April 2006
ISSN 1109-4028

Volume 7, Issue No 2
Pages 49 – 159

Chemistry Education Research and Practice

Published quarterly by The Royal Society of Chemistry

ROYAL SOCIETY OF CHEMISTRY

Chemistry Education Research and Practice

April 2006 | Volume 7, Issue no 2
ISSN 1109-4028 | Pages 49-159

This issue is in celebration of the work of the Centre for Science Education, University of Glasgow, under the leadership of Professor Alex Johnstone.

Guest editors: Alex Johnstone and Norman Reid

Contents

Indexed/Abstracted in
CHEMICAL ABSTRACTS (CA)
EDUCATIONAL RESEARCH ABSTRACTS ONLINE (ERA)
http://www.tandf.co.uk/era

Chemistry Education Research and Practice

The journals, *University Chemistry Education*, published by The Royal Society of Chemistry, (http://www.rsc.org/uchemed/uchemed.htm) and *Chemistry Education Research and Practice*, published from the University of Ioannina, (http://www.uoi.gr/cerp/) have merged with effect from January 1st 2005. The new, fully electronic journal is published by The Royal Society of Chemistry under the title: ***Chemistry Education Research and Practice,*** and it will continue to be available free of charge on the Internet**.** There are four issues per year.

The new journal is edited by Georgios Tsaparlis (gtseper@cc.uoi.gr) and Stephen Breuer (s.breuer@lancaster.ac.uk) and intends to maintain the high standards set by its predecessors. Its editorial policy will be the following.

'*Chemistry Education Research and Practice*' is the journal for teachers, researchers and other practitioners in chemical education. It is the place to publish papers on:

- research, and reviews of research in chemical education;
- effective practice in the teaching of chemistry;
- in depth analyses of issues of direct relevance to chemical education

Contributions can take the form of full papers, preliminary communications, perspectives on methodological and other issues of research and/or practice, reviews, letters relating to articles published and other issues, and brief reports on new and original approaches to the teaching of a specific topic or concept.

The new journal welcomes contributions of the type described above; these should be sent to cerp@rsc.org.

Chemistry Education Research and Practice

Chemistry Education Research and Practice

Guidelines for Authors

Submission of contributions

Chemistry Education Research and Practice (CERP) is the journal for teachers, researchers and other practitioners in chemical education. It is published free of charge, electronically, by The Royal Society of Chemistry, four times a year. It is the place to publish papers on:

- research, and reviews of research in chemical education;
- effective practice in the teaching of chemistry;
- in depth analyses of issues of direct relevance to chemical education

Contributions can take the form of full papers, preliminary communications, perspectives on methodological and other issues of research and/or practice, reviews, letters relating to articles published and other issues, and brief reports on new and original approaches to the teaching of a specific topic or concept.

1. The original contribution should be submitted electronically, preferably in Word for Windows format. Any associated diagrams should be attached in JPG or GIF format, if possible. Submissions should be made by e-mail as a file attachment to cerp@rsc.org, or directly to the editors: Stephen Breuer at s.breuer@lancaster.ac.uk or to Georgios Tsaparlis (gtseper@cc.uoi.gr).

2. Submitted contributions are expected to fall into one of several categories (listed above). Authors are invited to suggest the category into which the work should best fit, but the editors reserve the right to assign it to a different category if that seems appropriate.

   A word count (excluding references, tables, legends etc) should be included at the end of the document.

3. Presentation should be uniform throughout the article.

   Text should be typed in 12pt Times New Roman (or similar), with 1"/ 2.5 cm margins, double-spaced, unjustified, ranged left and not hyphenated.

   Always use an appropriate mix of upper and lower case letters: do not type words in uppercase letters either in the text or in headings. **Bold** or *italic* text and not upper case letters should be used for emphasis.

   All nomenclature and units should comply with IUPAC conventions.

   Tables and figures should be numbered consecutively as they are referred to in the text (use a separate sequence of numbers for tables and for figures). Each should have an informative title and may have a legend.

Equations should be written into the text using the word processing program, either as normal text or using the program's equation facility.

Structures should, wherever possible, be treated as a figure and not incorporated into text.

References should be given by the name of the author (or the first author, if more than one), followed by the year of publication. If an author has more than one reference from the same year, then it should be given as Smith 2001a, Smith 2001b, etc.

Footnotes should be generally avoided and important additional information may be referenced and included in the reference list.

4. A title page must be provided, comprising:
   - an informative title;
   - authors' names and affiliation, full postal address and e-mail; (in the case of multi-authored papers, use an asterisk to indicate one author for correspondence, and superscript a, b, etc. to indicate the associated addresses);
   - an abstract of not more than 200 words;
   - keywords identifying the main topics covered in the paper

5. Wherever possible articles should be subsectioned with headings, subheadings and sub-sub-headings. Do **not** go lower than sub-sub-headings. Sections should not be numbered.

   The introduction should set the context for the work to be described; include references to previous related work, and outline the educational objectives.

   A concluding section (which need not be headed conclusion) will include an evaluation of the extent to which educational objectives have been met. A subjective evaluation may be acceptable.

6. The formatting of references should follow the following practice:

   Books and Special Publications:
   Author A., (year), *Title of the book italicized*, Publisher, Place of publication, page no. if applicable.

   Journal Articles:
   Author A., Author B. and Author C., (year), Title of the article in Roman type, *Full Name of the Journal Italicised*, **Volume no. in Bold**, inclusive page numbers.

   For example:
   Osborne R. and Freyberg P., (1985), *Learning in science: the implication of children's science*, Heinemann, London.

   Jackman L.E. and Moellenberg W., (1987), Evaluation of three instructional methods for teaching general chemistry, *Journal of Chemical Education*, **64**, 794-96.

7. All contributions submitted will be refereed anonymously by two independent referees. In case of a disagreement a third referee will be consulted. The decision of the Editors on

the acceptance of articles is final.

8. Authors grant *CERP* the exclusive right to publish articles. They undertake that their article is their original work, and does not infringe the copyright of any other person, or otherwise break any obligation to, or interfere with the rights of such a person, and that it contains nothing defamatory.

9. Articles will be published on the Web in PDF format.

Professor Alex H Johnstone

Alex was born in 1930 in the city of Edinburgh and was brought up in Leith, the seaport of Edinburgh. He was educated at Leith Academy and was fortunate enough to be taught by excellent teachers. When the time came to go to university, he had a ‘problem’ because he was qualified to study Medicine, Science, English language, History and Divinity. He chose Science and studied Chemistry, Physics, Maths and Botany, graduating in Chemistry from the University of Edinburgh. He also holds a Doctorate in Chemical Education from the University of Glasgow, a postgraduate Certificate in Education and a Diploma in Biblical Studies. He joined the Chemistry Department at the University of Glasgow, and his first Chair was in Chemistry when he was appointed for his excellent teaching in inorganic chemistry (with a particular interest in bio-inorganic systems) and also for his extensive research in Chemical Education. Later he was appointed to the Chair of Science Education at the same University, but he continued to teach chemistry at all levels till his retirement.

His deep commitment to teaching, supported by his interest in psychology and his chemical expertise, led him to establish the Centre for Science Education within the Faculty of Science at the University of Glasgow. This is a research centre, but not a teacher training centre. Here he has supervised the work of over eighty researchers and has had, as co-supervisors, colleagues from most departments within the Faculty of Science, such as chemistry, physics, biology, maths, geography, statistics and psychology. In this way it has been possible to work on a broad, but co-ordinated front, to tackle the large research questions in Science Education, which are more fully described in the overview paper in this issue of CERP. Each research problem has been designed to be part of a whole research programme and each student has contributed to the solution of large problems. More than two hundred publications have appeared as a result of this work.

The Centre’s work was recognised when the Principal of the University appointed Alex to set up a Teaching and Learning Service within the University of Glasgow to help teachers in every Faculty and Department within the university to develop their teaching skills and to offer support for innovation in teaching and learning.

Recognition for his outstanding work has not been confined to Glasgow or to Scotland. Amongst the honours awarded to him were:

- The Nyholm Medal of the Royal Society of Chemistry
- The Mellor Medal of the Royal Australian Chemical Institute
- The Illuminati Gold Medal of the Italian Chemical Society
- The Brasted Medal of the American Chemical Society
- The Verhagen Titular Chair of the University of Limburg, Belgium
- The FECS Lecture of the Federation of European Chemical Societies
- Presidency of the Education Division of the Royal Society of Chemistry

But Alex is not a one sided person. He has enthusiastic interests and expertise in local history, genealogy and archaeology, the last of these acquired in extra-mural classes in the University of Edinburgh and supported by having a son who is a graduate in that discipline. His interest in language ranges from etymology and the origins of place names in English to a working knowledge of German and French. His knowledge of history, botany and geology allow him to lead walking groups to explore the countryside and the hills around Stirling where he lives, leaving them tired but much better informed than when they had set out. Alex has, regretfully, never learned to play a musical instrument, but he has been passionately interested in choral singing, having been a member of several choirs. He is a lay preacher in the United Free Church of Scotland where he has scope to use his earlier training in Biblical Studies and to follow one of the lines, which he had at the end of secondary school.

Now that he has ‘officially retired’, he is still involved on the editorial boards of several journals and actively trying, through the Royal Society of Chemistry, to influence the establishment of a chemistry curriculum based upon research in Science Education, which will take account of what is now known about learning, about the nature of chemistry and the interests and enthusiasms of young people. In this way it is hoped to restore chemistry to a place of importance and popularity within our schools and universities.

In all of this business, his wife Martha, who has encouraged, advised, restrained and supported him throughout, has preserved Alex’s sanity, productivity and sense of humour. Together they are enjoying grandchildren, travelling and pursuing many interests, which time now permits.

Chemical education research in Glasgow in perspective

Alex H. Johnstone

Centre for Science Education, University of Glasgow, Glasgow, G12 8QQ, UK
e-mail: alexjo@btinternet.com

Received 1 November 2005, accepted 5 January 2006

Abstract: This paper describes the work of a large science education group (80+ workers) which, from 1969, has been tackling teaching and learning problems over a broad front. For much of the time, the group has worked within a Faculty of Science and has tried to take a scientific approach to the research. This approach is still followed although the Centre is now in a Faculty of Education. At the start, time was spent in gathering facts, looking for common factors, raising and testing hypotheses, generating working models and applying findings to real teaching and learning situations. This paper seeks to present an overview of the work up to about 1997, with illustrations from later work. Although the research applies to all science subjects, the emphasis here is on chemistry. The other papers in this issue exemplify the ongoing research which has arisen from this basic ground-laying and which has spread worldwide. [*Chem. Educ. Res. Pract.*, 2006, **7** (2), 49-63]

Keywords: science education, integrated research, working memory model, language, laboratory, problem solving, multi-level nature of chemistry, secondary and tertiary levels.

Introduction

This is not a formal research paper, but rather a background presentation against which each of the other papers in this issue can be set and understood.

Research in Science Education began in the University of Glasgow in 1969 and was housed in the Faculty of Science for the next thirty years. The centre was staffed by practising scientists and science educators who were interested in solving research problems in the teaching and learning of the sciences at all levels: from early secondary school to post-graduate university.

In the early 1960s, in common with many other countries, Scotland adopted new curricula in chemistry, physics and biology for secondary schools (ages 12-18) and the author was heavily involved in the design of the chemistry curriculum.

Looking for difficulty

It was decided to begin research with the chemistry curriculum by questioning students who had undergone the new curriculum. As they arrived at the Universities of Glasgow and Strathclyde to begin their studies in chemistry, each student (N = 1000) was given a list of all the topics and sub-topics in the chemistry curriculum and asked to categorise each of them into one of four groups:

a) *“I understood this easily”*
b) *“I had some difficulty but I now understand it”*
c) *“I have never understood this and will need to be taught it again”*
d) *“I have never been taught this”*.

The frequency of the choice of (c) was plotted against the topic list for each of the two university samples and the results gave almost perfect agreement.

The experiment was then repeated with 17-year old pupils in secondary schools and once more the frequency pattern was the same as that derived from the two universities.

The troublesome topics which they selected were:

i) Writing formulae and equations, and doing calculations from them (Howe,1971)
ii) Volumetric work involving molarities (Duncan, 1973)
iii) Ion-electron equations (Garforth, 1976)
iv) Avogadro's Number and the mole (Duncan, 1973)
v) Heats of reaction, Hess's Law and thermochemistry
vi) Redox reactions and E^o values
vii) Equilibrium (MacDonald and Webb, 1977)
viii) Organic formulae (various forms), (Kellet, 1980)

All these topics had (c) frequencies in excess of 50% of all respondents. Clearly this was a very serious situation requiring further research.

Looking for causes

The size of the research group was expanded so that each of the above topics could be analysed by a researcher looking for possible sources of the difficulties and for common factors among them. The reference after each topic above gives a link to the published material arising out of this phase of the work.

Results

Howe (Howe, 1971) working on formulae and equations at school level, tried to work backwards through the operations underlying the writing of 'simple' inorganic formulae looking for weaknesses. He asked pupils for:

i) the elements in a given compound,
ii) the symbols for these elements,
iii) the charges on the ions of these elements,
iv) the formula of the compound.

In more than 50% of the responses for familiar binary compounds, pupils made mistakes in steps (i), (ii) and (iii) and yet got the formula correct! However, when he asked about less familiar, but analogous compounds, e.g. lithium bromide instead of sodium chloride the errors occurred in all four stages. Interview revealed that pupils did not try to construct formulae from first principles, but simply memorised the ones most commonly used. Similar results were obtained for writing balanced equations: memory, not reason.

Duncan (Duncan, 1973) working with equations and the mole found that this multi-concept exercise defeated the majority of pupils. On reflection, he realised that a 'simple' problem such as, "*What mass of calcium carbonate will exactly neutralise 100 ml of 0.2 M hydrochloric acid*", was far from simple. The following operations had to take place:

a) A pupil first had to recall the nature of the reaction
   acid + carbonate = salt + carbon dioxide + water.
b) Formulae for each compound had to be recalled or worked out.
c) The equation had to be balanced to establish the 1:2 ratio between $CaCO_3$ and HCl.
d) The number of moles of HCl in 100 ml of 0.2 M had to be obtained.
e) The required number of moles of $CaCO_3$ would be half of that.
f) A fraction of the gram molecular weight had to be calculated.

Many of the problems found in text books and examination papers yielded an even more complex analysis.

Ion electron equations were another area of difficulty which made formula and equation writing far more complex than situations exposed by Howe and Duncan. Such equations required decisions about, symbols, ions (simple and compound), charges on the ions, spectator ions, state symbols and balancing.

Thermochemistry and equilibrium considerations yielded analyses of conceptual complexity in line with the problems exposed above. Not only did they need the underlying support of accurate equation construction, but they then added additional layers of thermochemical and algebraic manipulation.

Looking for common factors

The common factor of 'information complexity' was clearly emerging, but we needed some theoretical framework to link it all together. This was provided by the work of Kellett (Kellet, 1980) on organic formulae and equations. She wanted to 'see' organic formulae through the eyes of students. She prepared pictures of organic formulae written in various forms: expanded structural formulae, condensed formulae, formulae in different orientations, formulae of varying complexity.

Each formula was projected on a screen for a short period and students were then asked to draw it from memory. Some examples produced a 100% correct response for all students while the *same* formulae, in a different format, were often very poorly recalled. Ethanol written as C_2H_5OH was successfully recalled while the extended structural formula was badly recalled. Ethanoic acid written as CH_3COOH was easy, but its extended structural formula was a disaster. Putting together CH_3OH and $HOOCCH_3$ to give the ester was difficult, but the extended structural form was impossible for most.

Examination of the student scripts showed that the vast majority read the formulae from left to right and wrote them that way. Most errors took place at the right-hand side of each attempt. Interviewed about strategy, students explained that they memorised every chemical symbol and every bond as a symbol. The C_2H_5OH formula had six symbols, while the extended structural formula had 14 symbols.

Looking for a model

This immediately suggested a link with the work of Miller (Miller, 1956), who had done a similar exercise with digits. Subjects could recall 7+/ -2 digits in his Digit Span Tests, and this led him to his idea of a limited Short Term Memory Span. This made us re-examine the work of Howe, Duncan, Garforth, McDonald and Kellett to see if we had a common factor and if we now had a working hypothesis.

We had begun in 1969 with no fixed theoretical stance other than our awareness of the work of Ingle and Shayer (Ingle and Shayer, 1971) who were critical of the Nuffield Chemistry Syllabus in England. Their criticisms were based on the developmental stages proposed by Piaget. They had analysed the syllabus and classified each topic in terms of the Concrete and Formal Operational stages and declared that some topics were unsuitable for pupils around age 16 and proposed that these should be postponed until the pupils had reached the Formal Operational stage of development. Since we were dealing mainly with students at university level, they should have been well into that level and yet the problems persisted. We therefore sought another model on which to base our thinking.

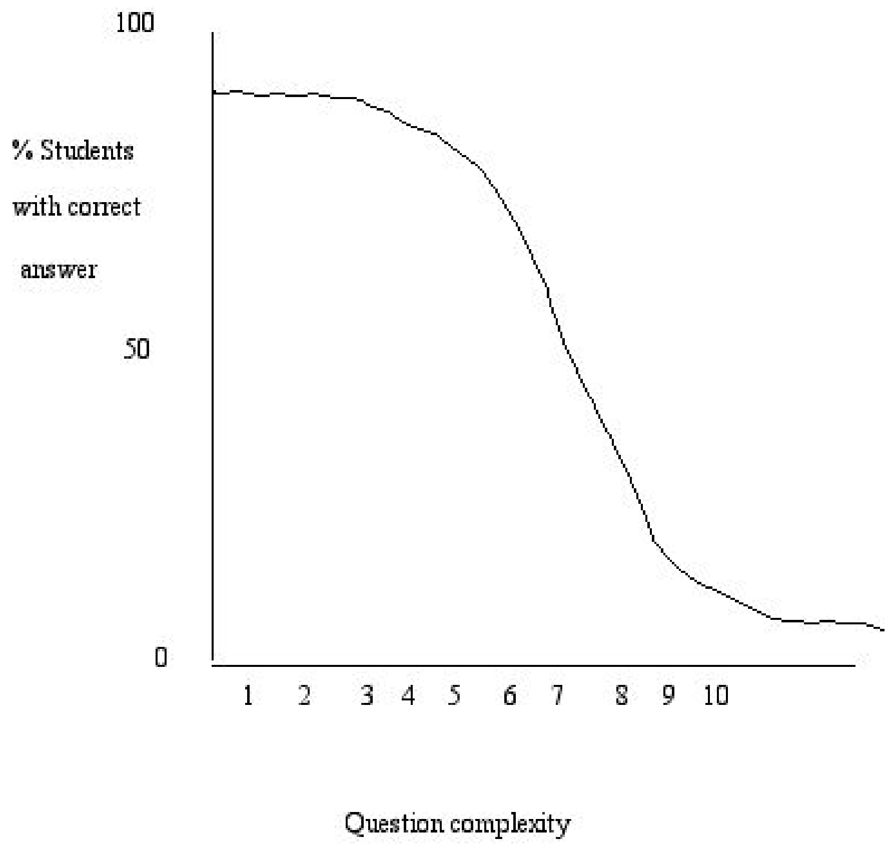
The work of Pascual-Leone (Pascual-Leone, 1970) and his group to seek a rational explanation for the Piagetian Stages, led to a neo-Piagetian treatment in terms of a limited, but growing, Working Memory in which increasingly complex operations could take place. This was not the same as Miller's Short Term Memory, which dealt with memorisation and recall (in exactly the same form). The Working Memory Space was a 'place' where information was temporarily stored and reworked before a response was made. A very full treatment of research into Working Memory can be found in the work of Baddeley (Baddeley, 1999).

To measure Working memory Space, Miller's Digit Span Test was modified to give a Digit Span Backwards Test (DSBT). A subject was given a series of digits and asked to give them back in *reverse* order. This involved a holding operation and a reversing operation. The number of digits retained correctly in this test was 5+/-2. The thinking operation was sharing the Working Memory Space with the holding of the information. This idea was closer to our situation in which students had been given, or had to recall, information and then convert it into another form to use in solving some problem.

El-Banna (El-Banna, 1986) tried to put our chemistry work alongside measures of Working Memory: Digit Span Backwards Test (DSBT) and Pascual-Leone's Figural Intersection Test (FIT). Five hundred students had their Working Memory Span measured by DSBT and sat a chemistry test in which the questions were of increasing complexity. The actual complexity was agreed by a panel of researchers, counting the pieces of information given in the question, plus the information to be recalled plus the operations required to produce an answer.

The Facility Values, (percentage of students getting a correct answer) for each item on the *y*-axis were plotted against the agreed complexity on the *x*-axis. We expected that the Facility Value (FV) would fall steadily as the complexity increased. It did fall, but in a surprising way (Figure 1)

Figure 1 Student success related to question complexity



All items with a complexity of 5 or less obtained a high Facility Value, but, as complexity increased, the FV plunged sharply, but not necessarily to zero. This is the kind of curve obtained in normal science when something is tested to a limit, e.g. tensile strength. It seemed that performance was good, provided the test stayed within Working Memory Capacity, but fell dramatically thereafter. El-Banna then replotted this curve, but this time he divided his sample into those with DSTB scores of 5, 6 and 7.

In the first curve, he found that the drop occurred when complexity exceeded 5; in the second when the complexity exceeded 6 and in the third when the complexity exceeded 7. At last we had our hands on something substantial; a model to direct our further research. However, this model raised some problems.

a) Although the fall in the graphs was steep, it was not vertical.
b) It did not drop to zero.
c) Miller's work suggested that the Short Term Memory could undergo an *instant* overload, but the questions in the chemistry test could be solved by a series of small sub-steps each of which would not cause an overload.

Further thought and experiment, and the consideration of Miller's ideas of chunking, helped to make sense of this. Depending upon how students had been taught previously, a question of complexity 7 could be reduced to 6 or less by some 'tricks' or 'shortcuts' and so these students were able to solve problems beyond their nominal Working Memory Capacity.

Those who were performing on the down slope were coping with problems slightly beyond their capacity and those who were on the low line parallel to the *x*-axis were a very small minority who were chunking problems well beyond their measured capacity. This might explain objections (a) and (b), but what about (c)? A question will be at its highest complexity when a student first reads it. If he is unable to apply some demand-reducing strategy to break the problem and organise the sub-problems, the problem will exist for that student at its maximum demand. Experienced teachers will have witnessed this in class when a problem is presented. Nobody knows where to start, but if the teacher indicates what to do first and then next, the problem ceases to be a problem. If students are taught algorithms they can effectively reduce the demand of a question, but, without prompts of this kind, problems beyond the Working Memory Capacity of the student are apparently insoluble.

Writing formulae and balancing equations could well be intractable problems for pupils at school because Working Memory (WM) Space is age dependent. Pascual-Leone suggested that Working Memory Space increases by one unit for every two years reaching a maximum by about age 16.

Ion-electron equations would also exceed Working Memory and so would mole calculations. Electro and thermo-chemistry would likewise be liable to overload late school and early university students. Patterns were emerging which led to further developments.

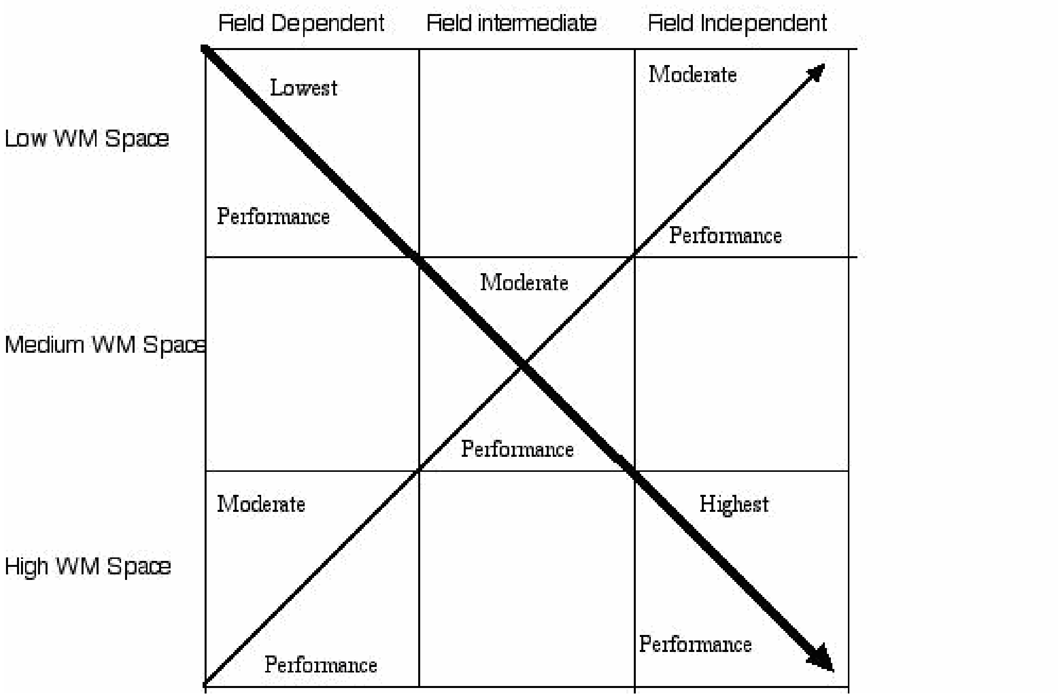
Further work by El-Banna and later by Al-Naeme (Al-Naeme, 1991) opened up other factors which affect students' effective use of their Working Memory.

Refining a model

Witkin's work (Witkin, 1977) on Field Dependence led us to examine its effect on the efficient use of potential Working Memory Space as measured by Digit Span Backwards Tests (DSBT). It was observed by El-Banna that a minor but significant proportion of students failed to solve questions the demand of which was within their DSBT measures. These were examples which did not seem to support our hypothesis that, "*Students should be able to solve problems if their demand did not exceed the measured Working Memory Space*".

However, when Al-Naeme (Al-Naeme, 1991) applied Witkin's Test of Field Dependence and the DSB Test to students and then analysed their scores on conventional chemistry tests, a very interesting trend appeared (Figure 2).

Figure 2. Effect on Chemistry performance of Working Space and Field Dependence



The underlying idea of Field Dependence (FD) and Field Independence (FI) is that some students, while learning, are easily affected by the 'field' against which the learning is done. They are easily distracted by irrelevant material and have difficulty in discriminating between the 'signal' and the 'noise'. At the other extreme, there are students who are Field Independent and can focus sharply on the relevant and ignore the irrelevant 'noise'. There is a continuum between these extremes. The scores in the FD/FI test were categorised as follows: Students who came within plus or minus one half of a Standard Deviation (SD) on either side of the mean score were classified as Field Intermediate (FInt). Those attaining less than the mean minus one half SD were classified as FD and those who attained more than the mean plus one half SD were classified as FI. This division gave us roughly equal numbers of students in each cell.

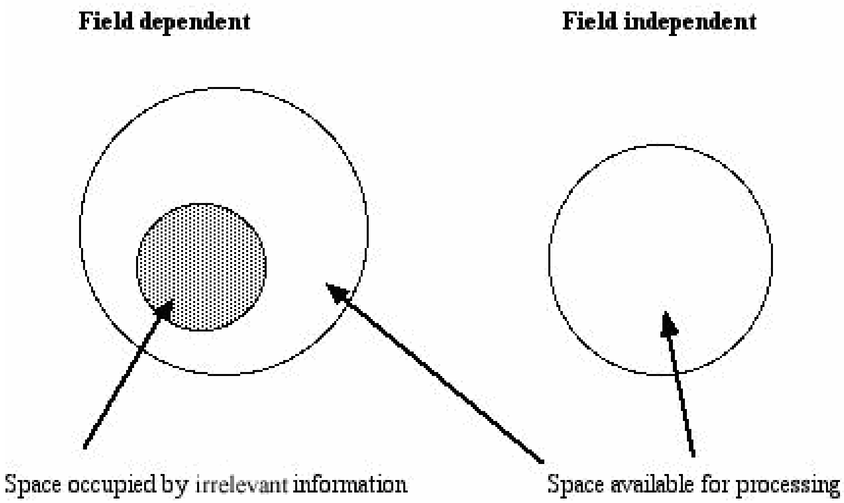
Looking at the results in Figure 2, it is possible to see the interaction of Working Memory Space and Field Dependence/Independence in their influence on scores obtained on a chemistry test. The vertical cells of Table 1 show improvement in chemistry scores as Working Memory Space increases. Also, the horizontal cells show improvement in chemistry scores as FD gives way to FI.

The strongest improvement trend is that from Low Working Memory Space coupled with FD (top left-hand corner) to High Working Memory Space linked with FI (bottom right- hand corner). Another striking trend appears along the other diagonal. The chemistry scores for High Working Memory Space linked with Field Dependence are almost the same as those for Low Working Memory Space coupled with Field Independence.

These patterns were replicated in ten different studies in chemistry, physics and biology and at different educational levels in schools and universities. How can this be interpreted? We saw from El-Banna the relationship between Working Memory Space and complexity of question, but it was not a perfect fit. Al-Naeme then showed that students, although having high Working Memory Space, could perform less well than expected when they were Field Dependent. In fact, they performed almost exactly as well as students who had low Working Memory Space, but who were Field Independent (Figure 3).

