Science Education International
Vol.21, No.1, March 2010, 31-46

International Council of Associations for Science Education

Questioning Chemistry: The role of level, familiarity, language and taxonomy

Susan Rodrigues, Neil Taylor, Margaret Cameron, Lorraine Syme-Smith, Colette Fortuna
School of Education, University of Dundee, Scotland

Abstract
This paper reports on data collected via an audience response system, where a convenience sample of 300 adults aged 17-50 pressed a button to register their answers for twenty multiple choice questions. The responses were then discussed with the respondents at the time. The original dataset includes physics, biology and chemistry questions. The questions were derived from the Assessment of Achievement Programme, Sixth Survey of Science 2003 [AAP 2003] (Scottish Executive Education Dept, 2005). The findings presented in this paper are a subset of the data collected as the paper discusses the designs seen in terms of the style and nature of chemistry questions and the degree of perceived confidence in, and the answers provided by, 300 adult participant responses. The focus of this paper is not on misconceptions or participants' strong or weak understanding of science. Instead the paper analyses the demands made by the questions. The findings suggest that structuring multiple choice chemistry questions is complex. What may appear to be a simple question in terms of the curriculum prescribed levels may be rendered more complex by phrasing, familiarity and taxonomy. Furthermore, if international studies rely on multiple choice questions which favour testing recall of information, then it is possible that pedagogy that supports recall rather than comprehension or application may be endorsed.

Key words: *Familarity, language, nature of chemistry, multiple choice chemistry questions*

Introduction

Many of us are probably familiar with the programme "Who wants to be a millionaire?" The programme uses an audience response system to collect responses when the contestant uses the 'Ask the audience' help line. This paper reports on data collected by using a similar type commercially available audience response system. Twenty questions derived from the Assessment of Achievement Programme, Sixth Survey of Science 2003 [AAP 2003] (Scottish Executive Education Dept, 2005) were presented to a convenience

sample of 300 adults aged between 17-50 who were either undergraduate or postgraduate students. The responses from the participants were discussed with the participants at the time and the collected data was made available to a team of 5 researchers from the science, mathematics and information communication technology (ICT) area. This paper provides some insight into the responses seen when respondents answered Chemistry questions. The participants' answers and their perception of their confidence in the accuracy and appropriateness of their answer form the body of this paper.

In terms of organization, this paper has four main sections. Section one provides an introduction by reporting on recent and relevant literature. Section two describes the process used to collect the data. Section three documents findings that pertain to the chemistry questions. Section four identifies implications for those interested in assessing chemistry through the use of multiple choice questions.

Audience response system

Audience response systems are being used more frequently and integrated more effectively in learning environments (Freeman, Bell, Comerton-Forde, Pickering & Blayney, 2007), possibly to encourage participation (Kay and Knaack, 2009). While Alexander, Crescini, Juskewitch, Lachman & Pawlina, (2009) suggest staff and students found the devices easy to use, Kay and Knaack (2009) suggest that around 2% of their sample encountered difficulty in handling the devices. Concern was also raised by some, for example Simpson and Oliver (2007) about inexperienced staff using audience response systems. Freeman *et al* (2007) identify other disadvantages in terms of:

- Technical support to set up equipment
- Time to prepare questions, issue handsets and organise equipment
- Cultural change in teaching to incorporate an interactive style.

The positive benefits in using an audience response system outweigh the disadvantages mentioned. Kay and LeSage (2009a, 2009b) in their literature review identified benefits of using audience response systems and suggested advantages could be classified under three categories: classroom environment, learning and assessment. Gauci, Dantas, Williams and Kemm (2009) posit that strategies that engage all participants while receiving anonymous answers from them, rather than responses from the most vocal participants, is an advantage. The responses are also made confidentially. So this means that the less confident participant may not be inhibited by other responses or persuaded to follow a popular response. Kay and LeSage (2009) also showed how audience response systems were used to inform staff and students about how well concepts were understood. These systems may therefore be effective as formative assessment tools while generating a record of this assessment process.

