do you know what a sample is?	U
It is a selection of periwinkles. He chooses some of the periwinkle shells to measure from each area.	U
He chooses 20 from each area.	U
Why would measuring 20 be better than measuring one?	A
Would all the periwinkles in one area be the same size?	A
If they aren't all the same size, why is it helpful to measure 20 of them?	A
What can the scientist find out if he measures 20 that he cant find out if he measures only one?	A
If he only measures one it might be an extra big one or an extra small one. He needs to measure 20	G
9d	
Why should they go back where they were found?	U
What would happen to the periwinkles if they were put back in a different place?	A
Why might they die if they were put back in a different place?	A
Do you know another word for an animal's home?	A
Why do they need to be where they normally live?	A
They may not survive in a different place.	G
They may not be adapted to living in a different place.	G
9e	
Do you understand what is happening here?	U
What has happened to the water?	A
Where has some of the water gone?	A
That's not it. Where else could it have gone?	A
It has gone into the air.	G

Do you know the scientific word for this? W
It has evaporated. G

Appendix 3 - A sample transcript

Straw Sounds

Reads Q 6a with help with **straw, caused, vibration, vibrate**
Silence
Do you know what vibrations are? *U*
Yeah when you tap things it vibrates.
That's right.
If we tap the lead it vibrates.
You're right so can you name two things that vibrate to cause the sound? *A*
The straw.
Perfect and what else? *A*
Her mouth.
You're right but what comes out of her mouth? *A*
The air.
Fantastic.

Reads Q 6b with help with **changing, note, scientific**
We turn volume up.
Yes volume is how loud it is but do you know the word for how high or low a note is?A
No..
Do you know what is meant by a highnote and a low note? *A*
Yes.
Do you know the word for that in science? *W*
No.
Shall I tell you?
Yes.
It's called the pitch. Have you heard that before? *G*
Yeah in out big class.

Reads Q 6c with help with **identical, different, gently, describe, affects**
Do you know which straw will produce the highest note? *A*
I think the longest one
Actually it's the other way round. Good guess but the shortest straw will give the highest note. The longest straw will give the lowest note. *G*
Low note.
That's right.

Reads Q 6d(i) with help with **different**
Do you know what's meant by a fair test? *U*
People need to be the same.
That's right.
The straw needs to be the right length.

Yes so why might it not be a fair test? *A*
They are from different families.
Ok, think about ... will everyone blow the same? *A*
No some might be low notes and some might be high notes.
That's right and do you think some people might blow harder than others? *A*
Yeah
If they don't all blow the same we can't tell what is causing the different notes so it is not a fair test. *G*

Reads Q d (ii) help with **even**
Um it wont be a fair test.
Yeah that's right and can you think why? *A*
Silence
This time the same person is blowing into all of the straws. *U*
Do you think she will always blow the same? *A*
No
If she doesn't blow the same each time we cant tell what causes the different notes. *G*

Fish Tank

Reads Q4a with help with **tropical, clown loach fish, algae, also, dark corner, tick**
I've got a gold fish.
I think stop growing.
Do you lets think about that. Why do you think sunlight will make it stop growing? *A*
Because the sun will stop them growing.
Mm What happens to plants when they get sunlight? *A*
It will dry them out.
Yes too much sunlight will dry them out but what if they get just the right amount of sunlight? *A*
Silence
If they had lots of water and just the right amount of sun? *A*
Grow.
Yeah so that's our answer well done.

Reads Q 4b with help with **algae, water snails, he, then, clown loach fish, chain**
Do you know what a food chain is?
No
You need to write down what is getting eaten by what. The arrow means 'is eaten by'.
Can you put them in the right boxes? *A*
Snails
Snails. Who's going to eat the snails? *A*
The fish.
So where do we need to put snails and where do we need to put the fish? *A*
I think that one.
Yeah.
What goes in there? What gets eaten by the snails? *A*
Fish.
Not quite, read it again.
Water snails eat algae – algae.

Yes you've worked it out. Fantastic.

Reads Q 4c with help with **Neil, uses, magnets, algae Magnet B, outside**
Chats about his fish tank at length.
Right why?
Because magnet B is inside and the other one is outside so it sticks together.
Do you know what it's called when magnets stick together? *W*
No.
The magnets attract. *G*
Well done. Perfect.

Reads Q 4 d(i)
Silence.
What temperature does it say the water is in Neil's tank? *U*
30 degrees.
30?
35.
35 and what does it need to be to keep the fish healthy? *U*
25.
Right so...
No.

Reads Q 4d(ii)
Because the fish will die
Yes why will they die? *A*
Because the temperature if it goes right up to the top they will be dead.
Yes so is the water too... *A*
Hot. He needs to tip some cold water in.
That's right. Well done.

Periwinkles

Reads Q 9a with help with **animals, seashore, describe, other**
What looks different? *U*
Periwinkle B is the big one and Periwinkle A is the little one.
Thats right but can you think of something that isn't to do with size? *A*
The other one's taller?
That's also to do with size isn't it? Anything else that looks different? *A*
One is white one is grey.
That's right the colour is different. Well done.

Reads Q 9b with help with **have, bodies, shell, describe, function.**
Do you know what a function is? *U*
No.
It's something the shell is used for.
Silence.
Why might it need a shell? *A*
Keep it protected from birds.
Absolutely perfect well done.

Reads Q 9c with help with **area, affects, measures, sample, different areas, seashore, instead**
Because they'll have different skin.
Why is measuring 20 better than measuring one? *A*
Because in case one's got a cold.
Do you think periwinkles get colds?
No.
Right so what can the scientist find out if he measures 20 that he can't find out if he only measures one? *A*
He needs to measure all the shells of different colours.
Yes OK and if he only measured one it might be an extra big one or an extra small one. *G*
Yup.

Reads Q 9d with help with **return, animals, same, collected**
Becasue they get allergic to something.
They might, what else might happen to the periwikles if they are put back in a different place? *A*
Get hurt.
Yes they might not be able to survive that's right. *G*

Reads Q 9e (very tired from reading now) with help with **level, during, leak**
Probably the animals did drink it.
Actually that's not what happened, something else has happened what do you think?
Sigh.
Where has some of the water gone? *A*
Well it has been sinking a little bit?
Yes well it does say none has leaked so there's somewhere else the water could have gone. *A*
Over.
It can go up is that what you mean ?
Into the Air!
Yes that's right! And do you know the scientific word for this? *W*
Make a rainbow?
Yes it might make a rainbow, but when water goes up into the air like that it's called evaporation. *G*

Ofqual wishes to make its publications widely accessible. Please contact us if you have any specific accessibility requirements.

First published by the Office of Qualifications and Examinations Regulation in 2010

Office of Qualifications and Examinations Regulation
Spring Place
Coventry Business Park
Herald Avenue
Coventry CV5 6UB

2nd Floor
Glendinning House
6 Murray Street
Belfast BT1 6DN

Telephone 0300 303 3344
Textphone 0300 303 3345
Helpline 0300 303 3346

www.ofqual.gov.uk