High WMS and FD = Low WMS and FI

Figure 3. Effect of Field Dependence upon available Working Memory Space

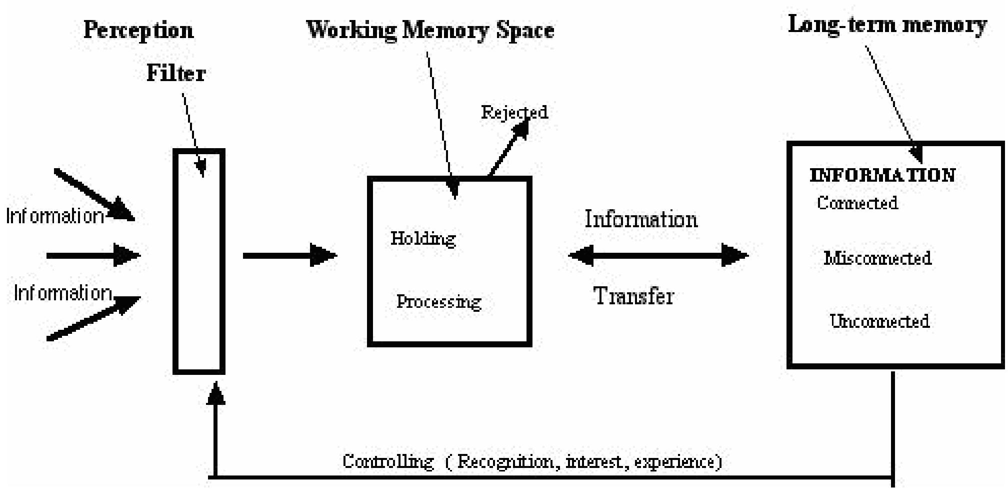


The potential usable Working Memory Space was effectively reduced by the space taken up by the irrelevancies introduced by Field Dependence. The low Working Memory Space students, who were Field Independent, had all the potential space available for use. These two models, El-Banna's and Witkin's, come together to rationalise the experimental results. Other factors that might moderate the available Working Memory Space were investigated, but FD/FI had by far the strongest effect.

We now had a working model to help us to interpret earlier results, to plan further studies and to look for solutions to the areas of difficulty isolated at the beginning of the 1970's. However, any model of learning which did not include the ideas of (a) perception (the initial interaction between the senses and the new information) and (b) of long term storage and retrieval, would be sadly incomplete. Our empirical model fitted well into existing models of Information Processing and we eventually adapted and adopted, as a working basis for our research, the model shown in Figure 4.

This indicates that external stimuli, such as those presented in teaching and learning experiences, are perceived by our senses and filtered. The learner attends to what is familiar, stimulating, interesting, surprising or exciting. To do this, the filter will be controlled, to a large extent, by what is already held in Long Term Memory. Something cannot be familiar, interesting or surprising unless it is being compared with some previous experience or expectation. What is held in the Long Term Memory store is crucial for this perception stage. This fits with Gestalt Theory (Koffka, 1933) and with Ausubel's (Ausubel, 1978) ideas that the most important factor in learning is what you already know.

Figure 4. Information Processing Model



The information admitted through the filter enters the conscious processing part of the mind, the Working Memory Space. The space has two functions:

a) to hold the incoming information in temporary store and
b) to operate upon the information to make 'sense' of it and prepare it for some response and/or to store it in Long Term Memory.

The sense-making operation will require information to be recalled from Long Term Memory to interact with the new incoming information. If the synthesis of 'sense' is achieved, it can be stored in Long Term Memory attached to (or filed along with) existing knowledge and understanding. Sometimes 'sense' is achieved by faulty attachment and this is very difficult to undo, because the learner has seen it as sensible and even satisfying. An Alternative Framework has been born. We shall look at the origins and consequences of this later. If a correct association is made, the learned material is more likely to be accessible and usable.

However, a third possibility exists for the fate of the processed material in the Working Memory Space. No 'sense' is made of it because no links can be found in the Long Term Memory, and yet the learner feels that it is important and must be stored. Such information enters Long Term Memory as rote learning, unattached and sometimes unlabelled. In this state, such information is often difficult to recall.

Before we leave the model and show how it informed our research, we must return to the Working Memory Space. Two functions of Working Memory Span were mentioned above: temporary holding and processing of information, but the Working Memory Span has a finite capacity. The consequence is that, if there is too much information to hold, there is not enough space for processing and the system overloads and seizes up. Similarly, if complex processing is needed, little information can be held for processing. This may also lead to overload and unsuccessful processing and storage.

In the Science Education literature there have been a very large number of studies reported under the general heading of Alternative Frameworks, Children's Science or Conceptual Misunderstandings, in which researchers have analysed the wrong ideas constructed by learners. These have generally been reported with little or no scientific explanation for the occurrence of such misunderstandings, and little indication of how the problems should be overcome. Our model indicates that the learner's frameworks occur as the products of the pupil's efforts to 'make sense' of the incoming taught information, but, in so doing, forming mis-linkages and storing them in Long Term Memory. The remedy suggested by our model is to help the learners to have the right attachment points in Long Term Memory alerted *before* teaching the new material so that correct interlinking will have a better chance of taking place in Working Memory Space before storage in Long Term Memory. This is just another way of looking at Ausubel's Advance Organisers which prime the Long Term Memory. In the light of this, Sirhan et al. (Sirhan, 1999) devised pre-lecture and pre-laboratory work to enable undergraduates to have Long Term Memory activation for new learning. Sirhan and Reid (2001) took this further and showed the overall picture of the development.

The Constructivist Movement has grown out of the Alternative Frameworks literature in an effort to help learners to construct their learning by correct preparation and shaping before storage. Since we, as researchers, are all looking at the same phenomenon, 'human learning', it should not be surprising that a comprehensive model of Information Processing embraces perception, processing and storage and in so doing provides a mechanism and a rationale for the research of Alternative Framework and Constructivist protagonists: to harmonise their work and to provide insight for their future research. The model suggests how to present material to avoid overload, to optimise processing and sense-making and to facilitate storage and recall. It also has a developmental, age-related basis which harmonises well with Piaget's work. What the model does not attempt or profess to deal with is 'attitude' and 'motivation' although it would suggest that constant overload which prevents 'sense-making' is a sure recipe for frustrating students and driving them along a series of steps: "*I don't understand*"; "*I can't understand*"; "*I shall never understand*"; "*I do not want to understand*". Many of the students voting with their feet and leaving the sciences may well have had this experience.

Using the model

Five important lines of research have opened up as a consequence of, and informed by, the Information Processing Model:

a) the function of language in science teaching and learning,
b) the problems of learning in a laboratory,
c) multi-level learning,
d) the assessment of science learning,
e) problem solving.

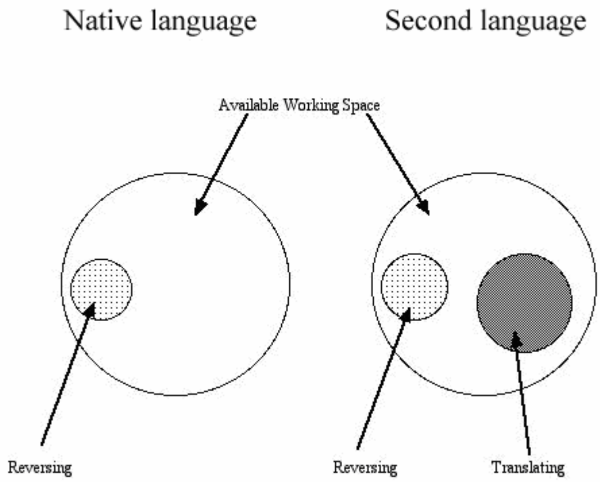
Language

The language study was conducted mainly by Cassels (Cassels,1985) and, 20 years later, by Selepeng (Selepeng, 2001). Cassels reasoned that the language held in Long Term Memory would affect the filter and the Working Memory processing, and he set out to find the vocabulary which might cause misunderstandings and has the potential for the construction of Alternative Frameworks. He eventually isolated more than 100 words, commonly used in school science, which caused trouble. These were words that teachers could easily assume that the pupils had a grasp of their scientific meaning. A word such as 'volatile' could be interpreted by pupils to mean 'unstable', 'explosive' or 'easily vaporised', based upon their common, everyday experience. All the meanings could make sense of a piece of chemistry, but two would have the potential for the construction of Alternative Frameworks. Another word is 'equilibrium' which carries with it a cluster of ideas from physics and from everyday experience, all of which contain the seeds of problems leading to Alternative Frameworks. 'Equilibrium' suggests balance and static state, an idea which is unhelpful in chemistry.

Cassels also found that the problem was even greater for pupils whose first language was not English. This had important consequences for those teaching ethnic minority groups, but even more so for pupils in ex-colonial countries where English was used as the instructional medium for science teaching despite the fact that the native language and culture was something very different.

Selepeng followed this problem further and measured the effective Working Memory Space for pupils when the Digit Span and Digit Span Backwards Tests were applied in the native language *and* in the second language. The effective Working Space was, on average, 1.6 units (20%) less in a second language than in the native language. In other words, pupils were handicapped in their science learning by the reduction of their Working Memory Space in a second language. The processing of the second language was taking up some of the valuable processing space needed for the understanding of the science (Figure 5).

Figure 5. Loss of Working Space when using a second language



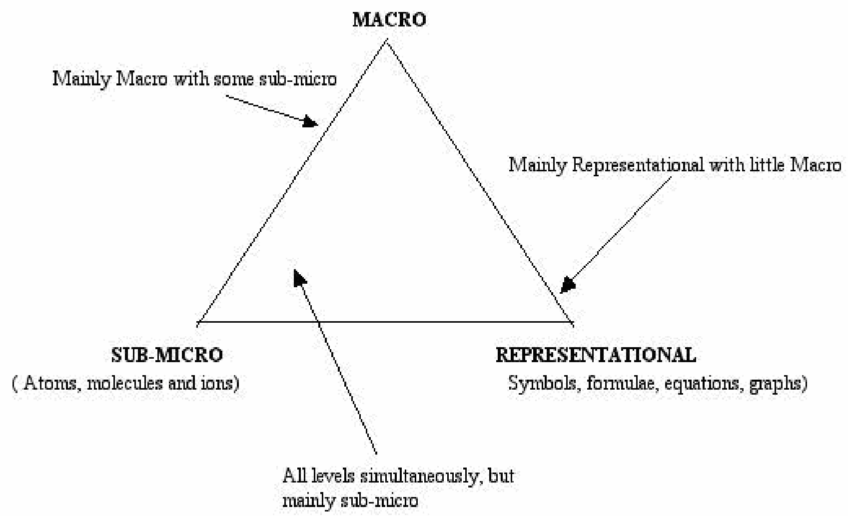
Laboratory work

Work by Wham (Wham, 1982), Letton (Letton, 1991) and Sleet and Vianna (Sleet, 1994) showed that there was little cognitive gain achieved in formal laboratory work at university level. Students gained hand skills and techniques, but lacked the connections to underlying theory. It became clear to those workers, that the structure of a conventional laboratory session had a gross potential for Working Memory Space overload. Written and verbal instructions, unfamiliar equipment and chemicals, observing and recording; all these together occupied Working Memory Space leaving no room for cognitive processing. Students, in an effort to reduce the discomfort of the overload, used the written instructions as a 'mind-in-neutral' recipe. This behaviour is often deplored by teachers, but the Information Processing Model indicates that the fault lies with the teacher for creating situations of gross overload. Several workers, such as Letton (Letton, 1991), Zaman (Zaman, 1998) Al-Shuaili (Al-Shuaili, 2001), began to design laboratory experiences with pre-lab preparation. This involved the students in thinking through the purpose of a laboratory, in planning some part of the experiment and, in so doing, activating the Long Term Memory in readiness to 'make sense' of the laboratory which followed. With the LTM thus primed, the students were in a position to distinguish 'noise' from 'signal' in the laboratory and to disregard the 'noise' as irrelevant and release space for thinking about the meaning of what they were doing. There was also an added bonus in an improvement in student attitude to laboratory work. How this developed, is reported in a recent paper in this journal (Al-Shuaili , 2001).

Multi-level learning

It was recognised that overload could occur at the early stages of learning chemistry because of the very nature of the subject (Johnstone, 1982 and 1991). Pre-1960, the atomic and molecular aspects of chemistry were not as evident at the early stages of learning the subject as they are now. Atomic and molecular structure and the nature of bonding tended to be kept until the later years in school. However, we now have a situation where the particulate nature of matter (the atomic and molecular development of this) and the introduction of ions and bonding are found early in introductory chemistry. Pupils are now confronted with the simultaneous introduction of unfamiliar substances on the bench, a description of them in molecular terms and a representation of them by symbols and formulae (Figure 6).

Figure 6.The three conceptual levels of chemistry



This figure is semi-quantitative in that any corner of the triangle indicates 100% treatment of the subject in that medium. For example, a totally macro approach would be represented by the macro corner. However, when the teacher introduces an experience on the bench interpreted by an equation, the treatment is somewhere along the right side of the triangle depending upon the emphasis given. Similar considerations would apply to the other two sides.

In many lessons, there is a blend of all three experiences simultaneously, represented by a point within the triangle, its position being determined by the relative proportion of the three components. Inside the triangle lies the potential for gross overload of Working Memory Space.

Teachers, and other chemists, flit around and inside the triangle with ease, giving us a powerful way of thinking about our discipline, but can early learners follow us inside the triangle without the onset of overload or with 'rationalisations' which lead to Alternative Frameworks? We might have to rethink our curricula to begin with a treatment of one corner only followed by the use of a side, before we lead the students into the middle of the triangle.

This idea is now appearing widely in the literature, often quoted, but little of the consequences followed up. A notable exception can be seen in the work of Tasker (Tasker, 2002) who has devoted much study and ingenuity to tackling this problem.

This triangular aspect of the nature of chemistry also gets in the way of laboratory learning and partially explains the lack of conceptual development in the laboratory.

Assessment of learning

We have almost come full cycle to El-Banna's work (El-Banna, 1986). He showed that, as we increase the complexity of test questions, there is, for most students, a rapid fall in performance when their Working Memory Space is exceeded. Another way of looking at this could be that, provided the task is within the Working Memory Space, we are testing mainly chemistry, but when the task exceeds the Working Memory Space, we are measuring a psychological artifact; the Working Memory Space. But a minority of students seem to succeed beyond this limit, by chunking. They have been taught, or have devised for themselves, strategies for breaking the over-size problem into smaller sub-tasks (or chunks) which are well within their Working Memory Space and sequencing them in such a way as to arrive at the solution. In a class situation, where such strategies have been taught, it is probably fair to set problems, which, on their face value, would appear to have a complexity greater than Working Memory Space. There is a danger, however, that teaching such chunking strategies reduces problems to algorithms which may not warrant the description, 'problem'. Before the advent of the mole and molarity, we settled for 'normality' and were able to use the relationship $V_1N_1 = V_2N_2$ to solve volumetric problems. There was no need to write formulae, balanced equations, mole ratios and so on. Volumetric normality questions come well within Working Memory Space whereas Mole questions almost inevitably come well outside Working Memory Space, and so we see in the literature a plethora of papers lamenting students' inability to solve them. The answer may be staring us in the face.

One line taken by some workers is to abandon such data-loaded questions and settle for multiple choice. This might be a partial solution (in reducing apparent load), but does multiple choice really measure the skills we think we are testing? Friel (Friel, 1979a and b) , Ambusaidi (Ambusaidi, 2000 and 2001), Johnstone (Johnstone, 2004) and Danili (Danili, 2005) have cast doubts upon multiple choice by exposing very serious problems.

This area of assessment needs much work to be done on it, taking into account the psychology involved alongside the massive efforts being made to design more clever computer programs to reduce the scoring load on teachers. 'Recall' has to be distinguished from 'recognition'; 'free and creative reasoning' has to be recognised as different from 'algorithm'; 'clever programming' must not be confused with 'better assessment'.

Problem solving

This is almost a corollary from the section above, but there are further considerations to be made. Reid (Reid, 1979), Hadden (Hadden, 1989), Wood (Wood, 1993), Yang (Yang, 2002) and Tsaparlis (Tsaparlis, 2001), have all shown that problem solving, well above the level of algorithms, is possible for pupils at middle secondary school as well as for undergraduate university students.

Using Johnstone's classification of problems (Johnstone, 2001), (Table 1) involving the information given, the methodology and the goals, these workers have exploited all eight types of problem to stimulate learning, to enhance interpersonal discussion skills, and to show chemistry to be a subject that is not remote from everyday living. Students are exposed to a large variety of different problems. Some are entirely pencil and paper, while others are conducted in the laboratory. Hadden's work (Hadden, 1989) has breathed new life into otherwise dull, routine verification laboratories in schools. Other examples can be found in some of the papers in this issue of the journal. The kinds of problems in this section are not algorithmic. Where they are complex, there is structuring to avoid overload. In some cases there are pre-problems to activate the stored information required to tackle the 'real' problem. All this development has stemmed from the clear theoretical base provided by the Information Processing Model.

Table 1. Categories of problems

Type	Data	Method	Goal
1	Complete	Familiar	Clear
2	Complete	Unfamiliar	Clear
3	Incomplete	Familiar	Clear
4	Complete	Familiar	Unclear
5	Incomplete	Unfamiliar	Clear
6	Complete	Unfamiliar	Clear
7	Incomplete	Familiar	Unclear
8	Incomplete	Unfamiliar	Unclear

Overview

This paper is really a history of a large group tackling problems of Chemical Education on a broad front and in an integrated way. It has told the story of how we have moved from data gathering to hypothesis construction. A working hypothesis has been tested and then used to influence other research leading to better teaching and learning. The Information Processing Model, which could be dismissed as sterile and soul-less, has been shown to embrace and bring together the 'schools' of Piaget, Ausubel and Novak; along with the students of Alternative Frameworks and the practitioners of Constructivism. The Model has been shown to have predictive as well as explanatory power. Like all models, it will continue to be useful only if it continues to explain the sources of the problems of the present and to point to ways of solving them for the benefit of future students.

The other papers in this issue show how things have moved on in Glasgow and elsewhere through the work of those whom I have come to regard as valued colleagues and friends. It has been impossible to discuss fully the research of the eighty workers who have been in the group and to show how each one has contributed pieces, both big and small, to the construction of a rich corpus of work.

Attitudes

Much of what appears above deals with cognitive issues, but on the way along it has been indicated that attitudes to chemistry and the development of attitudes through chemistry are of major importance. So much of the gloom, which exists at present about the future of chemistry in schools and universities, is due to negative attitudes to the subject and these must be related directly to bad experiences which pupils and students have had in chemistry lessons. The curricular changes in the 1960s and since, have, by their very structure, overloaded young people with an indigestible diet of conceptual overload. The mixture of abstraction, symbolism and formal laboratory experience (often mediated through mind-deadening worksheets) has been a huge turn-off for many. The seminal work of Reid (Reid, 1981) and his students in Glasgow has attempted to study the factors underlying attitude change and to apply them to the teaching of chemistry. His work is complementary to the cognitive work described briefly in this paper and together, these two strands of work provide a strong basis for the rethinking of the chemistry curriculum in schools and universities. Perhaps it is time for the 30 years of science education research to break out from its introspectiveness and repetitiveness and be applied to making the learning of chemistry an enjoyable and exciting experience for all young people.

Nothing that has been said in this paper denies or denigrates the valuable work of many researchers in many parts of the world, but, to keep the size of the paper within bounds it has been confined to the 'Glasgow School', the subject of this issue of CERP.

I am grateful to the editors and the editorial board of this journal for affording us the opportunity to present a very large corpus of work in a coherent way, to show how it has arisen, how it has fitted together and how a powerful set of tools has been fashioned to allow Chemical Education to meet the needs of our students and ultimately contribute to the needs and well-being of society. I am also indebted to my colleagues and former students who have contributed so generously to this issue of CERP. I commend their papers to the reader's attention, illumination and enjoyment.

References

Al-Naeme F. and Johnstone A.H., (1991), Room for scientific thought, *International Journal of Science Education*, **13**, 187-192.

Al-Shuaili A. and Johnstone A.H., (2001), Learning in the laboratory: some thoughts from the literature, *University Chemistry Education*, **5**, 42-51.

Ambusaidi A. and Johnstone A.H., (2000), Fixed response questions: what are we testing? *Chemistry Education Research and Practice,***1**, 323-328.

Ambusaidi A. and Johnstone A.H., (2001), Fixed response questions with a difference, *Chemistry Education Research and Practice*, **2**, 313-327.

Ausubel D.P., Novak J.D. and Hanesian H., (1978), *Educational psychology: a cognitive view*, (2nd Edition), New York, Holt, Rinehart and Winston.

Baddeley A.D., (1999), *Essentials of human memory*, Hove, Psychology Press Ltd.

Cassels J.R.T. and Johnstone A.H., (1985), *Words that matter in science*, London, Royal Society of Chemistry.

Danili E. and Reid N., (2005), Assessment formats: do they make a difference? *Chemistry Education Research and Practice*, **6**, 204-212.

Duncan I.M. and Johnstone A.H., (1973), The mole concept in chemistry, *Education in Chemistry*, **10,** 213.

El-Banna H. and Johnstone A.H., (1986), Capacities, demands and processes: a predictive model for science education, *Education in Chemistry*, **23**, 80-84.

Friel S. and Johnstone A.H., (1979), Second thoughts in multiple-choice tests, *Journal of Chemical Education*, **56**, 326.

Friel S. and Johnstone A.H., (1979), Does the position of the answer in a multiple-choice test matter? *Education in Chemistry*, **16**, 175.

Garforth F.M., Johnstone A.H. and Lazonby J., (1976), Ionic equations; difficulties in understanding and use, *Education in Chemistry*, **13**, 72.

Hadden R.A. and Johnstone A.H., (1989), P*ractical problem solving for Standard Grade chemistry*, Glasgow, Scottish Consultative Committee on the Curriculum.

Howe T.V. and Johnstone A.H., (1971), R*eason or memory? The learning of formulae and equations*, Edinburgh, National Curriculum Development Centre Bulletin 1.

Ingle R. and Shayer M., (1971), Conceptual demand in Nuffield 'O' level chemistry, *Education in Chemistry*, **8**,182-183.

Johnstone A.H. and Reid N., (1981), Interactive teaching materials in science, *South Australia Science Teachers Journal*, No **812**, 4-15.

Johnstone A.H., (1982), Macro and microchemistry, *School Science Review*, **64**, 377-379.

Johnstone A.H., (1991), Why is science difficult to learn? Things are seldom what they seem, *Journal of Computer Assisted Learning*, **7**, 75-83.

Johnstone A.H., (2001), Can problem solving be taught? U*niversity Chemistry Education*, **5**, 12-18.

Johnstone A.H., (2004), *Effective practice in objective assessment*, Learning and Teaching Support Network Physical Sciences Practice Guide, Hull, LTSN Physical Sciences Centre.

Kellett N.C. and Johnstone A.H., (1980), Learning difficulties in school science – towards a working hypothesis, *European Journal of Science Education*, **2**, 175-181.

Koffka K., (1933), *Principles of Gestalt psychology*, New York, Harcourt, Brace and World.

Letton K.M. and Johnstone A.H., (1991), Practical measures for practical work, *Education in Chemistry*, **28**, 81-83.

MacDonald J., Johnstone A.H. and Webb G., (1977), Misconceptions in school thermodynamics, *Physics Education*, **12**, 248.

Miller G.A., (1956), The magical number seven plus or minus two: some limits to our capacity for processing information, *Psychological Review*, **63**, 81-97.

Pascual-Leone J., (1970), Mathematical model for the transition role in Piaget's developmental stages, *Acta psychologica*, **63**, 301-345.

Reid N. and Johnstone A.H., (1979), Bringing chemical industry into the classroom, *Chemistry and Industry*, **4**, 122-123.

Reid N. and Johnstone A.H., (1981), Towards a model for attitude change, *European Journal of Science Education*, **3**, 205-212.

Reid N., (1981), Interactive teaching materials in science, *South Australian Science Teachers' Journal*, No **812**, 4-15.

Selepeng D. and Johnstone A.H., (2001), A language problem revisited, *Chemistry Education Research and Practice*, **2**, 19-29.

Shayer M., (1970), How to assess science courses, *Education in Chemistry*, **7**, 182-186

Sirhan G., Gray C., Johnstone A.H. and Reid N., (1999), Preparing the mind of the learner, *University Chemistry Education*, **3**, 43-46.

Sirhan G. and Reid N.,(2001), Preparing the mind of the learner, Part 2, *University Chemistry Education*, 5, 52-58.

Sleet R.J., Johnstone A.H. and Vianna J.F., (1994), An information processing model of learning: its application to an undergraduate laboratory course in chemistry, *Studies in Higher Education*, **19**, 77-88.

Tasker R., Dalton R., Sleet R., Bucat R., Chia W. and Corrigan D., (2002), Description of VisChem: visualising chemical structures and reactions at the molecular level to develop a deep understanding of chemistry concepts. Retrievable from the *Australian Universities Teaching Committee (AUTC) Learning Designs* Web site: http://www.learningdesigns.uow.edu.au/exemplars/info/LD9/index.html

Tsaparlis G. and Stamovlasis D., (2001), Application of complexity theory to an information processing model in science education, *Non Linear Dynamics, Psychology and Life Sciences*, **5**, 267-286.

Wham A.J.B. and Johnstone A.H., (1982), Demands of practical work, *Education in Chemistry*, **19**, 71-73.

Witkin H.A., Moore C.A., Goodenough D.R. and Cox P.W., (1977), Field-dependent and field-independent cognitive styles and their educational implications, *Review of Educational Research*, **47**, 1-64.

Wood C.A., (1993), *Creative problem solving in chemistry*, London, Royal Society of Chemistry.

Yang M-J. and Reid N., (2002), Open-ended problem solving in school chemistry: a preliminary investigation, *International Journal of Science Education*, **24**, 1313-1332.

Zaman T.U., Johnstone A.H. and Watt A., (1998), The students' attitude, and cognition change, to a physics laboratory, *Physics Education*, **33**, 22-29.

Educational research

Cognitive factors that can potentially affect pupils' test performance

Eleni Danili and Norman Reid

Centre for Science Education, University of Glasgow, Glasgow G12 8QQ, UK
e-mail: n.reid@mis.gla.ac.uk

Received 23 August 2005, accepted 9 February 2006

Abstract. The two cognitive styles, 'field dependent/field independent' and 'convergent/divergent', were explored in relation to three formats of assessment (multiple choice, short answer and structural communication grid) in five classroom chemistry tests. The study was conducted in Greece with the participation of Grade-10 (upper secondary school) pupils (age 15-16). The field dependent/independent characteristic correlated with pupils' performance in all the tests, and in almost all the formats of assessment. The convergent/divergent characteristic correlated with pupils' performance in assessment where language was an important factor, but not in algorithmic types of questions or in questions where there is a greater use of symbols and less use of words. It seems that, in relation to the convergent/divergent characteristic, the chemistry content and presentation of the test is a factor affecting the type of questions being asked. This study suggests that some of the factors that affect pupils' performance might be: (a) the content and presentation of the test, (b) the format of the test, (c) the psychology of the individual. [*Chem. Educ. Res. Pract.,* 2006, **7** (2), 64-83]

Keywords: assessment in chemistry, cognitive factors, cognitive styles, field dependence/independence, convergence/divergence

Background

In a previous study, correlations between different formats of assessment were found to be between 0.30 and 0.71 (Danili and Reid, 2005). This is a wide range and even the maximum of the correlations is less than 1 by a significant margin. This suggests that the best student found by one method is not necessarily the best student by another method. If the two formats of assessment were simply testing the same content, then very high correlation would be expected. Of course it can be argued that different formats of assessment test different abilities of the examinees and, therefore, it is fairer to use several formats to assess student skills and knowledge. However, the fundamental issues arising from the above study were:

- Are the different formats of questions testing different abilities? Which format of assessment is more valid?
- Are the different formats related to differences between students in one or more psychological traits?
- It might be reasonable to suppose that the use of multiple formats of assessment tests students more fairly than the use of a single format but on what basis can it be justified?

To address some of these questions, a study was designed which engaged 476 pupils (Grade 10, age 15-16) of 12 public upper secondary schools in Greece during the school year September 2002-May 2003. In this study, the two cognitive styles ‘field dependent/field independent’ and ‘convergent/divergent’ were explored in relation to three formats of assessment (multiple choice, short answer and structural communication grid) in five classroom chemistry tests.

Assessment formats in a dynamic interaction with teaching and learning models

Teaching and assessment are inseparable in the learning process. Assessment does not stand outside teaching and learning but stands in a dynamic interaction with them. Shepard (1992) emphasised the importance for educators to understand the conception of teaching and learning when they make decisions about testing practice and to examine the implicit theories which guide their practice. In the traditional model of teaching (sometimes described as objectivism as distinct from constructivism) learning is seen in terms of a distinct body of information, specified in detail, that can be transmitted to the learner. Assessment, in this context, consists of checking whether the information has been received (Entwistle, 1992). However, isolated facts, if learnt, are often not easily recalled from the memory because they have no meaning and do not fit into the learner’s conceptual map. Students can succeed in objective tests without necessarily understanding the material they have learned. This may be true particularly in science where much research has shown that students carry widespread misconceptions and misunderstanding of scientific phenomena (Osborne and Cosgrove, 1983; Nurrenbern and Pickering, 1987; Sawrey, 1990; Andersson, 1990; Bodner, 1991; Gabel, 1999). The behaviourist learning theory requires practice, repetition and testing of discrete basic skills prior to any teaching of higher-order thinking skills (Shepard, 1992).