Assessing Science

Science literacy is a commonly used, yet difficult to define, and sometimes controversial, term. Some suggest that the term science literacy covers: (a) understanding the nature of scientific knowledge and science; (b) understanding science doctrine; (c) understanding the relationship between technology and science; (d) understanding the impact of science on society; (e) applying science knowledge and reasoning to non academic life and (f)

being able to communicate science (Shwartz, Ben-Zvi and Hofstein, 2005). Given these various aspects, assessing whether someone is scientifically literate can be problematic. Nevertheless there is a body of research from the early 1980s suggesting that some of the science debates are informed and fuelled by misconceptions. This may be because children and adults have misconceptions about basic science ideas, or because the responses are a consequence of the language used in the questions. In formal academic spheres examinations and tests are often used to gauge competence in science, with teachers held accountable for pupils' success. Unfortunately, in many instances, this success is ascertained and measured using assessment instruments which are often at odds with evidence-based best practice (Torrance, 1995; Harrison, Hofstein, Eylon & Simon. 2008). In this paper we consider the data generated when participants used and explained their responses to items from an assessment instrument used to assess students' science.

Method

Three hundred adults aged between 17-50 years used an audience response system (Turning Point; Turning Technologies, 2010^{TM}) to respond to a bank of twenty science questions. There were 7 chemistry questions, 7 physics questions and 6 biology questions. The response systems allowed participants to independently answer questions that were presented via powerpoint. The respondents pressed a button on individual hand held devices to select and register their answers. The first five questions on the powerpoint asked the participants for basic information about themselves. For example they were asked their gender, and whether they had previously undertaken postgraduate studies. These questions were relatively straightforward and allowed participants to gain familiarity with the audience response system and the hand held devices, before responding to the science questions. It also allowed us to check that the handsets were registering the user's choice.

The science questions presented to the 300 participants stem from the Assessment of Achievement Programme, Sixth Survey of Science 2003 (Scottish Executive Education Department, 2005). The research team felt the questions would be valid and reliable as the questions were developed by National bodies (the Scottish Qualifications Authority and, Learning and Teaching Scotland) that are responsible for supporting assessment and learning in Scotland. The questions were also written to sample pupils aged from 7-14 years old. Therefore, for the purpose of our study the science was assumed to be within the range and experience of the sample of adults (17-50 years old). The audience response system was also used to collect information about their level of confidence in answering biology, chemistry and physics questions.

The audience response system collected and collated responses to 20 science questions. The national guidelines for 5-14 (Learning and Teaching Scotland, 2000) were used to categorise the questions, though we acknowledge that with the introduction of a Curriculum for Excellence in 2010 (Scottish Government, 2004) the questions could also be analysed from the new learning outcome statements. However, as the Scottish Survey of Achievement (SSA) 2003 would have relied on the national guidelines for 5-14, the research team thought it fair to map the SSA questions against the 5-14 guidelines (a full

set of the questions can be seen in appendix 1). Attainment targets are grouped as six levels of progression, with the following level descriptions:

- Level A: attainable in the course of P1-P3 (age 5-7) by almost all pupils.
- Level B: attainable by some pupils in P3 (age 7-8) or even earlier, but certainly by most in P4.
- Level C: attainable in the course of P4-P6 (age 8-10) by most pupils.
- Level D: attainable by some pupils in P5-P6 (age 9-10) or even earlier, but certainly by most in P7 (age11).
- Level E: attainable by some pupils in P7-S1, (age 11-12) but certainly by most in S2.
- Level F: attainable in part by some pupils, and completed by a few pupils, in the course of P7-S2 (11-13).

It should be noted that level E and F are higher than levels A and B. The 7 chemistry questions used with the adult sample range from level C to level F. In addition to reviewing the questions in terms of the 5-14 levels, the questions were also classified using a modified version of Bloom's (1956) taxonomy. The analysis of each question looked at what Bloom identified as the Cognitive element: mental skills (knowledge) and not the affective element. Bloom's taxonomy for cognition, identified six major categories: knowledge recall (recall data or information), comprehension (understand the meaning, translation, interpolation, and interpretation of instructions and problems), application (use a concept in a new situation), analysis (distinguish between facts and inferences), synthesis (combines parts, with emphasis on creating new meaning and/or structures) and evaluation (make judgments about the value of ideas or materials). Bloom's original version has been followed by several revisions. For example, Anderson and Krathwohl (2001) suggested; remembering, understanding, applying, analysing, evaluating, and creating. Anderson and Krathwohl (2001) see creating new knowledge within the domain as the most demanding, and there is a move from the use of nouns to the use of verbs in their framework. Bloom's taxonomy has been criticised on several fronts, (with criticisms including it being hierarchical and linear, rather than seeing the mind as a concept map or web), however, the gist of these taxonomies allow for questions to be considered from a related angle that did not simply look the level of difficulty ascribed by curricula. In light of the criticisms directed at Bloom's taxonomy the questions were considered in terms of three facets: recall (of facts, terms or basic concepts); comprehension (demonstrate understanding by requiring respondents to compare, translate or interpret) and; the three remaining aspects -application, analysis, synthesis were considered under one broad heading of apply (apply rules or facts, techniques in a different way; make inferences; combine elements in a new pattern). The questions were also subjectively classified in terms of context familiarity, which was not ascribed prior to asking the participants to respond. Instead context familiarity was ascribed as a consequence of participant discussion about each question.