In contrast, in the constructivist and information processing models learning is seen as a process of personal knowledge construction and meaning making. In this approach, learning is a complex and diverse process and therefore requires assessment to be more diverse, and to assess in more depth the structure and quality of students’ learning and understanding (Gipps, 1994). In the information processing models, the structure of effective learning is seen in such a way that knowledge can be stored usefully in the long-term memory. Knowledge is seen as something cohesive and holistic which provides scaffolding for later learning (Atkins et al., 1992). In fact, cognitive processes indicate that there is an intimate connection between skills and the contexts in which they are used. This means that assessment should reduce the emphasis on the ability to memorise, and increase the emphasis on thinking and problem solving. Information processing approaches to learning require a new assessment methodology, and tests ought not to ask for demonstration of small, discrete skills practised in isolation (Gipps, 1994).

The importance of aligning teaching methods and assessment tasks is stressed in many publications pertaining to the curriculum (Osborne, 2004). However, over the last decade, the amount of assessment in schools has increased. Consequently, the assessment workload for the teachers has grown dramatically, and the time available to devote to assessing each student has fallen. It is tempting to reduce marking workload by using objective tests. Objective testing assessment policy is based on objectivistic theories and is greatly concerned with quantitative measurement (Biggs, 1996). The quality of such assessment is embodied in notions of reliability and validity (Broadfoot and Black, 2004). Unfortunately, objective assessment practices can inadvertently de-skill students in various ways. They focus attention on the immediate tasks of passing examinations or completing tasks and distract students from the more vital task of learning how to assess themselves (Boud, 2004). This tradition is very much opposite to a *constructivist* theory of learning, which regards learning more in qualitative than quantitative terms (Biggs, 1996). According to constructivist theory, assessments policy should be based on performance on open-ended tasks which can reveal a wide variety of insights of thinking processes in students' written responses.

Cognitive styles

Learning theories are the bases which help teachers and educators to understand diverse factors of individual differentiation in: perceiving information; encoding information; transferring information; scanning the representation of the information; and working memory capacity. There are also individual differences in styles of remembering, thinking, and judging, and these individual variations, if not directly part of the personality, are at the very least intimately associated with various non-cognitive dimensions of personality (Kogan, 1976). Differences in the above factors are brought together to suggest that individuals have different cognitive styles and are different in intelligence, ability, personality, and achievement. It seems that our cognitive style influences our: intellectual abilities; skills; personalities; teaching and learning; and performance. According to Messick's (1993) definition "*cognitive styles are characteristic modes of perceiving, remembering, thinking, problem solving, decision making that are reflective of information processing regularities that develop in congenial ways*".

There have been various arguments relating to the overlap between style and ability. Some researchers support the idea that 'ability' describes performance in a given task whereas 'style' describes the way the task is approached (Messick, 1994) While intellectual abilities are primarily concerned with the ability to learn, cognitive styles are primarily concerned with differences in the ways of learning. Riding and Cheema (1991) considered cognitive style to be a fairly fixed characteristic of an individual while cognitive strategies are the ways that may be used to cope with particular situations and tasks. Strategies may be learned and developed. Styles, by contrast, are static and are relatively in-built features of the individual.

In the literature there are various labels of cognitive styles. In this study attention was focussed on field dependent/independent and convergent/divergent cognitive styles. The reasons for that were:

- they are dominant over the other cognitive styles in the literature.
- previous work suggests that they are related to assessment.

Field-dependent /independent cognitive style

Hundreds, if not thousands, of articles pertaining to the field dependence-independence (FDI) construct have been published. This polar construct originated in Witkin's work (Witkin et al., 1974; Witkin et al., 1977; Witkin and Goodenough, 1981). Witkin and Goodenough (1981) investigated for many years the idea suggested by Gestalt psychology, that some people are dominated by any strong frame of reference or pattern in a stimulus field, to such an extent that they have trouble in perceiving elements that cut across the pattern. He investigated the personality in relation to the integrative process of making contact with the environment through perception.

Early studies of Witkin and Asch (1948a, 1948b) found that some individuals consistently tended to attend to different type of cues. Subjects who used visual cues were designated 'field-dependent', while those who used postural cues (such as tactile, vestibular and kinaesthetic cues) were designated 'field-independent'. Further probes of the subject's ability to perceive individual elements within an organised perceptual field have followed. It was thought that there might be a relationship between the individuals 'disembedding ability' and their 'cognitive restructuring'.

Within this framework, Witkin and Goodenough (1981) defined the main characteristic of the field-dependent and field-independent cognitive styles as:

- *Field - Dependent* (FD) individual: one who can insufficiently separate an item from its context and who readily accepts the dominating field or context.
- *Field - Independent* (FID) individual: one who can easily 'break up' an organised perceptual field and separate readily an item from its context.

A number of studies have followed in examining the correlation between field dependency/independency (FDI) and academic performance in disciplines such as language, mathematics, natural sciences, social sciences, art, music and computer science at secondary school level as well as at university level. Tinajero and Paramo's (1998) review concluded that "in general field-independent subjects perform better than field-dependent subjects, whether assessment is of specific disciplines or across the board".

Many research studies (e.g. Johnstone and El-Banna, 1986; Al-Naeme and Johnstone, 1991; Bahar and Hansell, 2000; Danili and Reid, 2004; Tsaparlis, 2005) have looked at FDI. Their findings are that students' performances are consistent with the conclusion of Tinajero and Paramo (Tinajero and Paramo, 1998). Overall, the field dependent/independent test is considered by many researchers a very powerful instrument to predict academic performance of individuals (e.g. Terrell, 2002).

Convergent / divergent cognitive style

Research on the Convergence-Divergence cognitive style has not received as much attention as the FDI cognitive style from educators and researchers. The idea of convergent-divergent cognitive style has its origin with Hudson (1966) who, as an undergraduate, had found himself better at some parts of intelligence tests than others: good at the diagrammatic questions, and relatively poor at the verbal and numerical ones. At that time, there was a growing feeling that typical intelligence tests did not measure all aspects of intelligence. It was argued that such tests only measured what was termed 'convergent thinking' and not 'divergent thinking'. Convergent thinking means that someone has to focus down (converge) on the one right answer in order to find the solution of a problem. Convergent thinkers score highly in problems requiring one conventionally accepted solution clearly obtainable from the information available (as in intelligence tests), while at the same time obtaining low scores in problems requiring the generation of several equally acceptable solutions. On the other hand, divergent thinking is the opposite of this approach. Divergent thinking deals with the capacity to generate responses, to invent new ones, to explore and expand ideas, and in a word, to diverge. Convergent thinking thus demands close reasoning; divergent thinking demands fluency and flexibility (Child and Smithers, 1973).

Hudson (1966, 1968) thought that he might be able to measure arts/science aptitude and made an attempt to devise tests of aptitude for arts and science respectively. In the traditional IQ test, the individual is required to find the one right answer for a problem, being invited to choose this right answer from a list of alternatives. The new tests do not require the respondent to produce one right answer, and like intelligence tests, can take different forms. Such questions can consider imaginative themes by asking questions such as:

- How many uses can you think of for each of the following objects?
- How many meanings can you think of for each of the following words?

- Draw a picture in the space below to illustrate the title ‘Zebra Crossing’.

According to Hudson (1966) “*the converger is the boy (*sic*) who is substantially better at the intelligence test than he is at the open-ended tests; the diverger is the reverse*”.

Most of the research related to convergent/divergent styles has concentrated on the relationship between divergent thinking and arts-science orientation. Research showed that there is a tendency for convergers to choose science subjects. Johnstone and Al-Naeme (1995) indicated that much science teaching is convergent and students are rewarded for convergent thinking leading to unique specific answers. However, this may not to be the case for biology because it attracts both groups of students (Orton, 1992; Bahar, 1999). Bahar’s statement was that, “*biology might be one of the science branches in which students might cope equally well with a convergent or a divergent bias”*.

Many researchers tended to equate divergent thinking with creativity and convergent thinking with intelligence. This has caused a great deal of controversy, with different research supporting different results (Nuttall, 1972; Bennett, 1973; Runco, 1986; Fryer, 1996). In the literature little research is reported on convergent/divergent cognitive styles and performance in science. However, Al-Naeme’s (1988) research showed that divergent students had higher scores than convergent students in mini projects in chemistry.

Bahar (1999) showed that divergent pupils/students did not perform better in all cases compared to convergent pupils/students. Thus, he suggested that the answer might be related with assessment techniques. He said “…*when one is looking at the relationship between students’ performance in any topic and their cognitive styles, the type of assessment techniques used, such as multiple choice type of questions, essay questions, projects and so forth should be reported because a particular type of assessment technique may favour a particular kind of cognitive style”*.

In general, it seems that the language and the format of questions in relation to the cognitive style of an individual may be able to influence his/her performance.

What should be the aims of educational measurement?

From the above, it can be concluded that cognitive styles influence the personality of the individuals, and affect the psychological behaviours that indicate how learners perceive, interact with and respond to the learning environment **(Fatt, 2000).** Accordingly, cognitive styles have an impact on pupils’ performance and achievement. Therefore, the concern of educators should be to understand, from the heterogeneous mix of pupils’ learning styles, the possible styles, so that teachers can best adapt their teaching style and assessment materials to suit the pupils’ preferred styles and help them to overcome their difficulties and display their abilities. This is a daunting prospect for the teacher!

Furthermore, if assessment is to be part of teaching, then first it has to be seen that way. Most areas of learning have both mental and physical aspects. However, all learning has an emotional aspect and numerous research studies emphasise the importance of learner confidence, motivation and self-esteem, which are prerequisites for successful learning, and need to be encouraged. Therefore, the negative or positive impact of different forms of assessment on motivation and self-esteem need to be considered seriously. Thus, there is a need to reinforce pupils’ motivation by assessing them with appropriate format questions and, therefore, to enable them to show their best performance (Gipps, 1994). Assessment must be humane (Johnstone, 2003). Humanity takes into account factors that affect pupils/students’ performance, such us cognitive and psychological traits of individual personality.

If the purpose of educational measurement is 'how well' rather than 'how many', then this requires a quite different approach to test construction. Gipps (1994) pointed out that we need a more measured, analytical approach to assessment in education. We need to resist the tendency to think in simplistic terms about one particular form of assessment being better than other: consideration of form without consideration of purpose is wasted effort. She called for wider understanding of the effect of assessment on teaching and learning and fostering a system, which supports multiple methods of assessment while at the same time making sure that each one is used appropriately.

Thus, an important question in education is: what should be the aims of educational measurement? Should the aim be to devise tests, which look at the individual and find out 'how well' or should look at 'how many'? Should the aim be to devise tests to support learning or to devise tests to evaluate teachers according to pupils' performance and achievement? The dangers of 'teaching to the test' are well known and, if only a limited range of facts and skills are assessed, and, if 'high stakes' are attached to the results in terms of the consequences of the publication of 'league tables', then we can expect teachers to teach to the test and restrict the curriculum accordingly.

Wood (1986) argued that educational measurement aims should be *"to devise tests which look at the individual and find out how well rather than how many"*. Wood's definition of educational measurement was that it:

- deals with the individual's achievement relative to himself rather than to others;
- seeks to test for competence rather than for intelligence;
- takes place in relatively uncontrolled conditions and so does not produce 'well-behaved' data;
- looks for 'best' rather than 'typical' performances;
- is not effective unless rules and regulations characteristic of standardised testing are maintained;
- embodies a constructive outlook on assessment where the aim is to help rather than condemn the individual.

There is also a need for distinction between competence and performance. Gipps (1994) said that *"Competence refers to what a person can do under ideal circumstances, while performance refers to what is actually done under existing circumstances. Competence includes the ability to access and utilise knowledge structures, as well as motivational, affective and cognitive factors that influence the response"*. Thus, according to Messick (1984), *"a student's competence might not be revealed in either classroom performance or test performance because of personal or circumstantial factors that affect behaviour"*.

It is important for educators to think of the impact on motivation and self-esteem if they use the wrong tools to assess their students. It is also important to find the assessment forms that are appropriate to individuals and to elicit the best performance from them. In order to do that, educators should be aware of the learning theories, which seek to understand why the students so often face difficulties and to align assessment with these theories. Furthermore, different types of assessment seem to encourage deep or surface approaches to learning (Struyven et al., 2002). For example, fixed response questions may encourage students to think dualistically even if designed to go beyond recall issues because, at the end of the day, students are asked to select one right answer. Therefore, it is argued that the content and style of a test have an important message to students about the nature of science and their intellectual development (Boud, 1995).

Understanding the psychological processes, which underpin learning, may provide useful information to avoid constructing questions which may be beyond any reasonable expectation of student's abilities. Crisp and Sweiry (2003) emphasised the importance of how a question is understood by subtle changes of certain aspects of a question such as diagrams or images, which are particularly salient and hence can come to dominate the mental representation that is formed. Many researchers (e.g. Oakhill, 1988; Davey, 1990) use the information processing model to explain the difficulties that pupils face when they try to answer negative questions.

The study of Lu and Suen (1995) showed that pupils' performance in different formats of assessment are related to their cognitive style. Pollitt et al. (2000) also addressed the problems related to the language barrier that students face when they study in a language which is not their mother tongue. They concluded that the problems are linguistic, contextual and cultural. However, language problems may simply reflect information overload (Selepeng and Johnstone, 2001).

Measuring instruments of the study

The following measuring instruments were employed to gather information from the pupils:

- Two cognitive tests:
  - Field Dependent/ Field Independent test
  - Convergent/ Divergent test
- Five chemistry paper-and-pencil tests

Measurement of field dependency

A version of the Witkin et al. (1971) group embedded figures test was used to determine an individual's degree of field dependency. It is almost identical to that used by Witkin and was used by Bahar (1999) in his study of cognitive structure. Known as the Hidden Figure Test (H.F.T.), it comprises twenty complex figures plus two additional, introductory figures that were used as examples. Simple geometric target shapes are presented on the back of a booklet. Pupils are required to recognise and identify one of the target shapes embedded within each of the complex figures. They do this by tracing its outline with a pen or a pencil. The main scoring scheme for the tests is to give one point for a correct simple shape embedded in a complex figure. The overall sum of the scores is the total mark that a student can gain. Thus the possible maximum score that can be obtained is 20.

Measurement of convergency/divergency

The Convergent /Divergent test consisted of 6 mini tests, described below.

- Test 1 was designed to find out the subjects' ability to generate words of the same or similar meaning to those given. At the beginning of the test an example was provided to show what the pupil was required to do. For example, if the word 'short' was given, a set of words such as 'abbreviated, limited, brief, concise, momentary, little, abrupt, petite, crisp, and compact' might be expected. This test included three questions and the time given for this test was 4 minutes.
- Test 2 asked the pupils to construct as many sentences as possible using four given specific words in each sentence, the words to be used in the form as given. Any sentences which did not make sense, received no credit. An example was provided at the beginning of the test and the time given for the test was 4 minutes.

- Test 3 is the only test which is not verbal. This is because there are some pupils who are pictorial learners and thinkers and, therefore, they perceive ideas more easily by pictures and diagrams. Thus, a pictorial test was included to give an opportunity to this type of student. In this test the student was required to draw up to five different pictures to relate to the idea of the given word. An example was given at the beginning of the test and 5 minutes was allowed.
- The purpose of test 4 was to see how many things the students could think of that are alike in some way. They were asked to write all the things that are round, or that are round more often than any other shape. An example was given at the beginning and 2 minutes was allowed for it.
- The objective of test 5 was to measure the student's ability to think of as many words as they could that begin with one letter and end with another. For example, students were asked about the words, which begin with the letter G, and end with the letter N. Names of people or places were not allowed and the time limit was 2 minutes.
- Test 6 aimed to find how many ideas the students could think of about a given topic. They had to list all the ideas they could about a topic whether or not it seemed important to them. An example was given at the beginning of the test and 3 minutes were allowed to complete this test.

The total time allocated for these six mini tests was 20 minutes. The researcher controlled the time limit for each test during the session. The test was translated (free translation) into Greek and two Language teachers checked the clarity of Greek carefully. The aim was to detect possible ambiguities and sources of confusion. In order to measure pupils' performance, one mark was given for every single correct response (Hudson, 1966). The test had been widely used in measuring extent of divergency by Bahar (1999).

Chemistry tests

Five paper-and-pencil chemistry tests were given to the pupils. Each chemistry test was designed to assess pupils by a range of question formats which tested the same knowledge and understanding in the same topics. The range of question formats that have been used in the project is shown in Table 1. This choice was made on the basis of the expectation that multiple choice questions would favour the convergent pupil, open-ended questions would be more congenial for the divergent pupils and that structural communication grids would appeal to both convergent and divergent pupils. For readers who are less acquainted with structural communication grids, examples of this format are shown in the appendix.

Because the teachers had to replace their classroom tests with the researcher's tests, the experimental tests were designed with the teachers' advice in mind, and an attempt was made to keep them short and appropriately demanding. The tests were constructed to be similar to the study questions of the Greek chemistry textbook (Liodakis et al., 1999), the Scottish Standard Grade Chemistry book (Renfrew and Conquest, 1995), and the book by Moore et al. (1998).

The tests were based on:

- Test 1: Atomic structure, classification of matter, solubility.
- Test 2: The periodic table and chemical bonds.
- Test 3: Mole concept.
- Test 4: Acids, alkalis, pH, neutralisation.
- Test 5: Solutions.

Table 1. Combination of different format questions.

	Test 1	Test 2	Test 3	Test 4	Test 5
Question Format	MC	SA	SA	MC	MC
	SA	SCG	SCG	SCG	SCG
					SA

MC: Multiple-Choice
SA: Short-Answer (Open-Ended)
SCG: Structural Communication Grid

Sampling method and administration procedures

The study was conducted in Greece during the school year September 2002 - May 2003 and 12 public upper secondary schools (*lykeio*) participated. There was more than one class in some schools and, therefore, the total number of classes was 23 and the total number of teachers was 12 (one teacher in each school). The classes were of different size: the smallest had 11 pupils and the largest had 29 pupils. Table 2 outlines the whole plan of the study.

Table 2. Schools and classes involved in the study.

Schools	Number of classes in each school	Number of pupils in each class				Total number of pupils
1	1	19				19
2	3	21	19	14		54
3	1	21				21
4	1	14				14
5	2	18	18			36
6	4	18	18	17	11	64
7	3	27	24	23		74
8	2	29	27			56
9	1	25				25
10	2	21	22			43
11	1	26				26
12	2	25	19			44
Total	**23**					**476**

It was decided to work with the pupils of Grade 10 (age 15-16) because, at that stage, pupils do not participate in national exams and teachers are more willing to be involved in research. Another very important reason to work with Grade-10 pupils was that all pupils have to attend chemistry lessons as, at that stage, pupils have not yet been split into arts or science streams. Thus, classes were heterogeneous, with pupils of different abilities and subject orientation.

The schools were not chosen at random because of the nature of the research. The researcher contacted teachers of different schools, before the beginning of the school year, and explained to them the plan of the study. The schools were selected in different geographical areas and of a different socio-economic background as much as possible. After receiving the teachers' agreement on the project, the researcher applied to the Greek Pedagogic Institute and Greek Ministry of Education for permission to have access to schools in order to administer the tests.

The schools and the number of pupils who participated in each test are summarised in Table 3. As can be seen from this table, the participating schools opted out of many tests, and this was unfortunate. The reasons for that were thought to be the lack of provision of organised training and educational studies for teachers as well as the very small amount of teaching time (just two 45 minute periods per week through the year). This makes teachers concerned about finishing the teaching units, and they are not willing to spend time on the evaluation of their teaching.

Table 3. Number of schools and pupils who participated in each test.

		Test 1	Test 2	Test 3	Test 4	Test 5
Schools	1	*				
	2			*		
	3	*				
	4	*				
	5	*				
	6		*	*		
	7	*				
	8	*	*		*	
	9	*	*		*	
	10	*	*			
	11					*
	12			*		*
Number of pupils		288	185	146	75	64

The cognitive tests were administered by the researcher. The class teachers in the various schools administered the chemistry tests.

Statistical methods used in the research

Validation and reliability of the instruments

The cognitive tests were based on well-established techniques. The field dependency test was almost identical to that of Witkin et al. (1971) test, while Bahar (1999) had developed and tested the materials used for the convergent/divergent test, based on standard tests for convergence-divergence.

Most statistical tests of reliability (other than test and re-test) indicate internal consistency. This procedure was not used in any of the chemistry tests because the tests consisted of sections having heterogeneous items assessing a mix of modes, degrees of difficulty and different depths of understanding. The chemistry tests were discussed with experienced class teachers in Greece to check face validity and minor adjustments were made.

Correlations

Both Pearson coefficient and Spearman's rho correlation between the different formats of the questions were calculated and were found to give similar values. However, because the distributions were frequently observed to deviate from a normal distribution, the Spearman's rho coefficient was more appropriate and this was used in all subsequent discussion.

Findings of the study

The H.F.T. results

The mean score of the H.F.T test was 7.8 (minimum = 0, maximum = 20) and the standard deviation was 4.2.

The CV/DV test results

The mean score of the convergent/divergent test was 47 (minimum = 0, maximum = 75) and the standard deviation was 10.7.

Chemistry test results

Table 4 shows the mean scores (%) and the standard deviations of each section of all the tests.

Table 4. Mean scores and s.d. of each format of each test.

	Test 1 (N=288)		Test 2 (N=185)		Test 3 (N=146)		Test 4 (N=75)		Test 5 (N=64)		
	MC	SA	SA	SCG	SA	SCG	MC	SCG	MC	SCG	SA
Mean	64.3	53.5	52.2	36.7	60.9	67.7	53.0	68. 0	67.5	68.6	50.8
s.d.	20.4	25.6	30.7	25.4	37.3	36.7	22.7	24.1	28.6	26.4	35.6

SA Short answer test
MC Multiple choice test
SCG Structural communication grid test

In general the Short Answer sections of each test were found to be more difficult than Multiple Choice and Structural Communication Grid sections of the tests. [Test 4 was a new test and was not part of the test battery used in our previous paper (Danili and Reid, 2005).]

Correlations between field dependent/independent cognitive scores and chemistry scores

Table 5 summarises the correlations for the field dependent/independent characteristic. It shows that the field dependent/independent characteristic correlated with pupils' performance in all the tests and in all the formats of assessment (although not always significantly). Being field independent seems to be a very important factor which influences whether pupils perform well in almost all types of assessments, and irrespective of the content of the questions. This result is consistent with the majority of the research in this field (e.g. Tinajero and Paramo, 1998; Danili and Reid, 2004; Tsaparlis, 2005) and the correlation values obtained here are very typical of previous work. Although significant, correlation values at this level indicate that only a small percentage of the variance is related to the field dependency skill. The short answer format of assessment favours field independent pupils more than the grid format of assessment. This is seen in tests 2 and 3, although the effect is not observed in test 5 with its smaller sample.

Table 5. Spearman's rho correlations for the field dependent/independent characteristic.

	Test 1 (N=288)		Test 2 (N=185)		Test 3 (N=146)		Test 4 (N=75)		Test 5 (N=64)		
	MC	SA	SA	SCG	SA	SCG	MC	SCG	MC	SCG	SA
rho	0.25	0.29	0.31	0.12	0.32	0.19	0.12	0.31	0.26	0.39	0.40
sig.	**	**	**	NS	**	*	NS	**	*	**	**

** Correlation is significant at the 0.01 level.
* Correlation is significant at the 0.05 level.
NS not significant

Although the field dependent/independent characteristic may develop naturally with experience, it may be difficult to teach someone to be field independent. However, attention should be given in the construction of the assessment to avoid confusion for those who are not able to separate the important information from the unimportant although, in some cases, the ability to separate the 'message' from the 'noise' may be an important skill to test. Thus, shredding (a process where a group of 'experts' scrutinise questions to ensure content validity, and to remove ambiguity and other errors) is a necessary process for quality assessment. Superficial clues, negative and double negative expressions, or subtle aspects, which can come to dominate the mental representations, should be avoided (Crisp and Sweiry, 2003; Johnstone, 2003).

Correlations between the convergent/divergent scores and chemistry scores

Table 6 summarises the correlations between the convergent/divergent characteristic scores and chemistry scores for different formats of assessment in the five chemistry tests. In general, the convergent/divergent characteristic correlated with pupils' performance in assessment, where language was an important factor to perform well (e.g. test 1, 2). Thus, in assessments that require pupils to have linguistic skills in order to elaborate and interpret a given text or to explain phenomena, ideas and concepts, or to describe differences, the convergent/divergent style is an important factor for pupils to perform well.

Table 6. Spearman's rho correlations for the convergence/divergence characteristic.

	Test 1 (N=288)		Test 2 (N=185)		Test 3 (N=146)		Test 4 (N=75)		Test 5 (N=64)		
	MC	SA	SA	SCG	SA	SCG	MC	SCG	MC	SCG	SA
rho	0.34	0.29	0.32	0.16	0.12	0.04	0.07	0.37	0.04	0.05	-0.13
sig.	**	**	**	*	NS	NS	NS	**	NS	NS	NS

** Correlation is significant at the 0.01 level.
* Correlation is significant at the 0.05 level.
NS not significant

It is possible to suggest that the short answer or open-ended questions favour divergent style pupils more than objective questions do. This is because, in short answer questions, pupils need to articulate their thoughts, and divergent pupils are better able to do this. In objective testing, if a question needs reading skill in order to elaborate and interpret a text given, then again the convergent/divergent style may be a very important factor for success. However, in algorithmic types of questions or in questions where there is more use of symbols and less use of words, such as test 3 and 5, or MC questions 1, 2, 3, 12 of test 1, the convergent/divergent characteristic does not relate to pupils' performance. In this case the format of the questions does not have an effect on pupils' performance.

From the above findings it seems that, in relation to the convergent/divergent characteristic, the chemistry content may well be a factor affecting the type of questions being asked, and may allow the question to be more easily tackled by, say, a divergent pupil. However, in almost all the tests the divergent pupils outperformed convergent pupils and, when there were short answer questions or open-ended questions, the differences in the performance between the convergent and divergent groups became larger. Hudson (1966) noted that the convergers tended to chose the sciences, but the divergers who did choose the sciences performed very well. The results here confirm the advantage of being divergent in examinations.

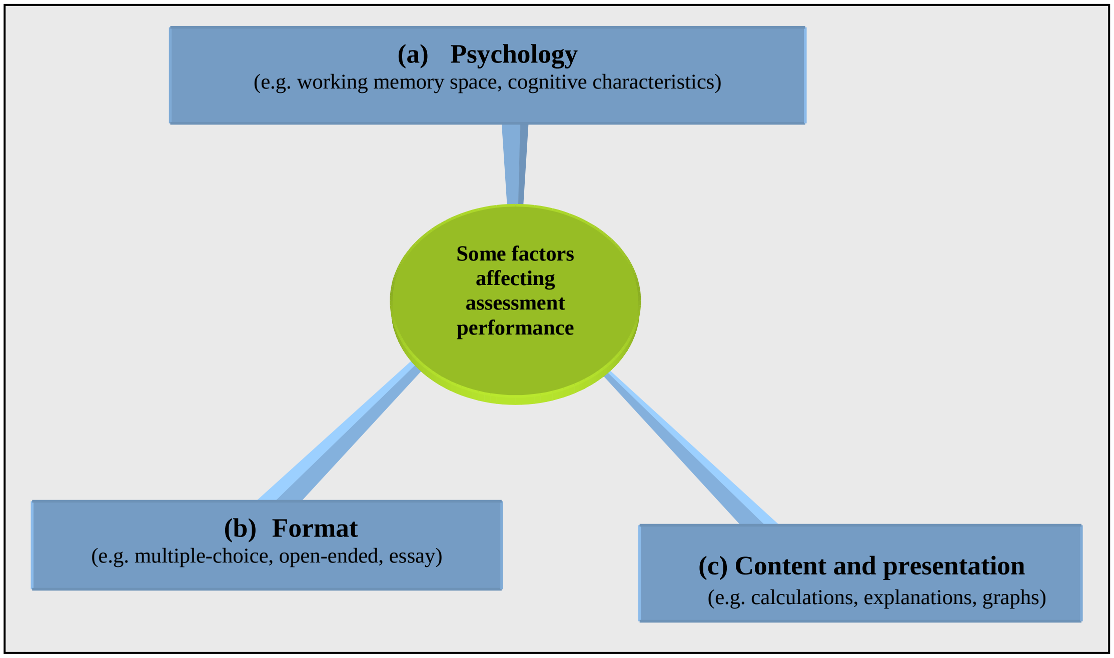
These outcomes are consistent with the work of Runco (1986), who indicated that there were particular domains of performance, for example art and writing, which were more strongly related to divergent thinking than other areas such as music and science. These results also might explain what Hudson (1966) pointed out: "*the convergence/ divergence dimension is a measure of bias, not a level of ability*". Thus, there is a tendency for those who are more divergent to feel more comfortable with some arts-orientated subjects. The closed nature of much early science teaching and learning may tend to attract those who are more convergent, while arts subjects offer opportunities for more extended thought, and attract those with good linguistic skills. However, if pupils who are good in linguistic skills choose science, it seems that they perform better than those who do not have such skills, because of their superiority in language. Linguistic skills such as comprehension and interpreting a scientific text are of paramount importance for reasoning in science (Byrne, et al., 1994). The results of Johnstone and Al-Naeme (1991) and of Field and Poole (1970) research offer some support for the claims.

In general, it seems that there is a relationship between the convergent/divergent characteristic and language. Although there is a debate as to whether thinking ability encourages divergence or divergent traits encourage thinking, it seems that the quality of a child's preschool language environment emerges as vital and, as Wittgenstein (1961) argued, the limits of one's language are the limits of one's world (Sutherland, 1992). Here is the importance of the teacher's role. The teacher should extend and challenge the child to go beyond where he would otherwise have been (Vygotsky, 1986). There is a need for teachers to encourage pupils to make their meaning explicit, and the use of the open-ended reports or essay assessments are useful tools for this.

Discussion and implications for good assessment practice

Based on all the above outcomes, some factors seem to have the potential to affect pupils' performance. Figure 1 reflects the observations from this study. There are, of course, other possible factors outside the scope of this work [see, for example, Taber (2003) for factors such as context, motive and purpose]. It also has to be noted that factors may well inter-relate with each other. For example, the test of convergence-divergence in this study correlated significantly, although at a low level, with the field dependent-independent test: those who are more divergent tend also to be those who are more field independent ($r = 0.19$, $p < 0.01$). This is not pursued further here.

Figure 1. Factors affecting psychological performance.



The study has raised many issues:

1. How do we decide about the validity of one format?
2. Is one format valid for one pupil but less so for another pupil?
3. Is there any format of assessment which is capable of being a more valid measure for most pupils?
4. What are we testing? Are we testing cognition or understanding of a particular discipline?
5. Do particular formats of assessment deskill the pupils?
6. Do particular formats of assessment frustrate pupils and therefore make them drop out of a subject or even out of study?

Assessment is a complex process. As Broadfoot and Black (2004) suggested, *"Educational assessment must be understood as a social practice, an art as much as a science, a humanistic project with all the challenges this implies and with all the potential scope for both good and ill in the business of education"*. In this situation Race (2003) suggests that *"Probably the best way to do our students justice is to use as wide as possible a mixture of assessment methods, … allowing students a range of processes through which to demonstrate their respective strengths and weaknesses"*. The issues are deeper as Thyne (1974), cited in Sanderson (1998), points out *"…it is axiomatic in the word of assessment that assessment tasks cannot measure 'cognition'…and the examiner must specify, at the outset, the performances to be accepted as evidence of Comprehending, or of Analysing, or whatever 'process' he wishes to assess, because examinations can measure only performance, not mental process"* Therefore, there are ethical issues about what formats of assessment need to be used to reflect pupils learning properly and, at the same time, to ensure a beneficial impact on teaching and learning practice (Gipps, 1994). It is important to be aware that testing may inadvertently favour a particular set of personal characteristics in the learner and thus test results may reflect possession of such characteristics as well as ability in the subject.

It follows that it is difficult to answer all the above questions fully, but it is possible to present the following guidelines for a good assessment practice:

1. Different formats may test different skills. Therefore we have to decide what we want to test. Do we want to test cognitive characteristics or to test knowledge and understanding?
2. It is impossible to design an assessment to suit each individual, and it is not wise to conduct all assessment by one method (e.g. objective testing, or open-ended).
3. The aims of the course must tie very closely to all aspects of assessment. If the aim of the course is the recognition of knowledge, then the test (probably multiple choice) must reflect that. If the aim is to transfer and apply knowledge, then problem solving, open-ended questions should be used. If the aim is to equip pupils with manual skills, then hands-on, or performance-based assessment should apply.
4. Teachers should be made aware that assessment is a skilful and demanding process, which has to be acquired through professional training. It is not just an appendage to teaching and learning.

Using a battery of different formats of assessment may help to ameliorate some of the problems outlined below:

- The use of objective tests may be of benefit to those pupils who, from nature or nurture, have not developed the cognitive processes needed to bring ideas together and present them clearly in open-ended formats. Also in a multicultural world, pupils may be assessed in a language which is not their mother tongue, and this means that may not perform well in some types of assessment where facility and subtlety of language is a limiting factor in their ability to produce an adequate answer.
- The use of open ended or problem-solving tasks helps more intellectually developed pupils to demonstrate their knowledge, their learning strategies, and to show their independence of thought. It is likely that objective tests constrain the more intellectually developed pupils and deprive them of opportunities to expand ideas.
- The use of oral examinations, open-ended assessments, essays, performance-based assessments, reports, portfolios and general alternative assessments encourages pupils to make their meaning explicit, to expand and enrich their vocabulary and their linguistic skills. Objective testing may deskill them linguistically.
- Assessment should not be punitive and judgemental but empowering and humane, especially at the school level when the pupils are forming their personality, building their self-esteem, and testing themselves in a different environment from their home. Assessment practice should support human needs rather than frustrate them. This means that assessment should encourage less successful pupils in their self-esteem, and help them to be less anxious about their performance, and therefore make them feel more comfortable in the school environment and stay longer in the school. After all we are human beings and we are entitled to make mistakes and to learn from them. Assessment for learning is an important issue (see Sorenson, 2000).

References

Al-Naeme F.F.A., (1988), *A study of some psychological factors affecting performance in chemistry at secondary and tertiary levels*, M.Sc. Thesis, University of Glasgow.

Al-Naeme F. and Johnstone A., (1991), Room for scientific thought?, *International Journal of Science Education*, **13**, 187-192.

Andersson B.R., (1990), Pupils' conceptions of matter and its transformation on classroom activities (age 12-16), *Studies in Science Education*, **18**, 53-85.

Atkins M., Beattie J. and Dockrell B., (1992), *Assessment issues in higher education,* Newcastle School of Education, University of Newcastle.

Bahar M., (1999). *Investigation of Biology student's cognitive structure through word association tests, mind maps and structural communication grids*, Ph.D. Thesis, University of Glasgow.

Bahar M. and Hansell M., (2000), The relationship between some psychological factors and their effects on the performance of grid questions and word association tests, *Educational Psychology: An International Journal of Experimental Educational Psychology*, **20**, 349-363.

Bennett S.N., (1973), Divergent thinking abilities–a validation study, *The British Journal of Educational Psychology*, **43**, 1-7.

Biggs J., (1996), Enhancing teaching through constructive alignment, *Higher Education*, **32**, 347-346.

Bodner G.M., (1991), I have found you an argument, *Journal of Chemical Education*, **68**, 385-388.

Boud D., (1995), *Assessment and learning: contradictory or complementary*? In P. Knight (Ed.), Assessment for Learning in Higher Education, London, Kogan Page Ltd.

Boud D., (2004), *Aligning assessment with long term learning needs,* Paper presented at the Enhancing the student experience in Scottish higher education. Constructive alignment of learning outcomes to assessment methods, Bridge of Allan, University of Stirling.

Broadfoot P. and Black P., (2004), Redefining assessment?, The first ten years of Assessment in Education. *Assessment in Education*, **11**, 7-26.

Byrne, M., Johnstone, A., and Pope, A., (1994), Reasoning in science: a language problem revealed? School *Science Review*, **75** (272), 103-107.

Child D. and Smithers A., (1973), An attempted validation of the Joyce-Hudson scale of convergence and divergence, *British Journal of Educational Psychology*, **43**, 57-61.

Crisp V. and Sweiry E., (2003, September), Paper presented at the British Research Association Annual Conference, www.ucles-red.cam.ac.uk, Edinburgh.

Danili, E. and Reid N., (2004), Some strategies to improve performance in school chemistry, based on two cognitive factors, *Research in Science & Technological Education*, **22**, 201-223.

Danili E. and Reid N., (2005), Assessment formats: do they make a difference? *Chemistry Education Research and Practice*, **6**, 204-212.

Davey B., (1990), Field dependence-independence and reading comprehension questions: task and reader interactions, *Contemporary of Experimental Education*, **52**, 119-206.

Entwistle A. and Entwistle N., (1992), Experiences of understanding in revising for degree examinations, *Learning and Instruction*, **2**, 1-22.

Fatt, J.P.T., (2000), Understanding the learning styles of students: implications for educators, *International Journal of Sociology and Social Policy*, **20**, 31-45.

Field T.W. and Poole M.E., (1970), Intellectual style and achievement of arts and science undergraduates, *British Journal of Educational Psychology*, **40**, 338-341.

Fryer M., (1996), *Creative teaching and learning*, London, Paul Chapman.

Gabel D., (1999), Improving teaching and learning through chemistry education research: a look to the future, *Journal of Chemical Education*, **76**, 548-553.

Gipps C., (1994), *Beyond Testing: towards a theory of educational assessment*, London, Routledge Falmer.

Hudson L., (1966), *Contrary Imaginations: a psychological study of the English schoolboy,* Great Britain, Penguin books Ltd.

Hudson L., (1968), *Frames of mind*, London, Methuen.

Johnstone A.H., (2003), *LTSN Physical sciences practice guide: effective practice in objective assessment.* Hull, LTSN.

Johnstone A.H. and Al-Naeme F.F., (1995), Filling a curriculum gap in chemistry, *International Journal of Science Education*, **17**, 219-232.

Johnstone A.H. and El-Banna H., (1986), Capacities, demands and processes: a predictive model for science education, *Education in Chemistry*, **23**, 80-84.

Kogan N., (1976), *Cognitive styles in infancy and early childhood*, New Jersey, Lawrence Erlbaum Associates, Inc.

Liodakis S., Gakis D., Theodoropoulos D., Theodoropoulos P. and Kallis A., (1999), *Chemistry A*' Lyceum Athens, Didaktika Biblion.

Lu C. and Suen H.K., (1995), Assessment approaches and cognitive styles, *Journal of Educational Measurement*, **32**, 1-17.

Messick, S., (1984), The psychology of educational measurement, *Journal of Educational Measurement*, **21**, 215-238.

Messick S., (1993), *The matter of style: manifestations of personality in cognition, learning, and teaching*, Princeton, NJ, Educational Testing Service.

Messick S., (1994), The matter of style: manifestations of personality in cognition, learning, and teaching. *Educational Psychologist*, **29**, 121-136.

Moore J., Stanitski C., Wood J., Kotz J. and Joesten M., (1998), *The chemical world concepts and application*s (2nd ed.), New York, Saunders College.

Nurrenbern S. and Pickering M., (1987), Concept learning versus problem solving: is there a difference? *Journal of Chemical Education*, **64**, 508-510.

Nuttall D.L., (1972), Convergent and divergent thinking, *The British Journal of Educational Psychology*, **43**, 1-7.

Oakhill J., (1988), *The development of children's reasoning ability: information processing approaches*, in K. Richardson, and Sheldon S. (Ed.), Cognitive development to adolescence, LEA/OU.

Orton A., (1992), *Learning mathematics issues, theory and classroom practice*, London, Cassell Education Series.

Osborne M., (2004) *Constructive alignment of assessment methods with learning outcomes*, An overview for the assessment workshop on 27th February 2004. Paper presented at the Enhancing the student experience in Scottish higher education, Constructive alignment of learning outcomes to assessment methods, Bridge of Allan.

Osborne R.J. and Cosgrove M.M., (1983), Children's conceptions of the changes of state of water, *Journal of Research in Science Teaching*, **20**, 825-838.

Pollitt A., Marriott C. and Ahmed A., (2000), *Language, contextual and cultural constraints on examination performance*, Paper presented at the International Association for Educational Assessment, Jerusalem, Israel, http://www.ucles-red.cam.ac.uk/conferencepapers.htm.

Race P., (2003), *Designing assessment to improve Physical Sciences learning – exams*, http://dbweb.liv.ac.uk/ltsnpsc/guides/designing_assessment/exams.htm [2003, 6/10/2003].

Renfrew R. and Conquest N., (1995), *Standard Grade Chemistry*, London, Hodder and Stoughton.

Riding R.J. and Cheema I., (1991), Cognitive styles: an overview and integration, *Educational Psychology*, **11**, 193-215.

Runco M.A., (1986), Divergent thinking and creative performance in gifted and non gifted children, *Educational and Psychological Measurement*, **46**, 375-383.

Sanderson P., (1998), *Cognitive process and discourse in the assessment of written text*, Paper presented at the European Conference for Educational Research, University of Ljubljana, Slovenia, http://www.ucles-red.cam.ac.uk/conferencepapers.htm.

Sawrey B.A., (1990), Concept learning problem solving: revisited, *Journal of Chemical Education*, **67**, 253-254.

Selepeng D. and Johnstone A.H., (2001), A language problem revisited, *Chemistry Education Research and Practice*, **2**, 19-29

Shepard L., (1992), *What policy makers who mandate tests should know about the new psychology of intellectual ability and learning*, In B. Gifford, and O'Connor, M. (Ed.), Changing Assessments: alternative views of aptitude, achievement and instruction, London, Kluwer Academic Publishers.

Sorenson P., (2000), *Drowning in numbers? The need for formative assessment*, in Sears, J and Sorenson, P, Issues in Science teaching, London, Routledge Falmer, pp. 123-132

Struyven K., Dochy F. and Janssens S. (2002), *Students' perceptions about assessment in higher education: a review*, Education-line, http://www.leeds.ac.uk/educol/documents/000025.htm.

Sutherland P., (1992), *Cognitive development today: Piaget and his critics*, London, Paul Chapman Publishing Ltd.

Taber K.S., (2003), Examining structure and context – questioning the nature and purpose of summative assessment, *School Science Review*, **85**, 35-41.

Terrell S.T., (2002), *The use of cognitive style as a predictor of membership in middle and High School Programs for the academically gifted*, Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, Louisiana.

Thyne R.L., (1974), *Principles of examining*, London, University of London Press.

Tinajero C. and Paramo F.M., (1998), Field dependence-independence cognitive style and academic achievement: a review of research and theory, *European Journal of Psychology of Education*, **13**, 227-251.

Tsaparlis G., (2005), Non-algorithmic quantitative problem solving in university physical chemistry: a correlation study of the role of selective cognitive factors, *Research in Science and Technological Education*, **23**, 125-148.

Vygotsky L., (1986), *Thought and language*, (A. Kozulin, trans.), Cambridge, MA, The MIT Press (Originally published in Russian in 1934).

Witkin H.A. and Asch S.E., (1948a), Studies in space orientation III. Perception of the upright in the absence of visual field, *Journal of Experimental Psychology*, **38**, 603-614.

Witkin H.A. and Asch S.E., (1948b), Studies in space orientation IV. Further experiments on perception of the upright with displaced visual fields, *Journal of Experimental Psychology*, **38**, 762-782

Witkin H.A., Dyk R.B., Paterson H.F., Goodenough D.R. and Karp S.A., (1974), *Psychological differentiation - studies of development*, New York, Wiley.

Witkin,H.A. and Goodenough D.R., (1981), *Cognitive styles: essence and origins*, New York, International University Press.

Witkin H.A., Goodenough D.R., Moore C.A. and Cox P.W., (1977), Field-dependent and field-independent cognitive styles and their educational implications, *Review of Educational Research*, **47**, 1-64.

Witkin H.A., Oltman P.K., Raskin E. and Karp S.A., (1971), *Embedded figures test, children's embedded figures test, group embedded figures test: Manual*, Palo Alto, CA, Consulting Psychologists Press.

Wittgenstein L., (1961), *Tractatus logico-philosophicus*, London, Routledge and Kegan Paul.

Wood R., (1986), *The agenda for educational measurement*, In D. Nuttall (Ed.), Assessing educational achievement, London, Falmer Press.

Appendix:

Structural Communication Grids

This is an objective form of assessment, which has distinct advantages over conventional multiple choice testing. The grid is a database on which a number of questions are based. The student is not told how many pieces of data are required to answer each question and so every piece has to be read and weighed up for its relevance to the question. The element of guessing is greatly reduced and students are seeking classification patterns rather than recognising one correct answer out of four or five. Successive questions will need some of the grid elements already used, and so answers cannot be correctly found merely by elimination. This process may involve subtlety of language interpretation, which may appeal more to the field independent and divergent students.

Each question is answered by an array of letters corresponding to the letters in the corner of each box in the grid, which has been selected by the pupil. Two examples are given below.

1. Each box in the grid below refers to a compound. Look at the boxes and answer the questions that follow. *(Boxes may be used as many times as you wish)*

A Sodium oxide Na_2O	*B* Lead nitrate $Pb(NO_3)_2$	*C* Phosphorus trioxide P_2O_3
D Barium iodide BaI_2	*E* Calcium oxide CaO	*F* Sodium nitrate $NaNO_3$
G Sulphur dioxide SO_2	*H* Magnesium sulphate $MgSO_4$	*I* Nitrogen dioxide NO_2

Select the box(es) that contain compounds which:

a) Produce alkaline solutions when dissolved in water..............................
b) Produce acidic solutions when dissolved in water ....
c) Cause acid rain………………………
d) Can react with the salt in box D and give a precipitation reaction……………..

2. Each box in the grid below refers to an element. Look at the boxes and answer the questions that follow. *(Boxes may be used as many times as you wish)*

A The element with the electron arrangement: 2.8.3	*B* Sodium	*C* Ar
D Magnesium	*E* The element that is a brown liquid at room temperature	*F* The element that has one electron in each atom
G The element with atomic number 19	*H* Chlorine	*I* Nitrogen

Select the box(es) that contain:

a) Elements that are gases in room temperature………………
b) Elements in the same group of the periodic table………………
c) Two elements that will combine to form an ionic compound with the formula X_3Y_2 ....
d) Elements that form a covalent compound with the element which is in box F ………………

Concept mapping in problem based learning: a cautionary tale

Alex H. Johnstone and Kevin H. Otis

Centre for Science Education, University of Glasgow, Glasgow G12 8QQ, UK
e-mail: alexjo@btinternet.com

Received 14 November 2005, accepted 7 February 2006

Abstract: Problem Based Learning (PBL) and Concept Mapping (CM) have parallel purposes, both based on a constructivist view of learning. In a Faculty of Medicine, PBL and CM have been applied together as the main learning modes. This provided an opportunity to test several hypotheses about the interaction of CM and PBL. Among them were: (i) Students using CMs for their study and revision would perform better on their assessment tasks, than those who did not. This was supported, but not strongly. (ii) Students with 'good' maps would do better than those with 'poor' maps. This was not supported. Many students with apparently 'poor' maps treated them as a sufficient set of keys to unlock very large databases and these students did well. Other students with 'poor' maps confessed to having a tenuous grip on their work and this accounted for the quality of their maps. This raises problems about using maps for assessment purposes. It may be that maps should be treated as very personal learning tools for the writer's eyes only, analogous to a personal diary which could be easily misunderstood by a reader. [*Chem. Educ. Res. Pract.*, 2006, **7** (2), 84-95]

Keywords: Concept mapping, problem based learning, assessment

Introduction

One important strand in the research at the Centre for Science Education has been to use an Information Processing Model to explore how students lay down information in Long Term Memory in a readily retrievable and useable form. Problem Based Learning is one method used to facilitate this.

The term Problem Based Learning (PBL) has recently been appearing in Science Education circles, in conferences and in the literature (Overton,, 2001; Belt et al., 2002). Even in casual conversation the title PBL is being applied to what used to be called tutorials, problem solving workshops and group exercises, and indeed they all involve some measure of PBL. Exercises in chemistry designed to promote discussion and group problem solving have been around for a long time (Percival, 1976, Johnstone, 1982, Wood, 1993, Schwartz et al., 1994). These have tended to be addenda to a more conventional mode of teaching through lectures.

However, the use of PBL as the main medium for learning in a discipline, or cluster of cognate disciplines, has been addressed by some of our medical colleagues. This is not a new phenomenon and medical schools across the world have embraced PBL as their medium for facilitating undergraduate learning (Barrows and Tamblyn, 1980; Schmidt, 1983; Barrows, 1986; General Medical Council, 1993).

The substance of this paper arises from experiences with PBL in the medical school in the University of Glasgow. Each university has its own version of PBL, but the common factors are to facilitate students' self-learning (as opposed to a didactic teaching and learning mode) and to encourage learning in context. In Glasgow the learning takes place through a series of 'real-life' scenarios in which the students learn the various strands of medicine (chemistry, biochemistry, anatomy, physiology, etc) in context. There is no formal teaching of these separate disciplines. Each topic, each concept, each discipline is met many times and from different angles and with increasing complexity. Gradually 'paper' scenarios give way to clinical incidents as the students progress. The overall aim is to facilitate integrated learning within concepts and between concepts. The earliest scenarios draw upon the students' pre-university knowledge and understanding providing points of attachment for new learning in Long Term Memory (LTM). As the knowledge and understanding base grows the frequent revisitation with more complex ideas allows a flexible network of ideas to be established. One of the basic tools in this learning process is the Concept Map. Students are encouraged to use the maps for three related purposes: to plan their work, to make explicit their own mental network and to prepare materials for later formal assessment. The maps are not themselves assessed, but they are aids for revision and inter-linking of knowledge. Concept mapping is associated with the work of Buzan (Buzan, 1974) and later with Novak and Gowin (Novak and Gowin, 1984). The suggested structure was to place a key concept (or node) in the middle of a page and surround this with closely related concepts (or nodes) linked by lines and some words to link them. The process was repeated with each of the new nodes and so on until a 'picture' of the knowledge and understanding structure was revealed. Such maps have been used in a range of learning processes such as planning, study, note taking, revision, problem solving and even for assessment.

Problem Based Learning and Concept Mapping can be seen as complementary (Table 1).

Table 1. The complementary nature of concept mapping and problem based learning.

	Concept Mapping	Problem Based Learning
Activate prior knowledge	Concept maps offer a method of visualising prior knowledge in the form of broad concepts and attaching the specifics of new information.	PBL activates prior knowledge by its application during the brainstorming session. This process also highlights gaps in knowledge.
Information supplied in the frame of a real problem	The use of a real life problem to form the first node of the map promotes the integration of academic and social data	PBL uses real life scenarios for two reasons: to tie new information to the likely cues for recall and to increase student interest by showing the relevance of new information to their work.
Elaboration on prior knowledge	Concept maps provide a structure on which new information may be assembled. The visualisation of this process allows its thoughtful integration into the students' expanding database.	The focus of PBL is the elaboration of prior knowledge. The students begin the process with what they already know. They then generate questions based on what they need to know to understand the scenario.

To complete this scene-setting introduction, we need to look at a typical week in the life of a PBL student.

Monday: PBL (2 hours). A group of eight students work with a facilitator. The first hour is devoted to discussion of the outcomes of the previous Thursday's tasks. The second hour is for the introduction and analysis of a new scenario and, during a 'brainstorming' session; students prepare a communal concept map, showing what they already know about the problems exposed in the scenario, and setting out what they need to know to tackle the problem fully.

Tuesday and **Wednesday:** Students work independently on the tasks arising from Monday, consulting books, computer data, research papers and any sources which might help them to meet the problem set out in the scenario. There are also laboratory sessions and workshops available, which are relevant to the problem.

Thursday: PBL (2 hours) as for Monday.

Friday and **weekend:** as for Tuesday and Wednesday.

Occasionally (not weekly) there is a lecture to integrate the work of the previous scenarios or to prepare the context for forthcoming scenarios. Almost half of the week is earmarked for private study, library work and report writing.

The facilitator (a member of staff drawn from medicine or science, trained to ask questions rather than provide answers) meets with the group of eight students. One student is appointed as chairperson and another as scribe. These ‘posts’ are regularly rotated round the group. Each student is presented with the scenario on about half a page of A4. This consists of a description of a situation, part of which might be familiar from previous work. The facilitator explains any unfamiliar terms and then the students, under the chairperson, have to decide on the main issues about which they need knowledge. The scribe records the ideas on a board in the form of a Concept Map to show linkages between the issues and to arrive at an agreed analysis of the problems. The facilitator can help with emphases on main concerns and deflect students from pursuing unprofitable lines. The students are then left with about six issues to be pursued. At the next PBL session students report back to communicate their findings, compare and resolve any conflicts and report on their information sources.

They are encouraged to produce their own Concept Map of the completed scenario and use it for further study, but this is an optional extra.

An evaluation study of the outcomes of the PBL in Glasgow was reported elsewhere (Mackenzie et al. 2003). This study ran alongside that of Mackenzie and her colleagues.

The Experiment

For administrative purposes the students in the first year PBL class were randomly assigned to one of three groups - A, B and C. Our experimental group (C) consisted of 82 students while the control was made up of groups A and B, totalling 160. The main thrust of the research was to observe students at work in a PBL situation in which they were encouraged to use concept mapping as a learning tool (Group C), but as this proceeded, several underlying behaviours appeared which will form the basis of this paper.

The initial research hypotheses were:

(i) Students who individually used concept mapping for study, planning and revision would have, on average, better scores in the battery of conventional assessment tests than those who used only the group concept map arising from each scenario.

(ii) As student knowledge and understanding increased, the maps for successive scenarios would contain more nodes and the inter-linkages would increase in number.

(iii) Students with the fullest (most complete) maps would be those who did best in their final assessment.

Group C students were given instruction in how to construct and use concept maps and they were strongly urged (but not obliged) to use them throughout their studies.

For each scenario, each student in group C was given four sheets stapled together. The top sheet contained the instructions for the experiment (Figure 1), while the other three blank sheets were non-carbon reproduction paper (NCR), a white, a yellow and a pink for ease of administration.

At the end of each scenario, before they had consulted books or any other resources, the students were asked to make a map of their present understanding of the scenario. This copied through all three sheets. The bottom pink sheet was detached, placed in an envelope and returned to the researchers.

Figure 1. Instructions for the conduct of the experiment to yield concept maps for each scenario

When you have finished answering your questions, put away your notes and books and follow these instructions.

- Turn the paper to landscape position and put your name and number in the top right hand corner.
- Reread the scenario.
- Use the name of the 'patient' as the central node.
- Expand the central node in concept map fashion.
- Expansion should be done one level at a time.
- Use all your previous knowledge (from primary school to last week) to explain the scenario.
- Try to limit the concept nodes to a couple of words and/or symbols; but this is your map, and so represent the ideas as you choose. If a list or a graph appeals to you, use it.
- When your map is complete, tear off the bottom pink page and return it in the envelope provided.

Now use your notes and books to make any additions or corrections to your map.

- When you have finished, tear off the bottom yellow page and return it in the envelope provided.

The remaining map on the white page should be kept and added to your notes for this scenario.

After the students had consulted literature and other sources, they were asked to elaborate or change their original map. This copied through both remaining sheets. The yellow sheet was detached and returned to the researchers. The students retained the white NCR sheet as part of their personal notes. They could elaborate this further as they progressed through other scenarios and noticed linkages between scenarios when the same concepts were revisited.

Results

1. Concept map returns

The total number of maps returned from ten scenarios was 546 pairs out of a possible 820 pairs. The commonest reason for the incomplete return was that some students used felt-tipped pens, which did not copy through to the lower sheets. Others used ball-point for some part of the maps and coloured felt-tipped pens for other parts. Students sent apologies, but did not make time to redraw the maps. Some students returned one sheet, but not the other, pleading forgetfulness. Only a very few students were completely uncooperative.

2. Analysis of maps

These were 'scored' by nodes, layers and linkages as suggested by Novak and Gowin (Novak and Gowin, 1984).

Nodes were concepts, ideas, sketches, graphs. The central node was the name of the main character in the scenario. The layer of nodes linked directly to the central node was layer 1. The next array of nodes linked directly to nodes in layer 1, were layer 2 nodes and so on. A 'rich' map would contain many nodes validly linked to nodes in inner layers.

There was a wide variation in the number of nodes appearing in maps based on the same scenario (table 2), and exemplified in the two maps shown in the Appendix. However the general pattern was that most maps had four level 1 nodes and each of these branched three times. Maps were four layers deep with specific detail around the edges. Nodes on the periphery had 2-3 times as many branches as interior nodes.

Table 2. Sample of Variation of the number of nodes for a given scenario

Student identifier	Scenario	Number of nodes
Student 202	4	83
Student 209	4	31
Student 234	4	46
Student 274	4	106
Student 278	4	79

However, there was a surprising finding. For about 25% of the students, the number of nodes in their maps was almost constant regardless of the changing complexity of the scenarios (table 3). The value of the 'constant' varied greatly between students. When students were interviewed near the end of the year, they were asked about this individual 'constant'. Typical replies were:

'I stopped when I had covered the topic.'

'I knew I couldn't learn everything, so I tried to be realistic and took on only what I thought I could remember.'

Table 3. Samples of nodes over eight scenarios of different complexity

Student	Sc1	Sc2	Sc 3	Sc 4	Sc 5	Sc 6	Sc 7	Sc 8
Student 210	47	44	46	47	46	–	44	45
Student 213	68	67	71	72	69	68	70	–
Student 234	45	45	46	45	43	44	46	44
Student 274	110	114	106	114	106	108	102	106
Student 279	59	58	58	59	61	60	59	58

As can be seen from table 3, the results from a random sample of this 25% of the students, based on eight very different scenarios, showed 'constants' varying from the mid-forties to over one hundred. If these 'constants' were a reflection of the students' coverage of the topic, then their perceptions of what was needed to comprehend the 'message' of each scenario differed considerably.