Adults were invited to participate and their participation was taken as consent. However, it should be noted that by simply not pressing a button on the audience response system, participants could opt out. The software collected the responses anonymously hence maintained confidentiality. The software collated the responses in situ and the graphs

generated, which depicted response patterns for each question, were then shared and discussed with the participant cohort. Clarke (1998) devised a complementary accounts methodology and this methodology was used to analyse the data collected: The anonymous data was shared with members of a Science, Mathematics and ICT (SMICT) research group, all of whom use ICT in their teaching and research practice, and this generated a complement of accounts for the same data set. The data presented in this paper provides an opportunity to identify relationships seen in responses to the chemistry questions, and to provide an opportunity to reflect on participant rationale for answer selection.

Findings

Twenty of the participants had a degree in the sciences of which 4 were in the physical sciences and 16 in the biological sciences. It is worth noting that the participants who had already achieved a science degree did not necessarily perform well in this test. Within the sample in the group without a first degree, 28% had Higher Biology, 12% possessed Higher Chemistry and 13% had Higher Physics. ('Highers' are a final year school examination, which students can use to gain access to University.) From a general perspective, analysis of all questions suggests:

- Participant ability to answer questions correctly was not matched in confidence to answering questions
- The question language influenced the answers given.
- Participants thought they had to use the numbers that appeared in questions, in some way.
- The level of difficulty in the nature of the questions was a function of cognitive demand in terms of curriculum level, language, familiarity and taxonomy.

In the national AAP 2003 report (Scottish Executive Education Department, 2005) a 65% threshold was used to indicate pupils' 'secure knowledge' of the topic at the prescribed level of cognitive difficulty. In the adult sample, there were only 6 of the 20 questions answered with more than 80% of the adult sample choosing the correct option, and only one of these was a chemistry question. Only two of the chemistry questions obtained the 65% threshold, one barely (65.8% identified the correct response for question16) and one clearly (84.5% identified the correct response for question 13).

Chemistry questions

Table 1 presents the 7 chemistry questions (represented by the number indicating the position of this question in the full survey, so as to avoid confusion should readers review the full survey in appendix 1). The second column in Table 1 presents the number of respondents (as a percentage) who provided the scientifically correct response. The third column reflects the curriculum level ascribed to this question, the fourth column signals context familiarity (as determined by participant discussion rather than title) and the fifth column identifies the taxonomic level.

Table 1. Summary of Results for the Chemistry questions

Question No.	% Correct	Curriculum 5-14 Level	Context Familiar (F) - Unfamiliar (U)	Shrunken Taxonomic scale
6	27.4	Level D	Candle flame (F)	Comprehension
8	6.4	Level C	Boiling water (F)	Apply
13	84.5	Level C	States of matter (F)	Recall
16	65.8	Level E	Reactivity series (U)	Apply
18	30.8	Level F	Chemical reaction (U)	Apply
20	57	Level E	pH	Comprehension/Recall
23	34.9	Level F	Atomic mass	Apply

Nearly 85% of the participants responded correctly to question 13. (In contrast, a similar level was achieved for three of the six biology questions (question numbers, 9,14,17 in appendix 1) and two of the seven physics questions (question numbers 10,15 in appendix 1). Sadly less than ten percent of the adult sample responded correctly to question 8, while under 35% responded correctly to questions 6, 8, 18, and 23. Only one of the questions in the chemistry batch could be considered a 'recall' question. The others warranted responses based on comprehension and analysis, and this in turn was guided by the taxonomic level, degree of familiarity, language used and the curriculum level. The correct answer is provided in 'bold' in the list of responses to each question.

Question with a high correct response rate (over 80%)
Question 13 was answered correctly by nearly 85% of the sample.

Question 13: A substance has a fixed volume but its shape depends on the container. What is this statement referring to?

- Solid
- **Liquid**
- Gas

The topic is found at Level C (attainable in the course of primary 4 –primary 6 (age 8-10) by most pupils) in the 5-14 curriculum guidelines. In terms of the shrunken Bloom's taxonomy, it was classified as a 'recall' question and it was classified as a familiar question. Participants explained the high correct rate for the cohort in terms of it being an easy question because they had many real life examples to recall with regard to this particular aspect of states of matter. They recalled for example that they poured liquids from container to container during the course of their everyday life, and were therefore able to use this recall and familiarity to answer a question on a topic deemed to be manageable for most 8-10 year old pupils.

Questions with a reasonable response rate (Between 35% and 80%)
Nearly two thirds of the participants answered Question 16 and over half of the participants answered question 20 correctly. Both of these questions were at Level E and while question 16 required application (apply rules or facts, techniques in a different way; make inferences; combine elements in a new pattern), question 20 could be considered to

be either comprehension (demonstrate understanding by requiring respondents to compare, translate or interpret what they had learnt about acids, alkalis and indicator levels in school) or for those who indicated recalling an advertisement for soap, a recall question.

Nearly 66% of the sample provided the correct response for Question 16. This question was classified as a Level E question, and using the shrunken Bloom's taxonomy, it was classified as an application question because the participants would need to apply facts about the periodic table and groups, and combine these elements to draw inferences. They needed to apply what they knew about the periodic table groups and series.

> Question 16: The following metals can be divided into two groups:
> Gold, Copper, and Mercury
> Calcium, Mercury and Sodium
> **Calcium, Sodium and Potassium**
> Mercury, Sodium and Potassium

In order, to answer the question using an understanding of concept approach, there were 3 steps;

(i) know what is meant by a group in the periodic table
(ii) apply this knowledge to identify which examples are in which groups
(iii) identify those that do not fall into a group

However, some of the participants indicated that they used a different process of elimination. In the list of possible responses provided for question 16, only one option did not include Mercury. Some of the participants based their responses on this process of deduction. As our participant explanations were generated during discussions with the adult cohort we cannot provide a breakdown with regard to how many of the 66% who provided a correct response did so on the basis of deduction.

In addition, the two response options that include both Sodium and Potassium, probably lowered the correct response rate, as some identified Sodium and Potassium as belonging to the same column in the periodic table, and hence interpreted the question in a general sense with respect to the word 'group' which resulted in them selecting the 'Mercury, Sodium, Potassium' option. These participants therefore drew accurately on some prior knowledge (commonality between sodium and potassium) but failed to recall the significance of the term 'group' from a chemistry perspective. Hence their comprehension and application elements required by the question were sound, but their chemistry curriculum level was not, and they did not have everyday contexts to draw upon. As our participant explanations were generated during discussions with the adult cohort we cannot provide a breakdown with regard to how many of those who provided an incorrect response including mercury, sodium and potassium did so on the basis of interpreting 'group' in its everyday sense.

Question 20 is a level E question, (attainable by some pupils the final year of primary school or first year of secondary school, (age 11-12) but certainly by most in S2, which is

the second year in secondary school). It was labelled a familiar context as participants referred to drawing on the pH messages promoted in advertisements involving soap. It was also classified as a comprehension question as it required participants to demonstrate understanding to interpret what they knew about acids, alkalis and indicator levels. However, in light of their comments regarding soap, it could also be considered a 'recall' question, if they simply recalled the pH statements made in the advertisements.

Question 20: When a pupil tested a liquid with litmus paper she got a value of pH 6. Which category does the liquid fall into?
Weak Acid
Strong Acid
Weak Alkali
Strong Alkali

Questions with a low correct response rate (Below 35%)
There were 4 questions (6,8,18,23) where less than 35% of the respondents were able to provide the correct response. Two were at level C/D and two were at level E/F. All were either comprehension or application type questions.