3. Map quality

Since concept maps are apparently gaining in popularity as potential assessment tools (Pendley et al., 1994, Regis et al., 1996, Edmundson, 2000; Stoddart et al., 2000, Nicoll et al., 2001) we took the opportunity to compare the analysis of the maps, using the scoring methods suggested by Novak and Gowin (Novak and Gowin, 1984) with the results of the normal battery of course assessments using multiple choice and extended questions. In fairness to these authors, they were not very enthusiastic about maps being used for assessment purposes and saw their function mainly as learning aids.

'Good' maps were characterised by many clear and validly linked nodes giving rise to clear layers, logically developed. 'Poor' maps were those with few nodes, weak linkages and indistinct layering.

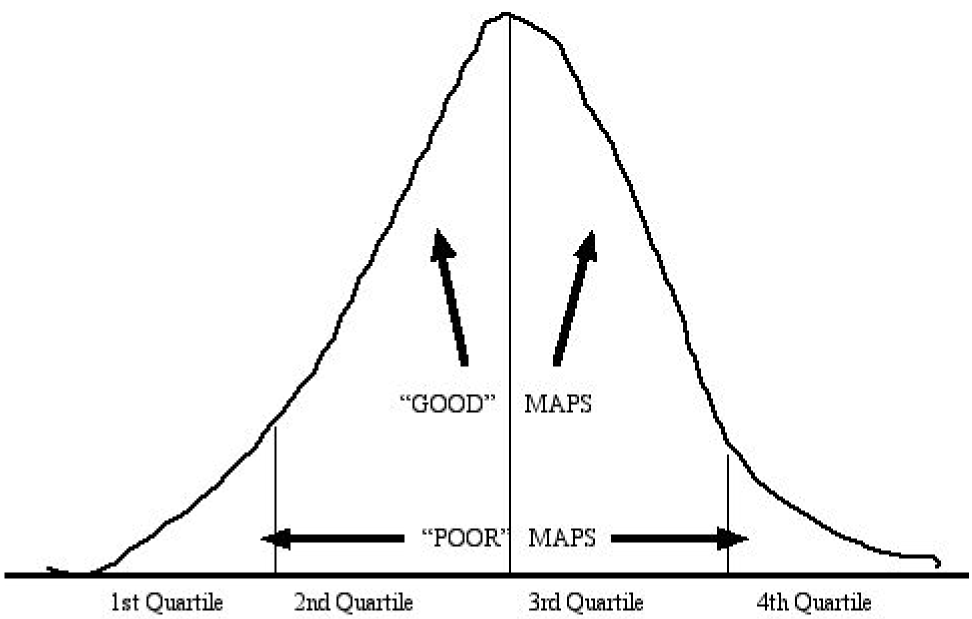
Such an analysis gave rise to three piles in our 546 maps. There were those which fell clearly into the two extremes set out above, but about one third came into intermediate positions such as those with many nodes, but set out in a chaotic fashion. It was difficult to do any rigorous correlation calculation between the map 'scores' and the normal assessment scores, but the following pattern emerged.

(i) Students, whose maps came into the 'good' and 'intermediate' categories, fell mainly into the middle two quartiles in the distribution of the normal assessment scores. This is not unexpected since the group represented 75% of the class and their distribution would approximate to 'normal'

(ii) However, students whose maps came into the 'poor' category, fell almost equally into the first and fourth quartiles of the normal assessment score distribution.

At first sight this seemed to be very strange, but interviews with twelve students in this category gave the explanation. Some of the students with 'poor' maps indicated that they did not have a good grasp of the scenarios and so their maps were sparse, ill-connected and unsatisfactory. Such students ended up in the lowest quartile. However, other students explained that they did not need a complex map, and that a simple, apparently 'poor' map was quite sufficient to act as a 'set of keys' to unlock their memory and reasoning store. Such students appeared in the highest quartile on the normal assessment battery of tests (Figure 2).

Figure 2. Distribution of maps across final exam results



Among the students who were interviewed were some whose maps were messy, ill-constructed, but node rich. On the scoring scheme, their maps fared badly and yet, when these students were asked to talk about the scenarios with the aid of their ‘poor’ maps, they could do so very cogently and well. What appeared to the observer to be a mess was a powerful tool for the constructor of the map.

These results have driven us to the conclusion that a map is a very personal thing, idiosyncratically constructed for the sole benefit of the individual student. It is likely to be misconstrued by an outsider; by anyone trying to assess what appears on paper.

The map is a memo; a set of keys; a collection of nodes which can spark off clusters of other nodes which need not be shown explicitly on paper. The best students may need only a skeleton map to unlock and activate a large body of knowledge and understanding. Does this imply that concept mapping is useless? On the contrary, it is a helpful learning tool which probably should not be pressed into service as an assessment tool.

When the results in the general course assessment were examined, group C students, who had used concept maps during the ten scenarios, scored (on average) about two percentage points higher than students in groups A and B who had acted as our control. (Table 4)

Table 4. Comparison of mean scores on conventional tests

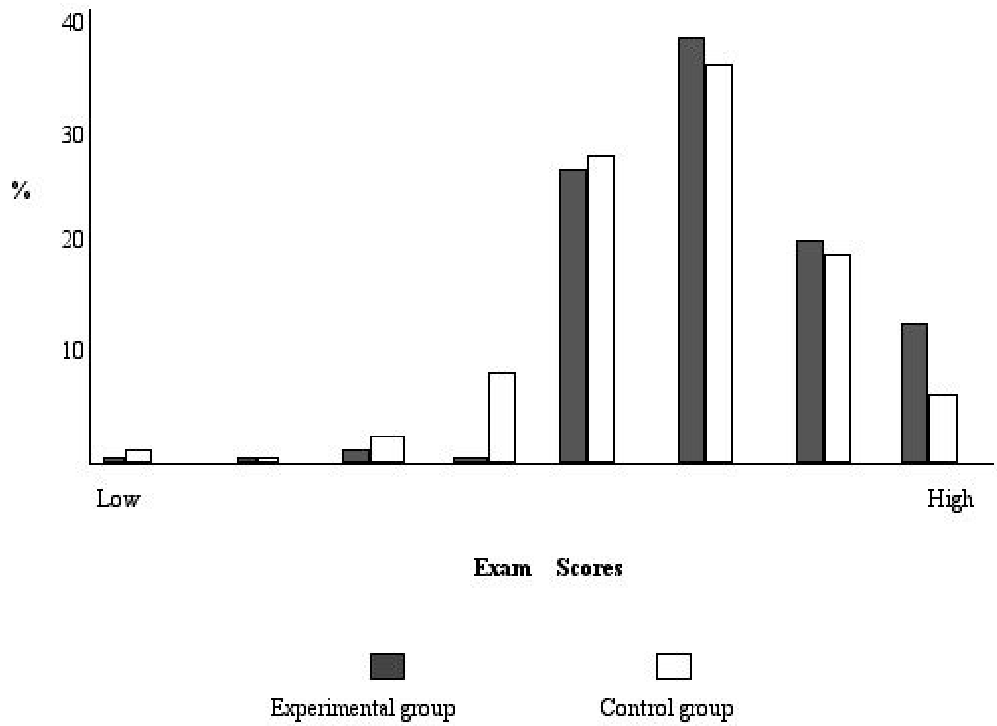
	N	Mean (%)	SD	SE
Control	160	66.3	5.5	0.4
Experimental	82	68.2	5.2	0.7

t = 2.7
Since t > 2.0, the means are significantly different at the 5% level

A final questionnaire administered to students in all three groups (90% return rate) showed that a significant proportion of the control group adopted concept mapping in their own studies of their own accord, having heard the recommendations from their colleagues in the experimental group.

It also turned out that some of the experimental group had given up mapping since they found it time consuming. The fact that the experimental and control groups were not wholly discrete would tend to reduce any differences in their mean scores, but the overall shift in the distribution in favour of the experimental group was significant at the 5% level, (student t = 2.7) (Figure 3).

Figure 3. Comparison of final exam results for experimental and control groups



Discussions and Conclusions

It is time to review our initial hypotheses.
(i) *Students who individually used concept mapping for planning, study and revision would have, on average, better scores in the battery of assessment tests than those who used only the group concept map arising from each scenario.*
This hypothesis has been supported by the results. The differences have been significant but not dramatic.
(ii) *As student knowledge and understanding increased, the maps for successively more complex scenarios would contain more nodes and inter-linkages.*
This was generally the case, but about one quarter of the students tended to produce maps with an almost constant number of nodes and linkages regardless of the nature of the scenarios.
(iii) *Students with the fullest (most complete) maps would be those who did best in their final assessment.*
This was not well borne out experimentally. Most students with 'good' maps fell into the quartiles immediately on either side of the mean, but were not well represented the highest and lowest quartiles.
Students with 'poor' maps were distributed, almost equally, between the lowest and highest quartiles.

Recommendations for Practice

1. From many studies (Pendley et al., 1994, Lee and Fensham, 1996, Herron, 1996, Sanger and Greenbowe, 1999) it has been shown that learning in the sciences has a tendency to be linear and boxed. Working from conventional notes and textbooks, topic A precedes

topic B followed by topic C. Each topic is in a sealed box with little or no interaction between them. Having taught bioinorganic chemistry, one of us (AHJ) has had student complaints about lectures in which thermodynamics, organic ligands and inorganic ions came together. The opening of the concept boxes and the mixing of their contents was upsetting and even painful for the students. Mother Nature seems to do this mixing with ease! Also students who could solve simultaneous equations in maths could not do so in physical chemistry.

Any device which aims to break down the linearity and compartmentalisation in our students' learning is to be encouraged. We believe that concept mapping has a contribution to make to this.

2. The skills of concept mapping have to be taught, but thereafter maps are the private, idiosyncratic aids and products of our students' learning.

3. It is very doubtful if concept maps should be used by a 'non-author' for assessment purposes. To be useful learning tools, concept maps may be unsuitable reading for another observer. They may be analogous to a personal diary with its abbreviations, allusions, selection of the important and memory jogging; rich in meaning for the writer, but easily misconstrued by a reader.

References

Barrows H.S. and Tamblyn R.M., (1980), *Problem based learning, an approach to medical education,* New York, Springer Press.

Barrows H.S., (1986), A taxonomy of problem based learning methods, *Medical education,* **20**, 481-486.

Belt S.T., Evans T., McCreedy T., Overton T.L. and Summerfield S., (2002), A problem based learning approach to analytical and applied chemistry, *University Chemistry Education,* **6,** 65-72.

Buzan T., (1974), *Use your head*, London, BBC Books.

Edmondson K.M., (2000), Assessing science understanding through concept maps. In Mintzes J.J., Wandersee J.H. and Novak J.D., *Assessing science understanding: a human constructivist view,* San Diego, Academic Press.

General Medical Council, (1993), *Recommendations on undergraduate medical education. Tomorrow's Doctor*, London, Safair Print Services Ltd.

Herron J.D., (1996), *The chemistry classroom: formulas for successful teaching,* Washington, D.C, American Chemical Society.

Johnstone A.H., (1982) (Ed.), Set of interactive discussion units: '*Nitrogen for the nineties*'; '*To market a drug*'; '*Zinc and you*'; '*Titanium the space age metal*'; '*Organic liquids containing oxygen*', London, Royal Society of Chemistry.

Lee K.W.L. and Fensham P.F., (1996), A general strategy for solving high school electrochemistry problems, *International Journal of Science Education*, **18**, 543-555.

Mackenzie A.M., Johnstone A.H. and Brown R.I.F., (2003), Learning from problem based learning, *University Chemistry Education*, **7**, 13-26.

Nicoll G., Francisco J. and Nakhleh M., (2001), A three-tier system for assessing concept map links: a methodological study, *International Journal of Science Education,* **23**, 863-875.

Novak J.D. and Gowin D.B., (1984), *Learning how to learn,* New York, Cambridge Press.

Overton T.L., (2001), *Problem based learning: an introduction. Learning and Teaching Support Network Physical Sciences Primer 4*, Hull, Learning and Teaching Support Network.

Pendley B.D., Bretz R.L. and Novak J.D., (1994), Concept maps as a tool to assess learning in chemistry, *Journal of Chemical Education,* **71**, 9-17.

Percival F., (1976), *A study of teaching methods in tertiary level chemical education,* Ph.D.Thesis, University of Glasgow.

Regis A., Albertazzi P. and Roletto E., (1996), Concept maps in chemistry education, *Journal of Chemical Education*, **73**, 1084-1088.

Sanger M.J. and Greenbowe T.J., (1999), An analysis of college chemistry textbooks as sources of misconceptions and errors in electrochemistry, *Journal of Chemical Education,* **76**, 853-860.

Schmidt H.G., (1983), Problem based learning: rationale and description, *Medical Education*, **17**, 11-16.

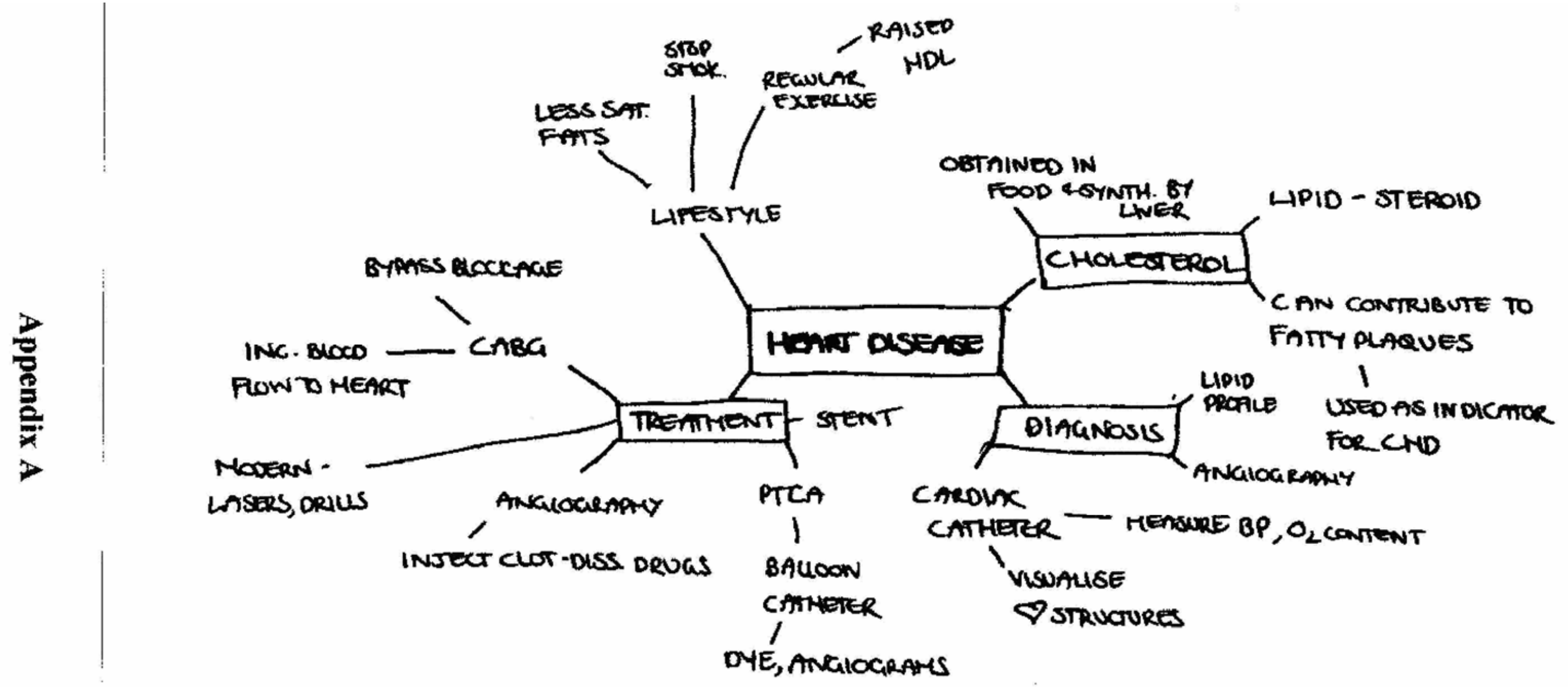
Schwartz A.T., Bunce D.M., Silberman R.G., Stanitski C.L., Stratton W.J. and Zipp A.P., (1994), *Chemistry in Context,* American Chemical Society, Indiana, W.C. Brown.

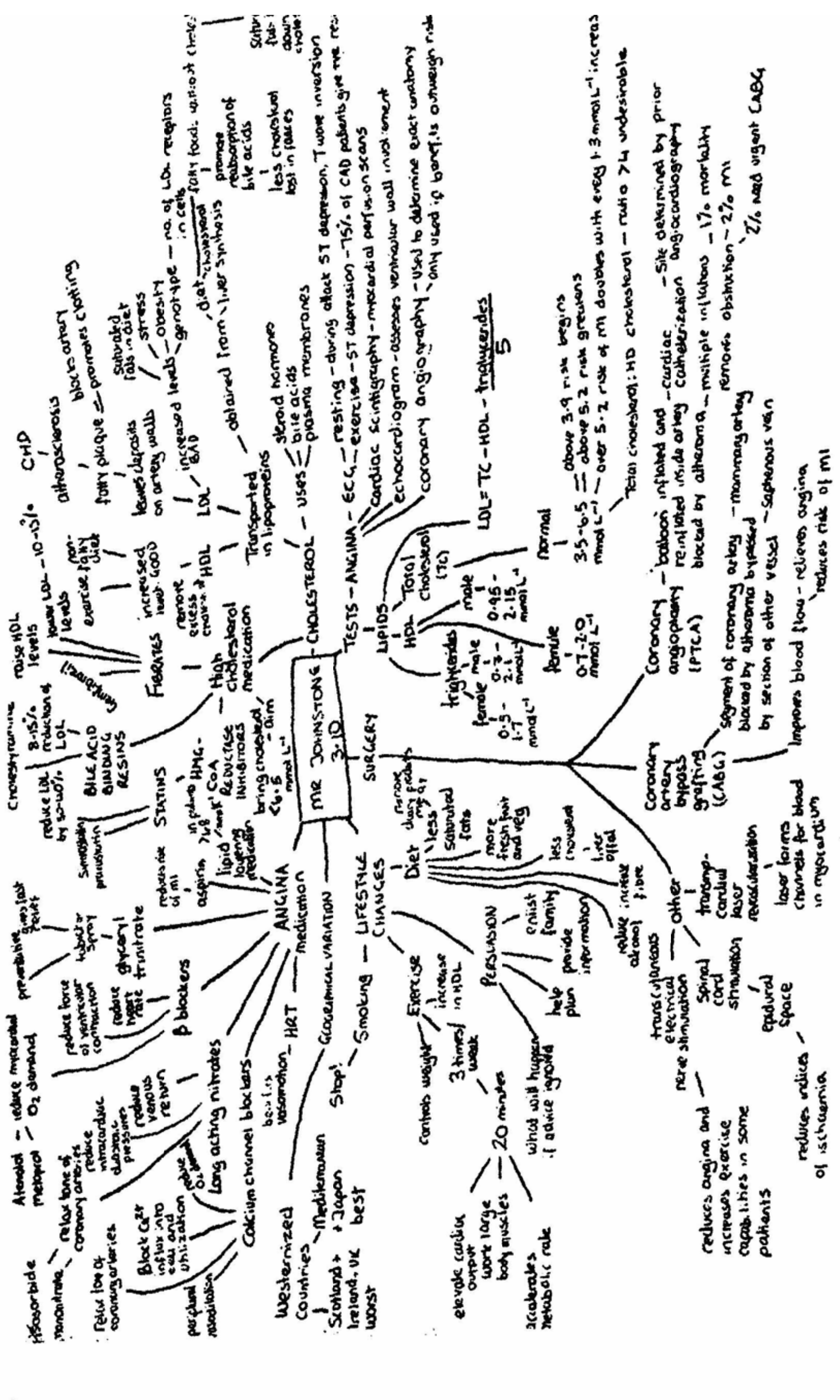
Stoddart T., Abrams R., Gasper E. and Canaday D., (2000), Concept maps as assessment in science inquiry learning – a report of methodology, *International Journal of Science Education*, **22**, 1221-1246.

Wood C.A., (1993), *Creative problem solving in chemistry,* London, Royal Society of Chemistry.

Appendixes

Appendix A




MR JOHNSTONE 3.10
ANGINA medication
Long acting nitrates
Calcium channel blockers
β blockers
glyceryl trinitrate
aspirin
STATINS
BILE ACID BINDING RESINS
FIBRATES
High cholesterol medication
CHOLESTEROL
CHD
atherosclerosis
fatty plaque
Transported in lipoproteins
TESTS – ANGINA – ECG
cardiac scintigraphy – myocardial perfusion scans
echocardiogram – assesses ventricular wall involvement
coronary angiography
LIPIDS
HDL
Total Cholesterol (TC)
triglycerides
LDL = TC – HDL – triglycerides/5
normal
SURGERY
Coronary artery bypass grafting (CABG)
Coronary angioplasty (PTCA)
other
LIFESTYLE CHANGES
Smoking
Exercise
Diet
PERSUASION
GEOGRAPHICAL VARIATION
Westernized Countries
Mediterranean + Japan best
Scotland + Ireland, UK worst
HRT
20 minutes
Improves blood flow – relieves angina
reduces risk of MI

Appendix B

Practice

The development of creative problem solving in chemistry

Colin Wood

Centre for Science Education, University of Glasgow [1]
e-mail: colin@meldr.co.uk

Received 19 January 2006, accepted 13 February 2006

Abstract: The object of this paper is to show how research has gone hand in hand with development to produce materials for teaching and learning which take problem solving well above the level of algorithmic manipulation and into the realm of creativity. To combat the common feature of school science and chemistry in particular, that problems have a unique solution, and to give students an appreciation of real science, problems of an open kind have been developed, which encourage the ingenuity, and idiosyncratic contribution of the students involved for their solution. In these, often there is no correct answer, only a 'best' answer, and there may be a variety of possible methods of finding it. Criteria for success are very different from the more common type of closed problems, but are much more difficult to define. The second strand of the work described here aims to make students aware of the benefits of group work and discussion by setting them objectives which are more likely to be achieved by groups of students working cooperatively together. Discussion is seen as of two types: task orientated discussion (how to solve the problem), and reflective discussion to consider in what ways their group was successful, and why; and in what ways it was less successful, and why. This helps students to realise that their success as a group is more than the sum of their individual contributions. [*Chem. Educ. Res. Pract*., 2006, **7** (2), 96-113]

Keywords: problem solving, creativity, discussion, teamwork

Introduction

In 1962, a document containing the so-called 'alternative' Chemistry syllabuses for use in Scottish schools was published (SED, 1962), covering O-grade (ordinary grade, age 15/16) and H-grade (higher grade, age 16/17) Alex Johnstone had made a major contribution to the development of these syllabuses, which were 'alternative' in the sense that teachers could adopt them or remain with the traditional syllabuses and examinations if they wished. However, they caught the mood of the moment, and most schools adopted them quickly, allowing the traditional syllabuses to be phased out. Later, Johnstone developed the groundbreaking Certificate of Sixth Year Studies (CSYS) syllabus for those post-H-grade students who opted to remain at school for a further year of secondary education. This was designed with deliberately low factual content to allow teachers to focus on wider skills of thinking (SEB, 1990). About one third of the marks for the assessment of this course were allocated to a project and practical work. These projects were organised and marked in a way similar to that of a final year university degree project, with an external assessor coming in to examine the students both in the practical work and the projects.

[1] Present address: 10 Melfort Drive, Stirling, FK7 0BD, UK

Teachers found that this increased students' maturity and self-confidence, easing the transition to university or work.

Research on problem solving and on why students find it such a problem was ongoing at the Centre, with publications starting in 1979. A series of teaching units was developed by Reid for school pupils (described in Johnstone and Reid, 1979, 1981b). Some of these were in problem format while many involved the use of small groups. They were designed, among other things, to take chemistry out of the classroom in a realistic way, and to foster thinking skills. However, the main aim was the development of positive attitudes to chemistry and to show pupils that chemistry requires critical thinking and risk taking (Johnstone and Reid, 1981a, 1981b).

Further work resulted in the publication, by the Royal Society of Chemistry, of five booklets written by school teachers (Johnstone, 1982), designed as an integral part of the Certificate of Sixth Year Studies course for ages 17/18. These were largely based round industrial processes (fertiliser production, drug design, zinc uses and extraction, titanium uses and extraction and properties of oxygen-containing organic liquids). Similar research and development produced problem-solving materials for students taking standard grade (replacing O-grade; GCSE equivalent) at 15/16 years of age. These problems were entirely based upon laboratory work. They were designed to complement normal laboratory work and most of them were intended to occupy the last twenty minutes of a laboratory period. The clues for their solution were to be found in the preceding conventional laboratory. The effect was to supplement and reinforce the main points of the conventional work and to give the students room to exercise their own ideas. The output from this research was the book '*Practical Problem Solving for Standard Grade Chemistry*' which was circulated to all schools in Scotland (Hadden, 1989).

All this represented the starting point of this project. What was new was using group dynamics as a means of encouraging creativity. The aims agreed with the Royal Society of Chemistry reflected this. These were refined as work progressed and eventually became the following.

- To improve students' ability to work and communicate with others, and to develop an awareness and control over their own thinking processes;
- to give students the opportunity to develop their problem solving skills;
- to give students the opportunity to be creative and use divergent or lateral thinking;
- to show students that science is more than 'getting the right answer', and that it can involve using one's judgement, being creative and using lateral or divergent thinking.

We intended to meet these aims by

- presenting problems with a variety of possible solutions;
- getting the young people to work in groups to discuss their solutions critically and present their agreed solution to their peer group in the form of a mini lecture.

Thirty problems (seventeen of them lab based) were designed to meet these criteria. They were trialled by 16/17-year-old school students in Central Scotland. Before the problems were attempted, it was found to be important to explain to students the purposes of these new learning experiences. They were told that these materials were designed to enhance their ability to solve new or unusual problems and to give them opportunities to communicate and co-operate with others in a small group.

Appendix 1 outlines how the problems were to be used while Appendix 2 gives two examples of the problems.

Problem solving

A problem is sometimes defined as a situation where at present the answer or goal is not known. For the problems normally encountered in educational situations, the way to that goal is not known initially. This, of course, assumes that a goal has been specified: ‘find the volume of’; ‘how many ...?’; ‘show that ...’. Some problems of this kind can be reduced to routines or algorithms in which students can be drilled, meaning that they can eventually become exercises where one set of numbers is substituted for another. For the student who can recall the method of tackling the problem, no problem remains.

However, in real life, problems may have quite different shapes which are not amenable to algorithmic manipulation and which demand a degree of lateral thinking for their solution. Criteria for success are very different from the more common closed problems. Teachers and students are so used to getting ‘the correct answer’ in academic problems, that they can be misled into thinking that in science, there is a unique answer for every problem. There is a danger of cultivating within our students an ‘all is known’ view of science: a discipline to which students can make no personal contribution. All too often examination questions are of this type, reinforcing even more this distorted view of science.

This paper describes problems of an open kind, which invite the ingenuity, and idiosyncratic contribution of students to their solution. Sometimes there is no correct answer, only a ‘best’ answer judged against the criteria set by the students themselves. Students may end up with their best answer being within an order of magnitude, realising this is only a useful ‘guesstimate’. Sometimes, there is a correct answer but a variety of possible methods. Success in others may lie with economy of time, cost or scale.

The problems attempt to foster process skills such as data seeking and selection, choice of method, balance of criteria, and awareness of error as well as discussion and presentation skills. Underpinning our approach is the Information Processing Model of learning. According to this model, the process of problem solving will more or less cease if too many ‘chunks’ of information are competing for the students’ attention. Chunks are pieces of information coming into the mind’s ‘working space’. However, if there are too many of these, there is no space left to process the information and the problem solving process grinds to a halt. An upper limit of about five chunks is normal for comfortable manipulation.

The Information Processing Model as outlined by A H Johnstone in Wood (1993) acted as a theoretical basis for developing the strategies recommended to students attempting to find solutions to the problems. For example:

- The problems were designed to minimise overload by the way they were structured, and by using discussion to help students break the problem into processable chunks. Consider ‘Hair’ as an example (Appendix 2). Students are asked to estimate the approximate rate of growth of human hair, and to use this figure to estimate the number of amino acid molecules which are incorporated in the growing hair every second.
- The initial instruction to the students points out that for many problems the search for an exact answer may be a distraction, and tells them that they will have to make approximations on their way to devising a solution. This initial discussion can focus on familiar words in the problem like ‘hair’, ‘rate of growth’, ‘units’, helping to define the problem and providing the important framework (‘perception’) of the problem. The importance of the language used in framing the problems cannot be overemphasised (Johnstone and Cassels, 1980; Talbi, 1990).

- Such discussion often initiated key ideas, which led to a solution of the problem. In this case, it was words like ‘haircut’, ‘hairdresser’ or phrases like ‘about once a month’, ‘about 1 cm cut off’ that led to a solution.
- Students are encouraged to ‘brainstorm’ the problem thus giving access to the long-term memories of several people and their associated mental patterns and chains (Kempa and Nicolls, 1981; Reid and Yang, 2002).

Types of problem

Originally we saw problems as of four different types -

Type	Data	Goal
1	specified	specified
2	insufficient	specified
3	specified	not specified
4	not specified	specified

As work progressed, this classification evolved and expanded into the following (Wood and Sleet, 1993):

Type	**Data**	**Methods**	**Outcomes/goals**	**Skills bonus**
1	Given	Familiar	Given	Recall of algorithm
2	Given	Unfamiliar	Given	Looking for parallels to known methods
3	Incomplete	Familiar	Given	Analysis of problems to decide what further data are required
4	Incomplete	Unfamiliar	Given	Weighing up possible methods and deciding on data required
5	Given	Familiar	Open	Decision about appropriate goals; exploration of knowledge networks
6	Given	Unfamiliar	Open	Decision about goals and choice of appropriate methods; exploration of knowledge and technique networks
7	Incomplete	Familiar	Open	Once goals have been specified by the student, they are seen to be incomplete
8	Incomplete	Unfamiliar	Open	Suggestions of goals and methods to get there

The ‘normal’ problems usually encountered are of types 1 and 2. Types 3 to 8 represent much more the skills required for investigative work. Goals may not be absolutely clear at the beginning, and methods may be unfamiliar. Data may be incomplete and the student will then have to generate them from experimental work and/or by literature search.

Problems of type 3 would involve the student saying ‘*If you want me to do this, I shall need the following ...*’.

Type 4 could be exemplified by *“How many copper atoms are there in a 2p coin?”*. This would involve a reasoning chain like: *“If I knew the mass of the coin and if I assumed that it was pure copper and if I had the atomic mass of copper and Avogadro’s number, I could get an answer, but it would only be approximate. But if I have no balance and only a ruler, I could get* *its volume (approximately) and if I knew the density of copper, I could get a good estimate”*. This is very different reasoning from that in types 1 and 2.

Type 5 is much more open and is left to the judgement of the student as to what would constitute a reasonable answer. For example: ‘Given the formula $[Co(NH_3)_4Cl_2]$ deduce from it as much as you can’. This could yield a range of responses, including the oxidation state of the cobalt ion and its ‘d’-electron configuration, the name of the complex, its percentage composition, its isomers, its likely reactions, and so on.

Type 7 would require the students to specify the goals, but to achieve these, extra data would have to be requested.

Type 8 might be of the kind where the students were given a substance and asked to suggest uses for it. The students would have to ask for, or find out experimentally, its properties before deciding upon uses.

Type 6 would be similar to type 8 but the given substance would be familiar to the students.

Types 1 and 2 problems have their place, but students are short-changed if they are not also exposed to the other types. In all the types, thinking skills are exercised, and flexibility, branching, open-mindedness and creativity are encouraged. Students who perceive chemistry as a developing, changing and intellectual adventure with room for individual thought and contribution will be those who are potentially the creative thinkers we hope to encourage.

All thirty problems in the book were trialled in schools at senior levels (17-19). The behaviour of the students was observed carefully. All the problems stimulated discussion and/or promoted the skills previously mentioned. Researchers involved in the trials were pleasantly surprised by the ability of students to argue and defend their presentations to their peer group and to the researchers who were total strangers to them.

The aim was to foster the development of the following process skills:

- Lateral thinking
- Balance of criteria
- Choice of method
- Data seeking and selection
- Awareness of error
- Discussion and presentation skills.

Discussion groups

“*The importance of problem solving for individual pupils has been widely emphasised ... However, it is also the case that scientific investigation, as it is practised in the wider scientific community ... involves teams or groups of workers. Co-operation and teamwork as well as effective leadership are likely to be qualities important among scientists. If we are to prepare the next generation of scientists for this, it will be necessary for schools to teach with these points in mind and not to leave it to chance.*” (Gayford, 1992).

A recent article “*Graduates unfit for work, say top firms’* reports that Britain’s biggest companies are finding that ‘many [UK] graduates lack the basic skills needed for employment. They *“are being let down by the [university] system … last year 598 positions were left unfilled as a third of employers said that they could not find candidates of sufficient quality*” (The Times, 2006).

Managers cited a series of shortcomings in potential recruits. These include:

- Too much time spent working on degrees and not enough joining clubs and societies, where students might work in teams;
- Not enough time spent on giving presentations in tutorials, leaving new graduates unable to communicate ideas in the workplace.

Discussion and presentation skills are important in all walks of life. They are not only important in their own right but are highly regarded by employers, and indeed in the community at large. Unfortunately, most students have to have these consciously taught, not necessarily by formal means, but by example, gentle encouragement and opportunity. Initially, students have to be encouraged to share tasks in groups, to pool their gathered information, to brainstorm, to listen, to criticise positively and to accept criticism as being constructive. Once they see the utility of group methods, even the most diffident of students make contributions in a peer group. There is a fine judgement needed on the part of the teacher as to when to intervene and when to remain silent, when to encourage and when to act as a consultant.

The importance of discussion

Discussion in schools

The traditional emphasis on content arising from examination constraints perhaps can mean that pupils tend to see science as an impersonal body of knowledge to be learned rather than as a co-operative activity where their knowledge, their skills and their value judgements are important and relevant. Such a content-driven approach to chemistry learning may well lead students to miss out on the excitement of chemistry and they may well see their studies merely as an experience to be endured on their way to some other area of study.

Primary pupils are accustomed to group work and to the associated project planning and allocation of different tasks within the group. The excitement and enthusiasm with which they approach the problems can be a great motivating force. This can continue at secondary stages. Extension material can build upon the success of group work in primary schools and can encourage groups of pupils to devise solutions to novel problems. They can see their learning as something generated, at least in part, from within the group rather than being imposed from outside.

Discussion in higher education

Group projects and presentations are becoming more common (Bell et al., 2002) in university and college courses, with the marks obtained contributing to the final degree classification. In addition to assessment by the responsible lecturer, students can assess one another using specific criteria set by the department concerned.

Discussion can help cultivate critical thinking skills

Engaging in discussion with others develops critical thinking. This can show students that they do have the ability, and even the obligation, to think critically. Resolving a difficulty, understanding a difficult topic, a flash of insight: these can and should be satisfying experiences. (Byrne and Johnstone, 1986a, 1986b, 1986c, 1987).

Discussion can help clarify ideas

Discussion can promote active learning because, in effective discussion, students can:

- express in their own words what they have learned;
- think critically about new knowledge and ideas;
- justify any decisions they make which are based upon the new knowledge and ideas;
- be prepared to admit to uncertainty and lack of knowledge;
- be prepared to admit to incorrect thinking and the superiority of another's ideas without loss of self-esteem.

Discussion groups can improve self-esteem

Schools are trying to get pupils to accept more responsibility and become 'managers of their own learning'. They are encouraged to become aware of their own learning processes, and as far as possible to be in control of them. At best this can be successful in helping young people to grow in self-esteem, to feel good about themselves as learners, and therefore to become more successful as learners (Higher Still Subject Guide: Chemistry, 1997).

Discussion can utilise group dynamics

Several minds working jointly on a problem can produce solutions that each on its own could not manage. As the group as a whole comes to realise this, it sees the necessity for encouraging the shy and the uncertain, stimulating the lazy and restraining the over-talkative.

Types of discussion

Two types of discussion are involved, both of which need a supportive and non-adversarial atmosphere. The first focuses upon the task itself and the second upon the factors, which contribute to effective group discussion and teamwork.

1. Task related discussion

a. Discussion before and during the problem

The problems are designed to increase problem-solving skills and to encourage co-operative working in small groups. In some, students can treat the initial group discussion as a 'brainstorming' session where all ideas are encouraged no matter how trivial or unrealistic they may appear at first sight. Divergent thinking is encouraged, as an apparently unrealistic idea from one group member can initiate a train of thought in other group members that could lead to the problem being solved.

Spending sufficient time on this initial discussion can save much time and effort during the problem itself.

b. Group presentations

In order to allow sufficient time for discussion, it can be better to give one problem to two groups of students rather than increasing the number of problems beyond the two or three recommended (Appendix 1). Each group makes its presentation to the whole class, and then discusses its findings with the class. This worked best when the teacher exercised informal control as required.

c. The teacher's role

The teacher's role during the problems is to provide sufficient support to ensure that the problem solving is at least partially successful. Students should be allowed to make mistakes and explore blind alleys provided that this leaves sufficient time for them to accomplish enough to feel some degree of success. They should be continually reviewing their solutions. Questions such as, *"What are you doing?" "Why are you doing it?" "How will it help you solve the problem?"* help focus the students' minds and can improve their problem solving performance dramatically. They also need encouragement: *"That sounds like a good idea, but have you thought about ... ?"*

2. Reflective discussion

Once the group have completed the problem, each student and the group as a whole are encouraged to be introspective and asked to consider -

- in what ways their group was successful, and why;
- in what ways it was less successful, and why.

This discussion follows on from the presentation and the ensuing discussion about the problem and the chemistry involved. Students are encouraged to look at what they achieved, and how they achieved their solution. This could give the participants valuable insight into themselves and could allow them to start to identify some of the elements that make for successful teamwork.

The aim of this discussion includes encouraging the ability to

- give and accept constructive criticism;
- value one another's contributions;
- be introspective and to consider one's own feelings and those of other group members.

Suggested approach to reflective discussion

People are often reticent when talking about factors that affect them on a personal level, and are more likely to enter into this type of discussion when in small groups. It was suggested that only a few minutes be spent on this type of discussion when it is first attempted, starting in a low-key way by asking a small number of questions and building upon success each time it is tried. The teacher should move the discussion imperceptibly on from task-related discussion to reflective discussion. Not too much should be expected at the first or even the second attempt.

As the students gain experience of such discussion, they start to realise that they are developing the skills and the abilities needed for successful teamwork, and that these are useful when devising solutions to the task in hand. The teacher can assist this process by asking questions regularly, usually resisting the temptation to offer answers. When asked a question by a group member the teacher should encourage another member of the group to answer it. The reply "*that's an interesting question*" is much used by leaders of team building courses!

Note that the teacher's role here is quite different from that during the problem-solving itself where he/she should give enough help to ensure success in tackling the problem concerned.

The following was suggested to encourage reflective discussion;

a. The group sits roughly in a circle with the teacher acting as informal chairman.
b. Reflective discussion can be started by asking a simple question such as *"How did you feel about the discussion?"* perhaps following this with *"Did it go well?"*, and *"Why do you think it went well?"*

An alternative is to start from the teacher's observations of the various groups' earlier discussions where the teacher has some idea of the more successful and less successful aspects. One of the successes could be used as a starting point, for example by asking direct questions such as *"John, would you agree that your group was successful in your discussion of ... ?", "Why do you think it was successful?", "What made it successful?", "What do the rest of the group think?"*

Once discussion is under way the teacher should ensure that some time is spent focusing on the feelings of the individuals in the group. This could be done by asking who felt that their contribution was never made (why?), or didn't have much heed paid to it, and what their feelings were at the time. An alternative is to choose a point in the earlier discussion where someone's point got lost in the discussion and asking something like *"Sheila, did you feel that your point about ... got sufficient attention?"*

c. At the end of the discussion, it is worth re-emphasising to the group the importance of teamwork and the interpersonal skills needed to work effectively in groups, and how important these are in industry, commerce and research. Most decisions are made, if not by committees, on the recommendation of committees.

Reflective discussion is designed to make students aware of their own feelings and emotions and of those of their peer group, and to make them realise how important these are in group work. This can boost students' self-confidence and increase their ability to contribute positively to group work.

Teamwork

The ability of each individual to function as a team member is a necessary but not sufficient condition for successful teamwork. Members of a particular team must be able to work effectively together. While there are certain attributes for all team members that increase the chances of this happening, members of successful teams usually have both different skills and different personalities.

1. Different skills

Different skills/ abilities should be represented in a team or in a committee, for example:

- a rugby team, where different balances of skills are required for forwards and for backs;
- a committee to plan a new chemical plant should include engineers, chemists, accountants and lawyers.

Students with a background in biology can have a different perspective from those who have studied only physical sciences.

2. Different personalities

Successful teams are likely to include a mixture of personality types. There are many ways of categorising people. One system (discussed by Johnstone and Al-Naeme, 1991, 1995) that works well in the science domain categorises by motivational trait -

- The ACHIEVER - sets out to do well, to be top of the class, is competitive, prefers expository methods of teaching and learning, dislikes being held back by slower students;
- The CONSCIENTIOUS - wishes to do well, content to please teachers or parents or whoever by working conscientiously within clear cut guide-lines, gains satisfaction from work duly performed;

- The CURIOUS - keeps asking questions, enjoys exploring, is the divergent thinker, the creative person;
- The SOCIAL - enjoys learning cooperatively and is not competitive.

There is overlap between the traits, but many of us show predominantly one or two of them. Most successful teams have an appropriate mix of these personality types, with some individuals moving flexibly between roles.

Discussion is a central part of the problem solving process and this suits the ‘social’ learner. The ‘conscientious’ are often hesitant when tacking the problems initially because no secure framework of thought is provided and because of the variety of possible methods and/or answers. Research findings (Johnstone, 1998, 2001) confirm the observations made during the trials that even this latter group is stimulated by the problems once the initial uncertainty has passed.

Analyses of school science courses (Johnstone and Al-Naeme, 1995) show that they generally provide opportunities for the achiever, for the conscientious and for the social learner, but provide little opportunity for the ‘curious’ pupil. Some ‘curious’ pupils consider that there is little in science for them because they perceive it as only about ‘getting the right answer’. Yet the ‘curious’ person can be the creative thinker, playing a key role in teams designing new products or devising new solutions to problems. Many of these pupils may gravitate to non-science courses in senior school and/or college/university; this is great loss to the scientific community.

Continuing work

This work on creative problem solving is not the end of the story. The Certificate of Sixth Year Studies Chemistry was the second most popular subject in the Scottish Sixth Year (17-18), coming second only to mathematics. While the projects were never intended to be original research (although that was done on occasion), it had to be *original to the pupil*; but the large number of pupils presented a challenge to teachers to devise valid projects for them all. This presented a clear logistical problem for teachers to find new problems for the projects and so an innovative series of ‘Starter Projects’ were devised by final year university chemistry students working with Johnstone. These were published as small booklets in two sets (Johnstone, 1993, 1995). They provided sufficient information for the pupils to get started on a project. If the pupil was making little progress, then further information was provided. As a last resort, fairly detailed instructions were available. The assessment of these projects allowed credit for initiative and resourcefulness, but students who needed a lot of support were not given this credit.

In 1997, the Certificate of Sixth Year Studies was replaced by Advanced Higher (Higher Still Development Unit, 1997). This course was more prescriptive than its predecessor, with much more content (similar to that in English A level) and unfortunately, this left less time for investigative and problem solving work. However, the project in a shortened form was retained with fewer marks being available for it, and the use of an external examiner, although retained initially, has been regrettably discontinued.

When these latest reforms were first proposed, The Royal Society of Chemistry (Scottish Education Committee) was concerned that the project work and its associated learning experiences would be marginalised or disappear, and it persuaded the Higher Still Development Unit to commission (along with The Royal Society of Chemistry) the two original authors of the Starter Projects along with the author of this paper to update the previous starters and write new ones. This built upon the knowledge and experience that had been gained when trialling the problems on creative problem solving described above. These were published jointly by the

Royal Society of Chemistry and Learning and Teaching Scotland (Support Materials, 2000, 2002) and were circulated to all Scottish secondary schools.

These materials were written specifically for the Scottish system, but demand from outside Scotland showed that they were of more general interest and usefulness. A new introduction was written and the material republished by The Royal Society of Chemistry (Education Division) (Robertson, Gray and Wood, 2001).

This paper has attempted to show how research and development can go hand in hand, with support from a Learned Society (RSC) and the chemical industry, to make a difference to the way in which chemistry is taught. “Education in chemistry” can be successfully linked with “Education through chemistry” to transform the subject from a largely passive experience to be accepted, into an exciting, demanding and participative experience to be enjoyed.

Epilogue

One Head of Chemistry wrote to the author after trialling some of the projects with his 17-year old pupils – *“I spoke to the pupils informally after the trial and everyone remarked how much they enjoyed the double period and how quickly the time had passed. Another comment was that you really had to know what you were talking about before you tried to explain it to a group. They all felt that it was much more demanding than simply answering an exam paper question. I thought that the material was excellent and intend to use it next year after the appropriate section of work.*

”There is no doubt that it is useful revision but that is secondary. The group discussion, team work and experience of presenting conclusions/ findings to their peers were invaluable.”

Acknowledgements

This work was carried out during a year’s secondment from school teaching funded by the Marjorie Cutter Scholarship from the Royal Society of Chemistry. The result of the year’s work was a book published in 1993 called *Creative Problem Solving in Chemistry* (Wood, 1993). The author would like to acknowledge the contribution by Ray Sleet of The University of Technology, Sydney, Australia who spent time working at the Centre of Science Education at Glasgow during the author’s secondment. Discussion with him helped with some of the individual projects and with the strategies developed.

References

Bell S.T., Evans T., McCredy T. and Overton T.L., (2002), A problem based approach to analytical and applied chemistry, *University Chemistry Education*, **6**, 65-72.

Byrne M.S. and Johnstone A.H., (1986a), *Powder v liquid: an exercise in the marketing of detergents*, London, Royal Society of Chemistry.

Byrne M.S. and Johnstone A.H., (1986b), *The Ardil project*, London, Royal Society of Chemistry.

Byrne M.S. and Johnstone A.H., (1986c*), So you think you know about fibres?* London, Royal Society of Chemistry.

Byrne M.S. and Johnstone A.H., (1987), Can critical mindedness be taught?, *Education in Chemistry*, **24**, 75-77.

Gayford G., (1992) Group behaviour in problem solving, *International Journal of Science Education*, **14**, 41-49.

Hadden R.A. and Johnstone A.H., (1989), *Practical problem solving for standard grade chemistry*, Glasgow, Scottish Consultative Committee on the Curriculum.

Johnstone A.H. and Reid N., (1979), Bringing chemical industry into the classroom, *Chemistry and Industry*, **4**, 122-123.

Johnstone A.H. and Cassels J.R.T., (1980), *Words that matter in science*, London, Royal Society of Chemistry.

Johnstone A.H. and Reid N., (1981a), Towards a model for attitude change, *European Journal of Science Education*, **3**, 205-212.

Johnstone A.H. and Reid N., (1981b), Interactive teaching materials in science, *South Australia Science Teachers Journal*, **812**, 4-15, (A series of 30 page interactive units).

Johnstone A.H., (1982), (editor of the following five booklets, written by various school teachers), all published by the RSC London. The titles were: *Nitrogen for the nineties*, *To market a drug*, *Zinc and you*, *Titanium the space age metal*, *Organic liquids containing oxygen.*

Johnstone A.H. and Al-Naeme F., (1991), Room for scientific thought?, *International Journal of Science Education*, **13**, 187-192.

Johnstone A.H., (1993), *Starter projects for sixth year studies*, Centre for Science Education, Royal Society of Chemistry and GlaxoWellcome, Glasgow.

Johnstone A.H. and Al-Naeme F., (1995), Filling a curriculum gap in Chemistry, *International Journal of Science Education*, **17**, 219-232.

Johnstone A H (1995), *More starter projects for sixth year studies*, Centre for Science Education, Royal Society of Chemistry and GlaxoWellcome, Glasgow.

Johnstone A.H., (1998), 'Learning through problem solving', in '*Problem solving in Analytical Chemistry'* (ed. A. Heaton), London, Royal Society of Chemistry.

Johnstone A.H., (2001), Can problem solving be taught? (Nyholm Symposium), *University Chemistry Education*, **5**, 12-18.

Kempa, R.F. and Nicholls C.E., (1983), Problem-solving ability and cognitive structure: an exploratory investigation, *European Journal of Science Education*, **5**,171-184.

Meyer R.E., (1992), *Thinking, problem solving, cognition*, Oxford, W H Freeman.

Reid N. and Yang M-J., (2002), Open-ended problem solving in school chemistry: a preliminary investigation, *International Journal of Science. Education*, **24**, 1313-1332.

Robertson K.J., Gray C. and Wood C.A., (2001), *Starter investigations for advanced higher chemistry*, London, Royal Society of Chemistry.

Talbi M., (1990), An information processing approach to the investigation of mathematical problem solving at secondary and university levels, *PhD thesis*, University of Glasgow.

SEB (Scottish Examination Board), Certificate of Sixth Year Studies, Dalkeith, 1990.

SED (Scottish Education Department), Circular 512, Edinburgh, 1962. Scottish support materials - *Higher Still subject guide Chemistry: sections 'rationale' and 'core skills'*, (1997), Edinburgh, Learning and Teaching Scotland (LTS).

Higher Still Support Materials 7134 *Chemistry starter investigations (Advanced Higher)*, Higher Still Development Unit and Royal Society of Chemistry, Edinburgh, 2000.

National Qualifications Curriculum Support Chemistry Starter Investigations Supplement, (Advanced Higher), Learning and Teaching Scotland, Dundee 2002. (These materials can be downloaded from www.ltscotland.org.uk).

The Times (London) 'Graduates unfit for work, say top firms', 7th February 2006, www.timesonline.co.uk/article/0,,3561-2028167,00.html reporting on a survey carried out by AGR (Association of Graduate Recruiters, www.agr.org.uk).

Wood C.A., (1993), *Creative problem solving in chemistry*, London, Royal Society of Chemistry.

Appendix 1: Using the problems

General

The problems are designed to improve problem solving skills, and to encourage cooperative working in small groups. A brief introduction by the teacher is required so that students know the objectives of the learning experience and realise what is expected of them.

Format of the problems

For flexibility of use each problem is set out in two parts, as a statement of the problem and as a teacher's guide, including a suggested approach.

Each of the problems starts on a new page. When the problems were trialled the statement of the problem and the suggested approach were photocopied and given to students; this approach worked well and is recommended to teachers. However, teachers are encouraged to modify the problems to suit their particular circumstances. It is unlikely that the suggested approaches will suit all circumstances, and the extended discussion may not always be necessary.

The practical work in some of the practical problems is small and could be carried out in advance if access to a laboratory is not possible for the main problem solving session; some of the non-practical problems have extensions which do involve laboratory work, and could be incorporated into the main problem solving session.

Objectives

During the trials it was found to be most important to explain the objectives of the learning experience to the students before they attempted the problems. These include:

- to improve students‘ ability to work and communicate with others and to develop an awareness and control over their own thinking processes;
- to give students the opportunity to develop their problem solving skills;
- to show students that science is more than ‘getting the right answer’, and that it can involve using one’s judgement, being creative and using lateral or divergent thinking.

Students should know that these skills are not only important in themselves but are highly regarded by employers and the community.

The chemistry underpinning many of the problems is relevant to senior school courses and can be readily integrated into class teaching. In addition, during trialling it was found that group discussion to find solutions, and the preparation of a presentation to the peer group, helped to clarify concepts, deepen understanding and correct misconceptions.

Student groupings and duration of problems

During the trialling of these problems, discussions within groups of students were found to be effective. The size of the class will determine how many groups there are and how many different problems are in use at one time. For classes of about twenty, two or perhaps three different problems might be used; whilst for classes of twelve or less, one or two different problems would be more suitable. To allow more time for discussion it may be better to give one problem to two different groups of students rather than increasing the number of problems beyond the number recommended.

Each group should make a presentation to the rest of the class and then answer questions and discuss their findings with the rest of the class.

It is not possible to give definitive time allocations for each problem, as this depends upon many factors including:

- the problem itself;
- the number of students;
- the ability of the students and their previous experience of these problems; and

- the importance the teacher attaches to different parts of the problem (tackling the problem, its presentation and the discussion following the presentation). In practice, the reverse may apply – the time available dictates the number of stages tackled.

When the lesson lasts for about 100 minutes or more it is possible to complete some of the problems along with the associated presentation and discussion in one day, otherwise the presentation and/or discussion will have to be left until later. The presentation and associated discussion is sufficiently important to warrant the extra time.

An overhead projector with an acetate transparency per group speeds up the presentations because each group can prepare their visual aids simultaneously; otherwise several groups can be competing for the use of the blackboard at the same time.

The teacher's role

The teacher's role during the problems is to provide sufficient support to ensure that the students are at least partially successful. Students should be allowed to make mistakes and explore blind alleys provided that this does not take up so much time that it jeopardises success. They should be made to think, but given sufficient help, not by giving them answers but by encouragement ('that sounds like a good idea'), by asking appropriate questions or by pointing them in the direction of a suitable textbook. In some of the problems, an old inorganic textbook like Parkes and Mellor proved more useful than a more up-to-date book like Cotton and Wilkinson.

After presentation of the problems the group is asked to consider how well they accomplished the task, and discussion becomes more reflective.

Notes

1. 'Communication and interpersonal skills ... are considered to need the greatest development at the commencement of ... employment' from CASupdate no. 1, published by the Careers and Appointments Service of the University of Technology, Sydney, Australia.

2. Some projects have been used for examination revision. Here too, much useful learning took place during discussion. On occasion a misconception would persist through to the presentation, but further discussion, with teacher input as required, would remedy the problem.

Appendix 2: Two sample problems

Problem 1: Hair

The problem as issued to students

i) Estimate the approximate rate of growth of human hair in ms^{-1}.

ii) Use this figure to estimate the number of amino acid molecules which are incorporated in a growing hair every second.

Many real-life problems do not need exact answers. Often an exact numerical answer would be misleading.

In science, it is useful to be able to make estimates. This can be important when checking whether an answer makes sense, for example in deciding whether an injection of 10 cm^3 of a drug solution is a reasonable dose.

In this pencil and paper exercise you will need to make approximations to get answers.

You should refer to any sources of information that you think might help such as your notebooks textbooks and data books. Ask for assistance if you get stuck.

Teachers' guide

Introduction

Teachers who have not used the problems before should read the section 'Using the problems' before starting.

Prior knowledge

Concept of bond length and knowledge of amino acids. A detailed knowledge is unnecessary as students are encouraged to consult textbooks and data books during the exercise.

Resources

Scientific calculators, tables of bond lengths and access to suitable textbooks in which to find the structures of amino acids should be made available for reference.

Group size: 3 or 4

Possible methods

Question (i)

One method is: 'I go to the barber every six weeks and he cuts off about 3 cm of hair' or 'I get my hair dyed: after about 2 weeks there is about 1 cm undyed.'

Therefore 1 cm (0.01 m) of hair grows in 2 x 7 x 24 x 60 x 60 seconds

or

1 m of hair grows in 2 x 7 x 24 x 60 x 60 x 100 seconds

This equals a growth rate of 8.3 x 10^{-9} ms^{-1}. Given the accuracy of the data, the growth rate of hair can be taken as about 10^{-8} ms^{-1}. The approximations have made mean that the answer is no more than a general indication of the value.

Question (ii)

The size of an amino acid molecule has to be estimated in order to calculate the approximate number of amino acids joining a hair per second. Most amino acid molecules have similar structures H_2N-CHX-COOH where X is different for each amino acid.

To get an idea of the length of a molecule, the bond lengths in the chain can be added together: this approximates to 0.5 x 10^{-9} m (0.5 nm). The number of molecules joining each chain per second is calculated by dividing this figure into the growth rate figure calculated in question (i):

10^{-8} ms^{-1}/ 0.5 x 10^{-9} m = 20 molecules joining each chain per second.

This is only the first stage: the number of chains growing along each hair have to be taken into account.

A reasonable guess for the cross section of a hair is 0.01 mm, 10^{-5} m. This equates to an area of about 10^{-11} m^2.

A reasonable guess for the cross section of a typical amino acid molecule is that it is about the same as the length calculated above of 0.5 x 10^{-9} m. This indicates a cross section area of about 10^{-19} m^2.

Therefore there are about 108 amino acid chains per hair. Using this figure, along with the 20 molecules joining each chain per second gives an estimate of 2 x 10^9 molecules joining each hair per second.

Suggested approach

During trialling the following instructions were given to students and proved to be effective:

1. Working as a group, discuss the first question and try the calculation.
2. Discussion can play a vital role in working out possible solutions to such problems. Several minds working on a problem together can stimulate the production of ideas that one on its own could not manage. About 10 minutes should be spent on this initially, with further discussion as required.
3. Because exact answers are not possible, you will have to use your judgement to make sensible estimates.
4. Write up what you did in note form. You should explain how you decided upon the estimates you had to make.
5. Repeat steps 1 and 2 for the second question.
6. Working as a group, prepare a short (ca 5-minute maximum) presentation to give to the rest of the class. If possible all group members should take part: any method of presentation, such as a blackboard, overhead projector, etc., can be used.

Outline the problem, describe what you did and explain your solution and the approximations you had to make. After the presentation, be prepared to accept and answer questions and to discuss what you did with the rest of the class.

Possible extensions

1. Estimate the number of hairs on the human body and calculate the total number of amino acids used each day to keep human hair growing.
2. Estimate the total food intake for an adult and the proportion of protein in this, and hence calculate the approximate protein intake per day. Consider what happens to all of this. The possibilities include: hair salon floors; shaving in the morning if you are male; hair and skin debris down the bath plug; not to forget old skin, the main constituent of house dust and excretion as urea.

Note

Students can be reminded that making hair is only one of many uses of amino acids in our bodies - all the fleshy parts of our bodies are replaced at the molecular level every four years or so. Thus we are literally new people every four years or so because every soft part is demolished and rebuilt at the molecular level, usually with little or no change of shape.