Question 6, is a level D question, which is a middle level 5-14 national guidelines attainment, supposed to be attainable by those aged age 9-10 years. It is a comprehension question, requiring the respondents to compare, translate or interpret prior experiences, and it could be considered familiar in terms of respondents having experience in lighting candles. However, less than 30% of this adult cohort was able to identify the correct response and the majority struggled with this question.

Question 6.) When a candle burns...
The wick burns
The wax vapour burns
The wick and the wax vapour both burn
The wax melts to let the wick burn

Three of these responses are acceptable, though the comprehensiveness/fullness of the response varies, and from a science perspective the most comprehensive response is 'The wick and the wax vapour both burn'. In terms of language, the question did not ask participants for the most comprehensive answer. While this may appear pedantic, it may explain the spread of response which indicates that participants were identifying acceptable responses rather selecting the most comprehensive response.

Question 8 had a very low success rate. It was a Level C, (pupils aged 8-10 are expected to attain this level), it was a familiar everyday context (boiling water) and it was deemed to be an application question because the respondents had to apply facts and make inferences.

Question 8: When water boils you see …
Bubbles of heat in the water
Bubbles of oxygen and hydrogen

Bubbles of air
Bubbles of steam

The majority of participants (70.6%) responded with, "You see bubbles of oxygen and hydrogen" and 18.4% responded "Bubbles of air". Respondents were asked why they chose oxygen and hydrogen, and their replies were: it 'sounded' the most scientific answer. Most participants knew that water comprises of hydrogen and oxygen, which they also know are gases. Given this prior knowledge and their reported need to provide a 'scientific answer' they deduced a response based on their understanding and interpretation of aspects of familiar knowledge and combined these elements to create new meaning, which unfortunately was incorrect.

Less than a third of our sample answered Question 18 correctly.

Question 18: Which of the following is not an example of a chemical reaction?
A candle burning
A firework exploding
A vitamin C tablet effervescing
Salt dissolving in water
An iron bar rusting

This is a Level F question (according to the 5-14 guidelines, a few pupils aged 11-13 would be expected to address this level). In addition, to answer the question there were 3 steps;

- know what is meant by chemical reaction
- apply this knowledge to identify which examples are chemical reactions
- identify the example that is left.

Interestingly, the participants said the wording confused them. The inclusion of the word 'not' also had a similar negative impact on the number of correct answers in response to question 7.

Question 23 provides a table (see below) and asks the participants to deduce the atomic number for Sodium. This question is a level F question in that only some pupils aged between 11 and 13 years old would be expected to complete it. It was considered to be an 'apply' question as it required participants to know how to determine the atomic number, if given information about constituent particles. It was not a familiar context in that participant conversation did not identify any every day contexts when explaining how they arrived at their answers.

Element	Number of protons	Number of neutrons	Number of electrons
Carbon	6	6	6
Sodium	11	12	11
Chlorine	17	18	17

The following were the response options

- 11
- 12
- 22
- **23**
- 34

Just over a third provided a correct response. However it is worth noting that 20.7% of the participants selected '34'. In the discussion regarding this question, many indicated their sense of dismay and apprehension when a table of information appeared on screen. For some participants it was a case of adding all the numbers as they remembered that atoms contained all of these components and arriving at the incorrect response (34). Others reached the correct sum for they too believed they had to perform a numerical calculation and were fortunate enough to add only the required elements. Some of those who provided the correct response also admitted that it was a guess.

Discussion

Using the audience response system meant that the participants were able to respond to questions with a degree of anonymity. In addition they were unaware of other responses until the responses were collated and presented for discussion.