Problem 2: Making copper

The problem as issued to students

Make some copper metal starting from copper(II) nitrate crystals. Normal laboratory apparatus and any other chemicals can be used provided that they do not contain any copper.

At least three different methods should be tested. Which method works best? You should refer to any sources of information that you think might help, such as your notebooks, textbooks and data books. Ask for assistance if you get stuck.

Safety

Normal safety procedures when handling chemicals should be adhered to and eye protection worn. There are particular hazards that could arise depending upon how you tackle the problem.

You must get your method checked for safety before starting on the practical work.

Teachers' guide

Introduction

Teachers who have not used the problems before should read the section 'Using the problems' before starting.

Prior knowledge

Reactivity series, interconversion of salts and the effect of heat on nitrates. A detailed knowledge is unnecessary as students are encouraged to consult textbooks and data books during the exercise.

Resources

The reactions should be carried out in test tubes or - in the case of the electrolysis - in small beakers.

Solid copper(II) nitrate should be provided at the start.

Students can request apparatus and chemicals during the practical session: these should be issued if they are safe to use. In particular, electrolysis apparatus will probably be required but should not be on view.

Group size: 2/3

Risk assessment

A risk assessment must be carried out.

Special safety requirements

If the nitrate is heated, note that poisonous nitrogen dioxide gas is produced; also the water released from the hydrated salt could run down and crack the hot glass.

There is a hazard of explosion with any gas reduction apparatus.

Possible methods

1. Dissolve the salt in water and electrolyse the solution.
2. Dissolve the salt in water and add a more reactive metal (preferably powdered) such as zinc or magnesium. It is usually necessary to stir the mixture with dilute acid after the reaction to remove excess displacing metal.
3. Heat the nitrate to get the oxide and reduce it with carbon or natural gas.
4. Dissolve the nitrate in water; add aqueous sodium carbonate; filter off and dry the copper(II) carbonate; heat to produce copper(II) oxide; and finally reduce this with carbon or natural gas.

If students do not think of the short methods (1 and 2), see if they are suggested in the final discussion. If they aren't, join in the discussion yourself!

NB. The thermite reaction is a possible method, but the inherent hazards mean that students must not carry it out. Teachers may wish to do so at their own discretion.

Suggested approach

During trialling the following instructions were given to students and proved to be effective:

1. Working as a group, discuss the problem and list as many different methods as you can. Ask for help if you can't think of at least three.
2. Discuss the advantages and disadvantages of each method.
3. Discussion can play a vital part in working out solutions to open-ended problems. Several minds working together on a problem can stimulate ideas that one on its own could not manage. About 10 minutes should be spent on this initially, coming together for further discussion as required.
4. Each person in the group should select a method and write it up in note form.
5. Get your method checked for safety and then carry out the practical work to find out how well it works.
6. Write a brief account of what you did; include discussions of the advantages and disadvantages of your method.
7. Working as a group, try and decide on the best method from all those tried.
8. Working as a group, prepare a short (ca 5-minute maximum) presentation to give to the rest of the class. If possible all group members should take part: any method of presentation such as a blackboard, overhead projector, etc, can be used. Outline the problem, describe what you did and explain your conclusion.

After the presentation, be prepared to accept and answer questions and to discuss what you did with the rest of tine class.

Note

Zinc powder can be a particular problem – some samples react well, others explosively, perhaps because different samples are oxidised to different extents.

Explicit teaching of problem categorisation and a preliminary study of its effect on student performance – the case of problems in colligative properties of ideal solutions

Vasileios Zikovelis and Georgios Tsaparlis*
Department of Chemistry, University of Ioannina, Greece
e-mail: gtseper@cc.uoi.gr

Received 19 December 2005, accepted 20 January 2006

Abstract. Research in science education suggests that the way knowledge is organised affects its availability for problem solving. It also constitutes the important difference between experts and novices. The retrieval of learned schemata from long-term memory is facilitated by categorisation of the problem. In this work, first we suggest a categorisation scheme for problems in the special topic of colligative properties of ideal solutions. Secondly, we report on the results when such a scheme was taught to an experimental group (n = 41) of eleventh-grade upper secondary students (age 16-17) in Greece. The group was reduced by considering only students who demonstrated knowledge of the categorisation scheme (n = 24), and was compared with a control group (n = 26) who were taught in the traditional manner. The experimental group showed a superior performance, but it was not statistically significant. Next, we divided the students into high-, intermediate- and low-achievement subgroups on the basis of their performance in two nationally-examined chemistry courses. No differentiation was found for the students of high and low performance. However, in the intermediate subgroup, students of the experimental group outperformed those of the control group. Because of the limitations (mainly small samples) in the research study, the findings should be treated as preliminary ones. The implications for teaching are discussed. **[***Chem. Educ. Res. Pract.,* 2006**,** 7 (2), 114-130**]**

Keywords: problem solving, problem categorisation, explicit teaching of problem categorisation, problems on colligative properties of ideal solutions

Methods and procedures of problem solving

Problem solving is considered an integral component in students' education in science. In school science, problems usually involve for their answer the use of mathematical relationships and the calculation of a numerical result. Such problems contain numerical data, and also the values of physical and chemical quantities and/or constants. Problems on colligative properties of ideal solutions are of this type; they constitute part of the upper secondary school general chemistry curriculum, and are the subject of this paper.

Of great interest is Johnstone's (1993) thorough classification of problems into various types. He has also emphasised a fundamental distinction, that between problems and exercises. A real/novel problem requires that the solver must be able to use what has been termed as higher-order cognitive skills (HOCS) (Zoller, 1993; Zoller and Tsaparlis, 1997; Tsaparlis and Zoller, 2003). As a rule, extensive practice in problems in a particular area can turn problems into exercises. For example, many problems in science can be solved by the application of well-defined procedures (*algorithms*) (Bodner, 1987) that can turn the problems into algorithmic exercises. Problems on colligative properties of solutions are of this type and particularly subject to this transformation. According to the Johnstone classification, in such problems data are given, the method can become familiar, and the outcomes/goals are given. If this is the case, then the problems can be solved by recall and application of algorithms. In this work, we study ways to make effective students' familiarisation with the method of solution. For this purpose, we concentrate our attention on problem categorisation.

In addition to the method, there are a number of other requirements that play important roles in the successful integration of the solution process, leading to the correct result. Relevant to this work are the differences between expert and novice problem solvers and especially: (a) the comprehensive and complete scheme of the experts, in contrast to the sketchy one of the novices; and (b) the extra step of the qualitative analysis taken by the experts before they move into detailed and quantitative means of solution (Simon and Simon, 1978, Larkin, 1980; Reif, 1981, 1983).

In this work, first, we present the theoretical basis for problem categorisation, and then we suggest such a scheme for problems in colligative properties. The scheme was the result of a systematic process. Finally, we report on a *preliminary* investigation of the effect of *teaching explicitly* (that is, offering students an explicit method for choosing between problem solutions) this problem categorisation on upper secondary students' problem-solving ability. The question behind the study is *whether the scheme is helpful to learners in terms of improving overall task performance.* The pressures on this kind of research (particularly in getting suitable samples) sometimes lead to non-ideal structures. This is the reason for the preliminary character of the investigation.

Rationale: Knowledge organisation and problem categorisation

Researchers in cognitive psychology stress that the organisation of knowledge affects the availability and the retrieval process of conceptual schemata during problem solving (Sternberg, 1981; De Jong and Ferguson-Hessler, 1986), and that it is important that a plethora of connections exist among various concepts (Chi and Koeske, 1983). Johnstone (1991) has pointed out that information that is well organised and connected in long-term memory is more easily recalled than specific information, which lacks organisation and connections. Tsaparlis (1994, 1998) has confirmed this in the case of solving very simple (two-step) organic chemical synthesis problems.

Problem categorisation in a specific topic involves organisation of the various problems into a number of categories and subcategories. A solution of a problem begins with a brief analysis of the problem so that it can be categorised mentally. A crucial fact is that experts and novices organise their knowledge differently in a domain. Experts' knowledge networks are characterised by a multilevel structure that is constructed in such a way that they are easily retrieved from long-term memory (Wilson, 1994). Each problem category functions as a cognitive structure that, at least for experienced solvers, includes possible solution methods.

Bunce and Heikkinen (1986) have proposed the explicit method of problem solving (EMPS) which aims to teach novice students the problem-solving analysis procedures used by experts. According to Reif (1981), this analysis helps students *encode* the pertinent information of the problem, which is a major difference in the problem-solving behaviour of experts and novices. *Encoding* is defined by Sternberg (1981) as the identification of each term in the problem, and retrieval from long-term memory of the attributes of these terms that are thought to be relevant to the solution of the problem.

According to Bunce, Gabel, and Samuel (1991), an important part of the encoding process is problem categorisation. If students cannot correctly categorise a problem, they will not be able to retrieve the relevant information from long-term memory. A subsequent step in EMPS leads students to relate the encoded parts of the problem in a schematic diagram of the solution path. After such an analysis, students can use mathematics to reach an algebraic solution and eventually a numerical answer. The above authors Bunce et al. further examined the effectiveness of EMPS and reported that specific instruction in problem categorisation techniques improved achievement scores for combination problems, requiring more than one chemical concept in their solutions, but not for single-concept problems. On the other hand, such training alone was found insufficient to lead to conceptual understanding.

The mechanisms of implicit learning, explored in the cognitive psychology literature, may offer another explanation whereby students become increasingly sensitive to patterns (regularities) in the problem conditions, without ever gaining conscious awareness of the links between conditions and solution process. It is important to note, however, that the research on implicit learning has only investigated simple perceptual rules in experimental psychology settings, as opposed to the learning that takes place in formal education [see Pacton et al. (2001) for a review]. The literature on second language acquisition also investigates the relative advantage of explicitly teaching grammar rules. Truscott (1998) reviewed research in this area and found little evidence to suggest that explicit grammar teaching is helpful in gaining language competence in naturalistic language situations.

Finally, of special importance in problem solving is the *logical structure* of a problem. Niaz (1995) specified this structure by the number of *operative schemata* entering the problem. According to Piaget (Inhelder and Piaget, 1958), a *schema* is an internal structure or representation, while the ways we manipulate schemata are called *operations.* The unit on colligative properties starts with a background sub-unit (A) on the vapour pressure of a single liquid substance (not a colligative property in itself), and is followed by the three sub-units (B-D) on colligative properties as follows:

A. Vapour pressure of a single liquid substance.
B. Vapour pressure depression of an ideal solution of a non-electrolyte.
C. Boiling-point elevation and freezing-point depression of an ideal solution of a non-electrolyte.
D. Osmosis and osmotic pressure of a solution.

The suggested categorisation classifies each problem on the basis of the type of colligative property involved. Each sub-unit involves one category in the scheme. The four categories, plus the sub-categories (see below) in each, constitute the logical structure of these problems.

Method

The problem categorisation scheme

At the start of this research project, the junior investigator worked out an initial categorisation scheme for each sub-unit under study (see above). Work started out with the sub-unit on the vapour pressure of a single liquid substance. Following that, the two investigators discussed and devised a scheme for the categorisation. This process was repeated several times until an agreement was reached. The same process was followed for each of the three colligative properties. The whole process for working out the categorisation schemes was very detailed and systematic, and took several months.

As mentioned above, the suggested categorisation classifies each problem on the basis of the type of colligative property involved, and involves four main groups (categories) of problems. One additional group was employed that contains problems combining two or more colligative properties. The suggested categorisation is shown in outline in Tables 1 and 2, and in full in the Appendix.

Table 1. Outline of the proposed problem categorisation* in the topic of colligative properties of ideal solutions

A. *Vapour pressure of a single liquid substance*	**B.** *Vapour pressure depression of an ideal solution of a non-electrolyte.*	**C.** *Boiling-point elevation and freezing-point depression of an ideal solution of a non-electrolyte.*	**D.** *Osmosis and osmotic pressure of a solution.*
A1. A single liquid substance not in equilibrium state.	**B1.** A non-electrolytic solution of a **non-volatile** substance.**	**C1.** A non-electrolytic solution of a **non-volatile** substance.**	**D1.** A non-electrolytic solution of a non-volatile substance.**
A2. A single liquid substance in equilibrium state.	**B2.** Dilution or concentration of a solution.	**C2.** Dilution or concentration of a solution.	**D2.** Dilution or concentration of a solution.
A3. A single liquid substance in two (or more) independent equilibrium states	**B3.** Mixing of solutions.	**C3.** Mixing of solutions.	**D3.** Mixing of solutions.
	B4. Two (or more) independent solutions at equilibrium state.	**C4.** Two (or more) independent solutions.	**D4.** Two (or more) independent solutions.
	B5. A solution of a **volatile** substance.		**D5.** Two non-electrolytic solutions separated by a semi-permeable membrane.
			D6. Pure solvent and a solution are separated by a semi-permeable membrane.

* Each problem involves just one colligative property.

** Unless stated differently (see Case **B5**), in all other cases, the solute(s) is (are) substance(s) that are non-electrolyte(s) and non-volatile.

Table 2. The case of problems that combine one or more colligative properties.

Problems that combine one or more colligative properties.	
1. Data include concentration. To be found: values of two or more colligative properties.	**2.** Data include value of one colligative property concentration. To be found: values of one or more other colligative properties.

Empirical investigation

The empirical investigation was carried out in the school year 1999-2000 in two public upper experimental secondary schools (*lykeion*) in the Greater Athens region. Experimental schools in Greece, while selecting their students by a draw among the applicants, are schools of higher standards in terms of their teaching staff, and consequently in terms of their student population. Five entire eleventh-grade classes were used, two from school 1 and three from school 2. The students of these classes followed the first out of two years of the ‘Positive Stream’ of Greek upper secondary school; this stream leads to higher education in medicine and other health related subjects, as well as to science and engineering. The five classes were divided into an experimental group (one class in school 1 and two classes in school 2) and a control group (one class in each school).

The material on colligative properties was the same and taught by the same teacher in each school (one teacher in school 1 and another teacher in school 2). Within each school, the same teacher taught both the experimental and the control group. Both teachers had done graduate studies in chemistry education, and also had long teaching experience.

Method of instruction in the experimental groups

Instruction in the experimental groups started with the theory of sub-unit A with the teaching methodology that the particular teachers usually follow, and within the teaching time that is dictated by the formal national curriculum. Following that, the teachers delivered instruction on the proposed categorisation of the problems in the sub-unit, providing in each case a brief suggestion/methodology for dealing with the corresponding problem. Note that these suggestions/methodologies are not shown in Tables A1-A5 in the Appendix, but two examples are included in the Appendix.

At that time, the students had in front of them a copy of the categorisation scheme that dealt only with the section dealing with the sub-unit under study (this was similar to one of the Tables A1-A5 of the Appendix). This categorisation scheme was prepared by the investigators and distributed by the teacher. Next, the teachers solved one problem on the chalkboard in collaboration with the students, after the students had assigned the given problem to the corresponding category and sub-category. Finally, for practice and consolidation of the methodology, the teachers distributed a limited number (2-3) of additional selected problems, requiring the students to categorise the problems by consulting the categorisation sheet. Students had to hand in both the solved problems and their categorisation at the start of their next chemistry class. The same teaching procedure was employed four more times, for each of the remaining three categories and for the 'combination problems'.

Method of instruction in the control classes

Instruction in the control classes started with teaching the theory of sub-unit A in the same way as in the experimental classes. Following that, the teachers solved *further* relevant problems on the chalkboard in collaboration with the students. Taking into account that the same teaching time was employed for the experimental and the control groups, as well as the fact that the instruction on problem categorisation took up a significant portion of the available teaching time, it is obvious that in the control group more problems were solved. Again, at the end of the class, the teachers distributed the same limited number (2-3) of additional selected problems for practice at home. Students had to hand in sheets with the solved problems at the start of their next chemistry class. The same teaching procedure was employed four more times, for each of the remaining three categories and for the 'combination problems'.

Comparison of performance

Students in all experimental and control classes were assessed on their ability to solve problems on colligative properties by administering to them the same test. This constituted the actual formal test in the chemistry course, dictated by the school regulations for the first term, and made a significant contribution to the formal marks for the first term. As a result, students treated the test seriously.

The test contained three problems. Each problem dealt with a different colligative property and involved simple numerical calculations. The problems are given in the Box. To

The three problems of the written test

Problems 1 and 2

The following two solutions, S_1 and S_2 are given:

S_1: A molar solution of substance A ($M_r = 40$) in a solvent X ($M_r = 80$) has a mass of 200 g.

S_2: A molar solution of substance B ($M_r = 90$) in the same solvent X has a mass of 150 g.

Problem 1. If the two solutions S_1 and S_2 have the same boiling point under the same external pressure, calculate the % w/w.

Problem 2. How many grams of solvent should be removed by evaporation from solution S_1 to produce a solution having a vapour pressure of 30 mm Hg?

Data: Vapour pressure of solvent X at the temperature of the experiment $P^o = 37.5$ mm Hg.

Problem 3

The following two solutions, S_3 and S_4 are given:

S_3: A molar solution of substance C ($M_r = 50$) in a solvent Y has a mass of 400 g

S_4: A molar solution of substance D ($M_r = 100$) in the same solvent Y has a mass of 400 g

What is the osmotic pressure at 27^{o}C of solution S_5 that results from mixing the two solutions S_3 and S_4.

Data: (1) Substances C and D do not react chemically with each other.

(2) The density of the solution S_5 is $d = 0.80$ g / mL at 27^{o}C.

(3) The value of the gas constant is $R = 0.082$ atm L / (mol K).

In addition, in the experimental classes, the following additional task was set for each of the three problems:

Find out the category and sub-category (for instance A.1.a) for the problem. (e.g. Problem 1 belongs to category, etc.

NOTE: You can consult the sheet with the problem categorisation.

In the control classes, students had with them and could consult the course textbook. The test sheet stated:

NOTE: You can consult your course textbook.

avoid the possibility of cheating among neighbouring students in class, four different, but equivalent, versions of the problems were used.

In the experimental classes, students had with them and could consult the sheets with the problem categorisation. In addition to solving the problems, these students had to categorise the three problems according to the suggested categorisation. In the control classes, students had with them and could consult the course textbook. Students had 70-75 minutes to solve the problems.

For our analysis, we took into account only those students who answered all three problems: n = 26 for the control group, and n = 41 for the experimental group. This experimental group (*the raw experimental group*) was further reduced by removing the students who *did not demonstrate* knowledge of the taught problem-categorisation scheme. Thus, in the *reduced experimental group* (n = 24), only the students who demonstrated knowledge of problem categorisation were included (as checked through their answers). For this reduction, we removed from the raw sample 2 (out of 13) students of high performance in the two national examinations (see below), 3 (out of 13) students of intermediate performance, and 8 (out of 16) students of low performance. Thus, 61% of the removed students were from those of lower performance. This shows that it was mostly students of lower performance in the national examinations who did not have (or did not demonstrate) knowledge of the problem-categorisation scheme.

Statistical analysis of the data

The small size of the samples, together with the lack of normality in the distributions in many cases dictate that non-parametric statistics should be used, in this case the Mann-Whitney test for independent samples (Cohen and Holliday, 1982, pp. 235-242). The small samples, plus a possible problem with the non-equivalence of the control and the reduced experimental group (see below), force us to treat our findings as preliminary ones. Note that because the chapter on colligative properties is excluded from the taught material of the subsequent years, it was not (and it is still not) possible to repeat the educational experiment to increase sample sizes.

Results and comments

Table 3 compares the mean scores (maximum 100) of the students of the control and experimental groups of our study in the five different levels of written testing (see above). For the experimental group, we report results both for the raw sample, and the *reduced experimental group* (see above). The comparison between the control and the raw experimental group shows very small differences in the scores in the three problems and in the mean score. In addition, the difference in the scores in National Examination 1 (in favour of the control group) is very small, while in National Examination 2, the control group had a superior performance. The latter difference might explain why the control group showed higher scores than the raw experimental group in Problems 1 and 2 and in the mean of the three problems.

Table 3. Performance* of the control and the experimental groups in the two national examinations and the problems of this study. (Maximum score: 100.)

	National exam 1	National exam 2	Problem 1	Problem 2	Problem 3	Mean, Problems 1, 2, 3
Control group (*n* = 26)	86.9 (11.7)	81.2 (12.3)	76.8 (26.2)	56.2 (33.3)	75.6 (28.9)	69.5 (22.1)
Raw experimental group** (*n* = 41)	85.5 (16.0	76.2 (22.7)	69.3 (32.1)	53.6 (36.9)	78.3 (26.2)	67.1 (27.4)
Reduced experimental group*** (*n* = 24)	92.8 (9.4)	85.5 (13.8)	82.4 (23.5)	74.0 (28.5)	89.0 (19.7)	81.8 (19.6)

* Mean scores, with standard deviations in parentheses.
** All participating students of the experimental group who answered all three problems.
*** Only experimental-group students who answered all three problems and also demonstrated knowledge of the problem-categorisation scheme.

Turning to the comparison between the control and the reduced experimental group, first we note that the latter has now higher scores in the two national examinations. This resulted from the fact that most of the weaker students in these examinations were excluded from the experimental group (see above). Performance of the experimental group is now higher by about 6% in Problem 1, by 18% in Problem 2, by 13% in Problem 3, and by 12% in the mean of the three problems. The lower difference in Problem 1 could be attributed to its order, since most students might have spent more time on it. By using the non-parametric Mann-Whitney test for independent samples (Table 4), the differences in Problem 2 and the mean of the three problems are statistically significant near the 95% significance level (p = 0.05). Significant (also in favour of the experimental group) is the difference in National Examination 1 at p = 0.02. Problem 2, as judged by student performance, proved more difficult, and this might be one reason for its showing a higher effect of the method. By using parametric statistical analysis of covariance with the scores of the two national examinations as covariates, we found that in none of the problems were the differences statistically significant. The F-ratio values (with significance levels in parentheses) are as follows: Problem 1, 0.00 (1.00); Problem 2, 1.82 (0.18); Problem 3, 1.08 (0.31); mean of Problems 1, 2, 3: 1.49 (0.23); (We repeat that parametric methods of analysis are not appropriate, as our samples in many cases did not follow the normal distribution.)

Table 4. Non-parametric statistical analysis (Mann-Whitney test) for comparisons of control (n = 26) versus reduced* experimental group (n = 24): large sample** test statistic Z, with two-tailed probability of equaling or exceeding Z in parentheses.

National exam 1	National exam 2	Problem 1	Problem 2	Problem 3	Mean, Problems 1, 2, 3
2.28 (0.023)	1.44 (0.15)	0.88 (0.38)	1.98 (0.048)	1.70 (0.09)	2.00 (0.045)

* Experimental-group students who demonstrated knowledge of the problem-categorisation scheme.
** The larger of the two samples (here both samples) has more than 20 data points (Cohen and Holliday, 1982, p. 239).

Comparison according to level of performance in the national examinations

The scores in the two national examinations provided a reliable means for dividing the students of our sample into three levels of performance (high, intermediate, and low). This division is independent of their performance in the tests of this study.

Table 5 shows the scores of the control group and the reduced experimental group in the five levels of testing for students of high, intermediate, and low performance respectively in the two national examinations. The criteria for this division are given in the tables. Because of the small number of students, these criteria were set to create groups that were comparable in size. In all cases, students of high performance outperformed students of intermediate performance in all three problems; similarly, the latter students outperformed the students of low ability.

The stability of the scores of the high-performing students is remarkable. These students achieved high scores in all problems, and consequently the taught problem-categorisation scheme had not affected them. At the other end, low-performing students also showed a lack of effect. The significant difference was due to the intermediate students of both the raw and the reduced experimental groups; although they fell a bit behind in (by about 4%) in Problem 1, in the other two problems and in the mean of the three problems, they scored considerably higher than the students of the control group. By using the non-parametric Mann-Whitney test for independent samples (Table 6), only the differences in Problem 2 are statistically significant (the corresponding values of statistic U are smaller than the critical values at the 5% significance level). Also, near significance is the difference in the mean of the three problems for the reduced experimental group. Again, Problem 2 showed a convincing effect of the categorisation. Note however the superiority of the reduced experimental group in National Examination 1 (which is near statistical significance).

Table 5. Performance* of the control group and the reduced experimental group in the two national examinations and the problems of this study for the students with *high, intermediate, and low performance*** in the two national examinations.

	National exam 1	National exam 2	Problem 1	Problem 2	Problem 3	Mean, Problems 1, 2, 3
High-performed students						
Control group (n = 7)	97.4 (3.4)	96.5 (2.2)	92.4 (8.1)	93.9 (16.2)	95.9 (8.1)	94.1 (6.1)
Reduced experimental group (n = 11)	98.3 (2.0)	95.2 (3.1)	93.8 (9.7)	87.3 (18.3)	93.8 (10.1)	91.7 (9.7)
Intermediate students						
Control group (n = 11)	89.0 (4.0)	83.0 (5.1)	82.1 (23.8)	45.8 (24.9)	72.7 (32.1)	65.8 (17.5)
Reduced experimental group (n = 9)	93.4 (5.4)	84.2 (5.2)	78.4 (27.7)	73.4 (26.5)	93.7 (11.9)	81.8 (17.8)
Low-performed students						
Control group (n = 8)	73.8 (11.6)	68.0 (8.1)	60.3 (32.1)	37.5 (29.3)	61.6 (28.3)	53.1 (18.4)
Reduced experimental group (n = 4)	76.4 (10.8)	61.6 (15.4)	59.8 (25.8)	38.4 (29.6)	65.2 (36.2)	54.4 (20.6)

* Mean values (%), with standard deviations in parentheses.

** For high-performed students, scores ≥ 92.5%; for intermediate students, 80% < scores < 92%); and for low-performed students 80% ≥ scores.

Table 6. Students of intermediate performance in the two national examinations: non-parametric statistical analysis (Mann-Whitney test) for comparisons of control and experimental groups (values of statistic U).*

	National exam 1	National exam 2	Problem 1	Problem 2	Problem 3	Mean, Problems 1, 2, 3
Control group (n = 11) versus raw experimental group (n = 12).	40	46	62.5	32.5	50	43
	Critical value of U for two-tailed test (p = 0.05) = 33					
Control group (n = 11) versus reduced** experimental group (n = 9).	25.5	35	49	21.5	29	24
	Critical value of U for two-tailed test (p = 0.05) = 23					

* The smaller of the two U values is given (Cohen and Holliday, 1982, p. 237).

** Experimental-group students who demonstrated knowledge of the problem-categorisation scheme.

The solution process

The kinds of errors committed

An interesting aspect of this work are the kinds of errors committed by the students, irrespective of the group (experimental or control) they were in. Of 141 main errors that were spotted 84 (59.6%) were due to deficiencies in understanding of fundamental concepts related to solutions (concentration, dilution and condensation, mixing of solutions). One explanation might be that these concepts were last to be taught during the previous (tenth) grade (taught in a hurry or not at all, and possibly also reviewed hastily at beginning of eleventh grade). An additional 18.4% (26 errors) were attributed to problems in understanding of colligative properties and their laws. The remaining 22.0% (31 errors) were due to lack of attentiveness, errors in numerical calculations, and various other reasons.

Numerical calculations in problem solving

Another interesting finding is the frequency of numerical errors committed by the students who had applied the correct solution procedure, irrespective of the group they were in. For this purpose, we identified all students who had applied the correct solution procedure, irrespective of whether or not they performed successfully the numerical calculations involved. There were 44 such students. Of these, 19 (43.2%) found the correct numerical result, while 25 (56.8%) made mistakes in calculations. This demonstrates that a considerable proportion of students fail to perform numerical calculations correctly, a finding that might be attributed to the fact that in their training students are not encouraged to finish the solution of problems that are given to them for practice both at school and home. Instead the emphasis is for practicing in doing as many problems/exercises as possible in the time available.

Concentrating on the way of performing calculations, some of them performed the calculations in a classical/serial manner, while others chose to use simplifications that speed up the calculations. Table 7 summarises these findings.

Table 7. Numerical errors committed by the students while performing numerical calculations by making or not making simplifications.

Number of students who *did not* make numerical errors: 25	Without simplifications: 11 (44.0%)
	With simplifications: 14 (56.0%)
Number of students who made numerical errors: 19	Without simplifications: 16 (84.2 %)
	With simplifications: 3 (15.8%)

Conclusions and implications for teaching and learning

In general, the results of this (preliminary) study did not reveal significant differences in performance between students who were taught problem categorisation explicitly and the control group. Problem categorisation is not an integral part of science teaching, so students do not have the experience or even a positive disposition toward using it. On the other hand, one can put the blame on the serious conceptual difficulties that the students of our study experienced in dealing with colligative properties themselves, as well as with the auxiliary concepts of solution chemistry (concentration, dilution and condensation, mixing of solutions). As a matter of fact, these weaknesses were responsible for 78% of the errors committed, so they left little space for the effect of the categorisation to be felt. Therefore, categorisation techniques are not by themselves capable or sufficient for conceptual understanding (Bunce et al., 1991).

A definite positive argument in favour of the categorisation scheme derives from the methodology used in this study. The experimental group had the sheets with the categorisation scheme and they had an extra task to be completed in the same time as the control group has for just the test. In addition, the control group had the course text available for consultation. From the test scores, it is fair to conclude that training in categorisation is at least equivalent to having more practice at doing these exercises (during the teaching), and the availability of the scheme in written form is at least the equivalent of, and for some students clearly better than, having the course text in terms of the help it can offer.