The discussion on the various questions and responses suggest that structuring multiple choice chemistry questions is complex and may not provide an accurate indication of whether participants understand the chemistry. The findings and follow up discussion with the participants also suggests that what may appear to be a simple question in terms of the curriculum prescribed levels may be rendered more complex by phrasing, familiarity and taxonomy. So for example, the 4 questions (6,8,18,23) included two questions at level C/D (middle curriculum level) and two questions at level E/F (high curriculum level), but all were either comprehension or application type questions, and less than 35% of the respondents were able to provide the correct response. A question that is considered less complex from a curriculum level perspective may still be challenging. Science often requires students to move between the macroscopic, microscopic and symbolic levels particular to science (Johnstone, 1982), and as Han and Roth (2005), suggest, there are challenges to face when making sense of microscopic, macroscopic and symbolic science. Not surprising then, that Question 8 had a very low success rate, despite the fact it is considered to be a curriculum Level C, (pupils aged 8-10 are expected to attain this level) and is a relatively familiar everyday context (boiling water). As it was an application question that required participants to apply facts and make inferences, moving from macroscopic to microscopic levels, it appeared to be more challenging, and participants drew erroneous conclusions in their attempts to involve scientific terms. Similarly, a question that appears to be complex from a curriculum level perspective may be rendered less challenging because it is making a simpler demand in terms of the language or taxonomy. There is evidence to suggest that perhaps some responses which may in the past have been considered misconceptions may not

necessarily be the case, and may instead be a consequence of the nature of the question. For example a question that includes numeracy in some form whether it is required or not (as in the case of a physics question in this study), leads participants to believe that they have to generate a response that involved some form of calculation. While the sample in this paper was an adult cohort, and international studies deployed in schools tend to involve school pupils, our findings suggest that the style of question influences response. We do not know what the balance is between the types of different multiple choice questions found within international studies, and accept that all the pupils participating in an international study will face the same question. However, we would suggest that if international studies rely on multiple choice questions which favour testing recall of information, then it is possible that pedagogy that supports recall rather than comprehension or application may be endorsed. The manner in which questions are posed may influence participant response and we would suggest that this has implications for studies that draw comparisons between nations. Furthermore, this may not be testing their understanding of the science.

References

Alexander, C., Crescini, W., Juskewitch, J., Lachman, N and Pawlina, W. (2009) Assessing the Integration of Audience Response System Technology in Teaching of Anatomical Sciences. *Anatomical Sciences Education* 2, 160 – 166

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

Bloom, B. S. (ed.) (1956) *Taxonomy of Educational Objectives, the classification of educational goals – Handbook I: Cognitive Domain* New York: McKay

Clarke, D.J. (1998) Studying the classroom negotiation of meaning: Complementary accounts methodology. Chapter 7 in A. Teppo (Ed.) *Qualitative research methods in mathematics education*. Monograph Number 9 of the *Journal for Research in Mathematics Education*. Reston, VA: NCTM, pp. 98-111 (ISBN 0-87353-459-X)

Freeman, M., Bell, A., Comerton-Forde, C., Pickering, J. and Blayney, P. (2007) Factors affecting educational innovation with in class electronic response systems. *Australasian Journal of Educational Technology*, 23(2), 149-170

Gauci, S., Dantas, A., Williams, D. and Kemm, R. (2009) Promoting student-centered active learning in lectures with a personal response system. *Advances in Physiological Education*, 33, 60-71

Han, J. and W-M. Roth. (2005) Chemical inscriptions in Korean textbooks: Semiotics of macro and microworld. *Science Education*, 90 (2), 173-201.

Harrison, C., Hofstein, A., Eylon, Bat-Sheva. & Simon, S, (2008) Evidence-based Professional Development of Science Teachers in Two Countries. *International Journal of Science Education,* 30(5), 577-591

Johnstone, A.H. (1982) Macroscopic- and microscopic- Chemistry. *School Science Review*, 64, 377-379

Kay, R. and Knaack, L. (2009) Exploring the Use of Audience Response Systems in Secondary School Science Classrooms. *Journal of Science Education and Technology*, 18, 382-392

Kay, R. and LeSage, A. (2009a) A Strategic Assessment of Audience Response Systems Used in Higher Education. *Australian Journal of Educational Technology*, 25(2), 235-249

Kay, R. and LeSage, A. (2009b) Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers and Education*, 53, 819-827

Learning and Teaching Scotland (LTScotland). (2000) *Environmental Studies: Society, Science and Technology. 5-14 National Guidelines*. Dundee, Scotland: Learning and Teaching Scotland.

Scottish Government. (2004) *A Curriculum for Excellence* Learning and Teaching

Scottish Executive Education Department. (2005) *Assessment of Achievement Programme|: Sixth Survey of Science 2003.* http://www.scotland.gov.uk/Publications/2005/03/20882/54916 (accessed 4/2/2010)

Shwartz, Y. Ben-Zvi, R. & Hofstein, A. (2005) The importance of involving high-school chemistry teachers in the process of defining the operational meaning of 'chemical literacy', *International Journal of Science Education*, 27(3), 323-344.

Simpson, V. and Oliver, M. (2007) Electronic voting systems for lectures then and now: A comparison of research and practice. *Australasian Journal of Educational Technology*, 23(2), 187-208

Torrance, H. (1995) *Evaluating Authentic Assessment: Problems and Possibilities in New Approaches to Assessment.* Buckingham: Open University Press

Turning Technologies. (2010) *TurningPoint* http://www.turningtechnologies.co.uk Accessed 03/03/2010

Appendix 1: The science questions used in the survey

6.) When a candle burns...

The wick burns
The wax vapour burns
The wick and the wax vapour both burn
The wax melts to let the wick burn

7.) Which of the following is not an Invertebrate?

Snail
Mussel
Spider
Worm
Starfish
Salamander

8.) When water boils you see …

Bubbles of heat in the water
Bubbles of oxygen and hydrogen
Bubbles of air
Bubbles of steam

9.) The part of the plant cell labelled A is...

Cytoplasm
Nucleus
Cell membrane
Cell wall
Chloroplast

10.) A pupil placed 3 kitchen tools in a bowl of hot water and left them for a few minutes. When he returned he noticed one felt much hotter than the others. Which do you think felt the hottest?

The wooden spoon
The plastic spatula
The metal ladle.

11.) Select the correct statement about the respiratory system.

The oxygen content of the air we breath out i...
Carbon dioxide in the lungs passes into the b...
Air sacs in the lungs are called alveoli
The oesophagus carries air to the lungs

12.) Astronauts appear weightless in space because …

There is no gravity in space.
The gravity of the Sun and Earth balance out...
The spacecraft is falling freely in orbit.
There is no air in space.

13.) A substance has a fixed volume but its shape depends on the container. What is this statement referring to?

Solid
Liquid
Gas

14.) A pupil noticed that birds feeding in the school playing field often chased away birds of the same species but hardly ever chased away birds of different species. Why do you think this happens?

Birds of the same species do not like each ot...
Birds of the same species compete for the sam...
Birds of the different species are always big...
Birds of the different species never compete ...

15.) The drawing shows an apple falling to the ground. In which of the three positions does the force of gravity act on the apple?

B only
A and B only
A and C only
A, B and C

16.) The following metals can be divided into two groups:

Gold, Copper, & Mercury
Calcium, Mercury and Sodium
Calcium, Sodium and Potassium
Mercury, Sodium and Potassium

17.) Which of the following is true?

The roots make food for the plant.
The roots of a plant make it green.
The roots of a plant take in water.
The roots of a plant make the seeds.

18.) Which of the following is not an example of a chemical reaction?

A candle burning
A firework exploding

A vitamin C tablet effervescing
Salt dissolving in water
An iron bar rusting

19.) In the circuit below all the bulbs are identical. When the switch is closed, which lamps will appear equally bright?

A, B and C
A and B only
B and C only
A and C only

20.) When a pupil tested a liquid with litmus paper she got a value of pH 6. Which category does the liquid fall into?

Weak acid
Strong acid
Weak alkali
Strong alkali

21.) Which of the following is a unit of force?

second
pascal
kilogram
newton
decibel

22.) Here are some statements about heat moving from one place to another by radiation. Select the correct response.

Heat rays are stopped by glass
A block of ice cream will radiate no energy.
Radiation can travel through places where the...
A black surface loses heat more slowly than ...

23.) The **following information is given in a data book.**

Element	Number of protons	Number of neutrons	Number of electrons
Carbon	6	6	6
Sodium	11	12	11
Chlorine	17	18	17

What is the atomic mass of sodium?

11
12
22
23
34

24.) Which of the following is a product of the process of photosynthesis?

Oxygen
Water
Compost
Sunlight
Pesticide
Carbon dioxide
Nitrogen

25.) The mass of an object on Earth is 4kg. Since the force due to gravity on Earth is 10 newtons per kilogram, the object has a weight of 40 newtons. On the moon the force due to gravity is about 1.6 newtons per kilogram. What will be the mass of the object on the moon?

4 kilograms
5.6 kilograms
6.4 kilograms
40 kilograms
64 kilograms