The most interesting finding resulted from looking separately at the performance of students of high, low and intermediate performance, as these three levels derived independently from their scores in the national examinations in chemistry. Students of high performance were good at solving problems irrespective of whether or not they were taught categorisation schemes. As a result, the methodology had no observable effect on them. These students had constructed on their own mental representations of these problems, resulting in the proper categorisation; hence the offered categorisation scheme was redundant to them. In addition, these students, being more industrious and attentive, should have had more practice in similar problems. At the other end, students of low performance also failed to benefit from the methodology offered. There are many possible reasons for this failure: insufficient conceptual learning, deficiencies in the proper understanding of the problem statements, weakness in manipulating the mathematical relations entering the problems, lack of interest in the particular lesson and/or the whole school process.

There was, however, a group of students who might have benefited from the taught categorisation. It was the group of students with intermediate performance. Students of intermediate abilities and overall performance are more likely to be receptive to the proposed problem categorisation. The above finding suggests that it is useful for teachers to follow a systematic organised approach to teaching problem solving. Problem categorisation, if applied in many different areas of problems, can result in a sound construction of a knowledge base that can contribute to successful problem solving, especially for intermediate-ability and performance students.

Note also that, while in this study, for reasons of economy in time, an explicit method of problem categorisation was used, that is, a receptive/passive model of teaching was adopted, consistent with Ausubel's meaningful learning theory (Ausubel, 1968), it can be predicted that the suggested methodology can be more successful if an active/constructivist model is adopted. To this end, students can work out the categorisation schemes on their own or collaboratively in groups under the guidance of their tutor. In any case, to become good problem solvers, students must be given ample opportunity to do it. This will not give them just practice, it will develop confidence. Further, students who have obtained good answers should not be ignored by the teacher; they also need guidance and help.

A problem-categorisation scheme essentially constitutes a set of rules that serve to make explicit the conditions in a problem that indicate the solution procedure. Although the assertion that explicit teaching will be helpful appears intuitive, further analysis and a review of the literature point variously to reasons either to support or question this claim. Intuitively, it might be expected that teaching students the link between the conditions and the solution would encourage this link to be made. Whether this link is likely to be invoked in real problem-solving situations remains an open question, but if it is learned, at least the possibility is there! However, it should be noted that at least some students are able to invoke the correct procedure without such teaching, and some probably do so without awareness of the link between conditions and procedure. Some learners may consciously notice and understand the pattern in the conditions, linking this to a particular solution procedure. On the other hand, one should take into account that the teaching of explicit rules adds a further layer of abstraction and complexity that students must process, and such additional information is liable to strain working-memory resources.

The main aim of research on problem solving in science is the development of a theory or models that can explain the interaction between a problem and the problem solver. The usefulness of such models is not simply their explanatory power, but mainly their predictive power. We hope that this study has contributed toward that aim, and that together with previous studies (Bunce et al., 1991) support the belief that explicit instruction in problem solving strategies can increase problem solving ability (Reif, 1981). It can also teach experts' methodology. Because of the errors likely to have occurred (due to the sampling and the 'trimming' of the sample to get the reduced experimental group), the conclusions are only preliminary, so a lot remains to be investigated further.

Acknowledgements

The authors thank the following high-school teachers who each helped in different ways in carrying out of this work: Spyridoula Angeli, Vasileios Angelopoulos, Erifyli Zarotiadou, Constantinos Kampourakis, Antonios Bakolis, Georgios Papaphotis and Spyridon Petsios. We are grateful to Dr. Norman Reid for his detailed feedback and particularly for the arguments he has supplied from cognitive psychology and the second language-acquisition literature. We also thank Dr. Stephen Breuer, co-editor of CERP, who read the manuscript and made a number of useful comments. Finally, GT wishes to acknowledge once again his debt to Professor Alex H. Johnstone for his inspiration for pursuing systematically the investigation of problem solving in science education, as well as for his constant critical judgement, advice, encouragement and help during his sabbatical in 1990 in the Centre of Science Education of the University of Glasgow and in all the subsequent years.

References

Ausubel D.P., (1968), *Educational psychology - a cognitive view*, Jolt, Rinehart & Winston, New York.

Bodner G., (1987). Role of algorithms in teaching problem solving, *Journal of Chemical Education*, **64**, 513-514.

Bunce D.M., Gabel D.L., and Samuel J.V., (1991), Enhancing chemistry problem-solving achievement using problem categorization, *Journal of Research in Science Teaching*, **28**, 505-521.

Bunce D.M. and Heikkinen H., (1986), The effects of an explicit problem-solving approach on mathematical chemistry achievement, *Journal of Research in Science Teaching*, **23**, 11-20.

Chi M.T.H. and Koeske R.P., (1983), Network representation of a child's dinosaur knowledge, *Development Psychology*, *19*, 29-39.

Cohen L. and Holliday M., (1982), *Statistics for social scientists*, Harper & Row, London.

De Jong T. and Ferguson-Hessler M.G.M., (1986), Cognitive structures of good and novice problem solvers in physics, *Journal of Educational Psychology*, **78,** 279-288.

Inhelder B. and Piaget J., (1958), *The growth of logical thinking from childhood to adolescence*, Basic Books, New York.

Johnstone A.H., (1991), Thinking about thinking, *International Newsletter on Chemical Education,* No. 36, 7-11.

Johnstone A.H., (1993), Introduction, in C. Wood & R. Sleet, *Creative problem solving in chemistry,* pp. iv-vi, The Royal Society of Chemistry, London.

Larkin J.H., (1980), Teaching problem solving in physics: issues in teaching and research, in D.T. Tuma and F. Reif (eds.), *Problem solving and education: issues in teaching and research,* New York.

Niaz M., (1995), Relationship between students performance on conceptual and computational problems of chemical equilibrium, *International Journal of Science Education,* **17**, 343-355.

Pacton S., Perruchet P., Fayol M., and Cleeremans, (2001), Implicit learning out of the lab: the case of orthographic regularities, *Journal of Experimental Psychology*, **130**, 401-426.

Reif F., (1981), Teaching problem solving – a scientific approach, *Journal of Research in Science Teaching*, **19**, 310-316.

Reif F., (1983), How can chemists teach problem-solving? *Journal of Chemistry Education*, **60**, 949-953.

Simon D.P. and Simon H.A., (1978), in R.S. Siegler (ed.), *Childrens' thinking: what develops,* Lawrence Erlbraum Associates, Hillstate NJ.

Sternberg R.J., (1981), Intelligence and nonentrenchment, *Journal of Educational Psychology*, **73**, 1-16.

Truscott J., (1998), Noticing in second language acquisition: a critical review, *Second Language Research,* **14**, 103-135.

Tsaparlis G., (1994), Blocking mechanisms in problem solving from the Pascual-Leone's M-space perspective, in H.-J. Schmidt (ed.), *Problem solving and misconceptions in chemistry and physics*, pp. 211-226, ICASE, Dortmund.

Tsaparlis G., (1998), Dimensional analysis and predictive models in problem solving. *International Journal of Science Education*, **20***,* 335-350.

Tsaparlis G. and Zoller U., (2003), Evaluation of higher *vs.* lower-order cognitive skills-type examinations in chemistry: implications for university in-class assessment and examinations, *University Chemistry Education*, **7**, 50-57.

Wilson J. M., (1994), Network representations of knowledge about chemical equilibrium: variations with achievement, *Journal of Research in Science Teaching*, **31**, 1133-1147.

Zoller, U., (1993), Are lecture and learning compatible? Maybe for LOCS: unlikely for HOCS, *Journal of Chemical Education,* **70***,* 195-197.

Zoller U. and Tsaparlis G. (1997), Higher and lower-order cognitive skills: the case of chemistry, *Research in Science Education,* **27**, 117-130.

Appendix:
The proposed categorisation scheme of problems in colligative properties of ideal solutions

Part 1: Tables setting out the problem categorisation

NOTE: These are the tables that were used by the students.

Table A1. Proposed problem categorisation in the topic of vapour pressure of a single substance.

A. *Vapour pressure of a single liquid substance.*	
A1. *A single liquid substance not in equilibrium state.*	**a.** A quantity of **liquid substance** is placed in an empty vessel at a given temperature. **No** dynamic **vapour-liquid equilibrium** is established. (The amount of the substance is inadequate, so that the liquid completely evaporates P.) P_{vapour} is the vapour pressure.
	b. A quantity of **gaseous substance (vapour of the substance)** is placed in an empty vessel at a given temperature. **No** dynamic **vapour-liquid equilibrium** is established. (The amount of the substance is inadequate, so that all the substance remains in the gaseous state.) P_{vapour} is the vapour pressure.
A2. *A single liquid substance in equilibrium state.*	**a.** A quantity of **liquid substance** is placed in an empty vessel at a given temperature. Dynamic **vapour-liquid equilibrium** is established. *The* vapour pressure is P^o.
	b. A quantity of **gaseous substance (vapour of the substance)** is placed in an empty vessel at a given temperature. Dynamic **vapour-liquid equilibrium** is established. The vapour pressure is P^o.
A3. *A single liquid substance in two (or more) independent equilibrium states.*	A quantity of the same **liquid substance** is placed in two (or more) empty vessels. Dynamic **vapour-liquid equilibrium** is established ($P = P^o$) or not established in each vessel at the same temperature ($P < P^o$).

Table A2. Proposed problem categorisation in the topic of vapour pressure depression of an ideal solution of a non-electrolyte.

B. *Vapour pressure* depression *of an ideal solution of a non-electrolyte.*	
B1. *A non-electrolytic solution of a* ***non-volatile*** *substance.**	A quantity of a solution of a substance which is not an electrolyte* (or more than one such substances that do not react chemically with each other) is introduced into an empty vessel at a given temperature. Dynamic **vapour-liquid equilibrium** is established. The vapour pressure is P.
B2. *Dilution or concentration of a solution.*	A quantity of a solution of a substance (or more than one substance that do not react chemically with each other) is **diluted** or **concentrated** at *constant temperature* (by *adding* or *removing* a quantity of solvent). Dynamic **vapour-liquid equilibrium** is established.
B3. *Mixing of solutions.*	A quantity of a solution of a substance (or more than one substance that do not react chemically with each other) is **mixed** with a quantity of *another solution* (or more solutions) at *constant temperature.* Dynamic **vapour-liquid equilibrium** is established.
B4. *Two (or more) independent solutions at equilibrium state.*	Quantities of two (or more) independent solutions of a substance (or more than one substance that do not react chemically with each other), in the same solvent and at the same temperature, are in state of dynamic **vapour-liquid equilibrium**.
B5. *A solution of a* ***volatile*** *substance.*	A quantity of a solution of a **volatile** substance (or more than one **volatile** substance that do not react chemically with each other) is introduced into an empty vessel, at a given temperature. Dynamic **vapour-liquid equilibrium** is established. The vapour pressure is P_{total}.

* Unless stated differently (see Case **B5**), in all other cases, the solute(s) is (are) substance(s) that are non-electrolyte(s) and non-volatile.

Table A3. Proposed problem categorisation in the topic of boiling-point elevation and freezing-point depression of an ideal solution of a non-electrolyte.

C. *Boiling-point elevation and freezing-point depression of an ideal solution of a non-electrolyte.*	
C1. *A non-electrolytic solution of a* ***non-volatile*** *substance.**	A quantity of a solution of a substance which is not an electrolyte* (or more than one such substances that do not react chemically with each other) boils or freezes at a given temperature.
C2. *Dilution or concentration of a solution.*	A quantity of a solution of a substance (or more than one substance that do not react chemically with each other) is **diluted** or **concentrated** at *constant external pressure* (by *adding* or *removing* a quantity of solvent).
C3. *Mixing of solutions.*	A quantity of a solution of a substance is **mixed** with a quantity of *another solution* (or more solutions) of the same substance (or a different substance that does not react chemically with the other substance) at *constant external pressure.*
C4. *Two (or more) independent solutions.*	Two (or more) independent solutions of a substance (or more than one substance that do not react chemically with each other) in the same solvent and at the same temperature. Each of these solutions boils or freezes at a given temperature (different for each one of them).

* In all cases, the solute(s) is (are) substance(s) that is (are) non-electrolyte(s) and non-volatile.

Table A4. Proposed problem categorisation in the topic of *osmosis and osmotic pressure.*

D. *Osmosis and osmotic pressure of a solution.**	
D1. *A non-electrolytic solution of a non-volatile substance.***	A given quantity of a solution of a non-volatile substance, which also is non-electrolyte* (or even more than one such substances that do not react chemically with each other), has osmotic Π, at a given temperature.
D2. *Dilution or concentration of a solution.*	A given quantity of a solution of a substance (or more than one substance that do not react chemically with each other) is **diluted** or **concentrated**, at a given temperature.
D3. *Mixing of solutions.*	A quantity of a solution of a substance is **mixed** with a quantity of *another solution* (or more solutions) of the same substance (or a different substance that does not react chemically with the other substance) at a given temperature.
D4. *Two (or more) independent solutions.*	Two (or more) independent solutions of a substance (or more than one substance that do not react chemically with each other) have osmotic pressures Π_1, Π_2 respectively, at the same temperature (or different temperatures).
D5. *Two non-electrolytic solutions separated by a semi-permeable membrane.*	Two different solutions of the same substance (or more substances that do not react chemically with each other) in the same solvent are separated by a semi-permeable membrane, at a given temperature. The phenomenon of osmosis occurs until a dynamic equilibrium is established.
D6. *Pure solvent and a solution are separated by a semi-permeable membrane.*	Pure solvent and a solution of a substance in the same solvent (or more than one substance that do not react chemically with each other) are separated by a semi-permeable membrane, at a given temperature. The phenomenon of osmosis occurs until a dynamic equilibrium is established. At equilibrium, the two columns of liquids have a difference in height, corresponding to the osmotic pressure of the solutions.

* In all cases, the processes are taking place under constant external (atmospheric) pressure.

** In all cases, the solute(s) is (are) substance(s) that are non-electrolyte(s) and non-volatile.

Table A5. The case of problems that combine one or more colligative properties.

Problems that combine one or more colligative properties.	
1. *Data include concentration. To be found: values of two or more colligative properties.*	Problem involves a solution of a non-volatile substance which also is non-electrolyte (or more than one such substances that do not react chemically with each other). The concentration of the solution in given, conditions are given. Values of two or more colligative properties of the solution are to be found.
2. *Data include value of one colligative property concentration. To be found: values of one or more other colligative properties.*	Problem involves a solution of a non-volatile substance which also is non-electrolyte (or more than one such substances that do not react chemically with each other). The value of one particular colligative property of the solution is given. Values of two or more other colligative properties are to be found.

Part 2. Examples of suggestions/methodology for dealing with the problems

EXAMPLE 1, CASE **A.** *Vapour pressure of a single liquid substance* **(A1)** *A single liquid substance not in equilibrium state***.**

(A1.a) A quantity of **liquid substance** is placed in an empty vessel at a given temperature. **No** dynamic **vapour-liquid equilibrium** is established. (The amount of the substance is inadequate, so that the liquid completely evaporates). P_{vapour} is the vapour pressure at the given temperature.

√ Apply the ideal-gas equation for the vapour:

$$P_{\text{vapour}} V_{\text{vapour}} = n_{\text{vapour}} R\, T$$

(N.B. $P_{\text{vapour}} < P^{o}$)

EXAMPLE 2, CASE **D**. *Osmosis and osmotic pressure of a solution;* **(D2)** *Dilution or concentration of a solution.*

A given quantity of a solution of a substance (or more than one substances that do not react chemically with each other) is **diluted** or **concentrated** under constant external pressure, at a given temperature.

√ Apply Vant's Hoff's law both for the initial and the final solution:

$$\Pi_1 V_{\text{solution, 1}} = n_{\text{solute, 1}} R\, T$$

$$\Pi_2 V_{\text{solution, 2}} = n_{\text{solute, 2}} R\, T$$

N.B. (a) in dilution: $n_{\text{solute, 1}} = n_{\text{solute, 2}}$

(b) in condensation with evaporation of solvent: $n_{\text{solute, 1}} = n_{\text{solute, 2}}$

(c) in condensation with addition of solute: $n_{\text{solute, 1}} + n_{\text{solute, added}} = n_{\text{solute, 2}}$

Fostering creative problem solving in chemistry through group work

Liberato Cardellini

Department of Material and Earth Science, Polytechnic University of the Marche, Italy
e-mail: libero@univpm.it

Received 30 November 2005, accepted 4 March 2006

Abstract: Although problem solving is a very important higher-order cognitive skill, our students seem to believe that this activity does not deserve too much effort and they develop the attitude that arriving at the answer is more important than understanding the process of solution. This is due in part to the way we teach problem solving. Usually, when teaching, we show them only some stages of the process, neglecting the analysis stage, because as experts we are now no longer able to recall the effort we had to expend the first time we tried to solve a problem, since it is now familiar to us. From our presentation, students see a clean, even elegant solution, having little in common with the uncertainty and the fuzzy thinking that they experience when they try to solve a problem by themselves. From research we know that it is quite difficult for students to develop creative solutions to problems. The results reported here are promising in making students active developers of original solutions. These results are tentatively attributed to the use of an active method of learning and to the students' motivation. [*Chem. Educ. Res. Pract.,* 2006, **7** (2), 131-140]

Keywords: problem solving; problem representation; creativity; cooperative learning; analysis skills; learning styles.

Introduction

Problem solving is very important for many subjects. Chemistry is no exception, combining in its problems characteristics of mathematics and physics problems, and adding its distinct chemical features (stoichiometry, chemical synthesis, chemical analysis, etc.). In stoichiometry, important quantitative relations constitute conceptual building blocks. But problem solving is a higher-order cognitive skill (Zoller, 1993) which demands many abilities, sometimes requiring much effort from the solver. Problem solving is a process in which various reasoning patterns are combined, refined, extended, and invented. It is much more than substituting numbers in well-known and practised formulas; it deals with creativity, lateral thinking and formal knowledge. Research has tried to correlate some cognitive variables, such as formal operational reasoning, working memory capacity, disembedding ability, specific knowledge, concept relatedness and idea association, to science achievement and problem solving ability (Mayer, 1975; Johnstone and El-Banna, 1986; Niaz, 1987; Camacho and Good, 1989; Niaz and Logie, 1993; Tsaparlis et al., 1998; Stamovlasis and Tsaparlis, 2000; Lee et al., 2001).

Suggestions from research are, that instructional methods should take into account the general strategies and methods of problem solving, thus providing a tool to increase reasoning skills in the problem solver. Life is, in essence, a continuous process of problem solving and selection from available and/or created options, and it is sobering to note that sometimes people who have received formal training in problem solving are not necessarily the best problem solvers. Nevertheless, problem solving abilities/decision making capacities are valuable and precious skills not only in academia, but also in the world of business and industry and in daily living. Furthermore, in science these skills play an important role in the acquisition and organisation of knowledge in a meaningful way. The way knowledge (both 'know-how' and 'know-what') is organized in memory is generally recognized to be critically related to the degree of success in problem solving (De Jong and Ferguson-Hessler, 1986). To help students organize and structure their knowledge, they are asked to draw concept maps (Novak and Gowin, 1984; Novak, 1998) for every topic presented in the course syllabus.

Access to knowledge

If an important prerequisite for being a successful problem solver is that knowledge must be activated in the Long Term Memory (LTM) when needed, this must have implications for the way we teach and the way in which students are encouraged to lay down knowledge in LTM. If the research on learning styles (Entwistle, 1988; Lawrence, 2000) is to be taken seriously, we have to admit that a lecture may not be the best way to facilitate the orderly linkage of knowledge in LTM. Not surprisingly, research shows that note-taking is affected by a person's information processing ability and by individual psychological factors such as field dependence/independence (Johnstone and Su, 1994). Moreover, students taught almost entirely in a way that clashes with their learning styles, do not learn as much as students taught in their preferred styles, and they retain less of what they learn (Felder, 1996a).

Many years ago, Whitehead warned about the danger of knowledge that is accessed only in a restricted set of contexts even though it is applicable to a wide variety of contexts, and he called it inert knowledge (Whitehead, 1953). Knowledge stored in sealed boxes without meaningful connections with other concepts is inert knowledge. This point is illustrated by Johnstone in another paper in this issue of CERP, where he reports on the problem of teaching senior undergraduates in bioinorganic chemistry. Students complain that in the same lecture they are meeting concepts of thermodynamics, complex formation and stability and organic ligands. It is clearly very uncomfortable for them to "*open several boxes at once*". (Johnstone and Otis, 2006).

Several factors influence the abilities in solving problems, from the nature of the problem, to the learners' developmental level and their knowledge base, to motivation and problem solving skills, to many individual and psychological factors. A deep and comprehensive analysis of those factors can be found in the literature (Reid and Yang, 2002).

This study looked at a group of engineering students (21 girls and 24 boys, aged 19-22) in the first term of their first year at university. Two psychological measurements were applied to the group to see if there was any relationship between these results and the quality of the creative problem solving resulting from this approach. These were (a) Formal Operational Reasoning and (b) Disembedding Ability.

The former was measured using the Group Assessment of Logical Thinking (GALT) test (Roadrangk et al., 1983). The scores ranged from 11 to 24 (out of 24) with a mean of 19.5 and standard deviation of 3.2.

The latter was measured by the Field Dependence/Field Independence test devised and calibrated by El-Banna (1987) based upon the original work of Witkin (Witkin, 1974; Witkin and Goodenough, 1981). Out of a possible score of 20, the range achieved was 4-19, with a mean value of 14.0 and a standard deviation of 3.4.

As we see in the literature, the development and improvement of education requires the fostering of creativity and critical thinking (CT) skills: "*Critical thinking is reasonable reflective thinking that is focused on deciding what to believe or do.*" (Ennis, 1987) There is no question that many teachers would like to improve the CT abilities of their students. How can this be done? Or, in the Schoenfeld words: "*The critical question is: can we train novices to solve problems as experts do?*" (Schoenfeld, 1980).

To become an expert in any domain, a lot of work is needed. It is estimated to take about 10,000 hours; a much longer time than any undergraduate course (Hayes, 1988). However, the experience described in this paper suggests that something valuable can be achieved in a short time. In their seminal work, Larkin et al. (1980) found four differences between experts and novices. Glaser and Chi (1988) enumerate seven characteristics of experts, but the difference, which is of major importantance for the novice, is that the expert spends a great deal of time analysing a problem qualitatively. In light of this thought, it was our hypothesis that students can be more successful in problem solving if they get accustomed to spending more time in analysing problems. The experience described in this paper is an attempt to encourage students to analyse problems and to apply novel methods to their solution.

Attempting to improve analysis skills

As teachers we believe that working on problems is an effective way to learn. Unfortunately, our students usually develop the attitude that arriving at the answer is more important than understanding the solving process. Many students start to calculate something from the text of the problem, without ever asking themselves whether this calculation will get them closer to the correct solution. Why might this be so? "*Textbook solutions to problems and solutions presented by instructors on the blackboard are always efficient, well-organized paths to correct answers*." (Herron, 1986) They apply algorithms developed by experts after repeated solutions of similar problems. "*They provide no indication of the false starts, dead ends, illogical attempts, and wrong solutions that characterize the efforts of students when they work in problem solving.*" (Herron, 1990; Bodner, 2003)

If we want to help our students, we have to find a different way to teach problem solving, a way that obliges students to spend more time analysing the problem. It was decided to try a fresh approach to teaching problem solving based upon the Analysis, Synthesis and Verification (ASV) method developed by the author (Cardellini, 1984) and incorporating the use of a cooperative approach involving small group working.

The initial problems tackled were non-chemical and non-algorithmic to emphasise the analysis and synthesis operations without the interference of chemical concepts which students may not have mastered as yet. Examples of the kind of problems are shown below.

Problem 1. Consider two containers, A and T. Container A has 10 mL of water in it from the Adriatic Sea; T contains 9 mL of water from the Tyrrhenean. Suppose that 1 mL of Adriatic water are removed from container A and put into T. After the liquid in T is mixed thoroughly, 1 mL of the mixture is removed and added to the contents of container A. Which container now has the greater amount of foreign water, the Adriatic water being foreign to T, or the Tyrrhenean to A? (Adapted from Case, 1975)

Problem 2. Two friends meet after a long time. While catching up on each other's news, one discovers that the first has married and has three sons. Then the second asks their age and the first one answers – "*the product of their ages is 36, and their sum is equal to that house number there*" – pointing to the number under the porch of the house. But the second replies – "*it is not sufficient.*" And the first one – "*OK. Then I will also add that the youngest has blue eyes.*"

Without any familiar algorithms to depend upon, students had to reason their way into the problems, draw diagrams (if necessary) to represent the problems and then, having analysed them, to plan a method for their solution. At the end came the verification of the results to be assured of their reasonableness. The ASV approach was aided by a number of questions which helped the students to subdivide the problems to reduce the load on Working Memory.

A key feature of the approach was small group working. (See Wood's paper in this issue; Wood, 2006). The class was organised into groups of three, giving seven groups of girls and eight of boys. Before starting on the problems, they were briefed in the use of concept maps (to aid the problem analysis) and in the techniques of group work. Effort has to be made to try to change the students' attitude toward problem solving: to shift the focus from looking at the product (the solution) to the importance of viewing problem solving as a conscious reasoning process.

Solving these problems requires time and it is not possible to apply a formula to them, so students have to use their reasoning. Maybe for the first time, students have to use some heuristic strategies such as decreasing the complexity of the problem, breaking down the task into subproblems, making the problem visible by translating it into pictures, diagrams or graphs, solving an analogous problem, or even working backwards, if this can help.

The first stage in problem solving is probably where the solver works hardest trying to understand the problem, extracting relevant information and translating it, or part of it, into a familiar form. "*This is a holistic or gestalt stage where relevant information is 'disembedded' from the problem, and the elements of the problem are juggled more or less simultaneously until the problem is 'restructured' or transformed into a problem that the student understands*" (Bodner and McMillen, 1986).

In this method of teaching problem solving, there is a special way of dealing with the students' errors. When the students have solved the problem in groups, one of them presents the solution on the blackboad, which is then discussed with the class. The same is done with the wrong solutions. At the end of every lesson, some problems are suggested as homework. Before starting each lecture, comments are made on the problems collected and corrected the day before, and the new problems are collected for correction.

The cooperative learning approach in detail

Cooperative learning is a method of active learning where students are involved in some activity beyond listening to the teacher. In the past students have been expected to learn how to solve problems by looking at the way teachers solve the problem at the blackboard. Instead, in this approach the students are actively involved in the learning process. Students tackle problems in groups, according to certain roles (Problem solver, Sceptic, Checker/Recorder) and follow a structured procedure under conditions that meet the five criteria of cooperative learning set out below (Johnson, Johnson and Smith, 1991; Felder, 1996b).

1. *Positive interdependence.* Team members must rely on one another to accomplish goals.
2. *Individual accountability.* Members are held accountable for (a) doing their share of the work and (b) mastering all material.
3. *Face-to-face interaction.* Some or all of the work is to be done by members working together.
4. *Appropriate use of interpersonal skills.* Team members practise and receive instruction in leadership, decision-making, communication and conflict management.

5. *Self-assessment of group functioning.* Teams periodically reflect on what they are doing well as a team, what they could improve, and what they will do differently in the future.

According to this method, the Problem solver has to think aloud, and the Sceptic has to understand the solution, asking for explanations when necessary. By working in groups, students are more likely to activate some critical thinking process. So the role of the Sceptic is very important, especially if he helps the Problem solver to move away from assumptions which might lead to a dead end. Because, for every new problem, the roles rotate, every student has the opportunity to improve her/his capacity in the analysis, synthesis and verification processes. Thinking aloud has the additional benefit of reducing the speed of the thinking process, discouraging jumping to conclusions and improving the chances of arriving at a more accurate conclusion. Constructing the solution in the group, negotiating the meaning, explaining inferences, teaching one another, making sense of the relationships facilitates the deep understanding of the solution. Making the groups work fruitfully, requires competence in applying the cooperative learning method, and effort from the teacher.

The students were told that there would be less class time spent on lectures and about half the time or more on problem solving. Since problem solving is so important, they were promised that I would correct *every* single problem they solved and give each of them my opinion and suggestions on their solution. I explained why this was being done, assuring the students that it was not an expedient to make my life easier – quite the contrary, but in this way they would learn more and they would become more confident in their abilities. Experience shows that students learn in a more meaningful way if they teach one another some of the time rather than just listen to lectures (Cardellini, 2000a).

They were reminded that this method of problem solving also would prepare them for their professional careers where they would certainly have to work in teams. Before each new lesson (three lessons amount to seven hours every week), students were asked to sit according to their groups, and the maps they had constructed and the problems they had solved as homework were collected. Then the maps collected and corrected the day before were given back, with advice on how to improve them, and from time to time, someone was asked to come to the blackboard to answer questions about the material reported on the map. This was done to be sure that they had studied the theory behind the problems, and to make sure that I had explained it properly. They had my cell-phone number and my e-mail address. They knew they could ask me about anything that was not clear, but they also knew they needed to be prepared when they came to the class.

Students were quite happy to work in this way and play according to the rules. In the 50 hours of the course more than 1900 solutions of problems were collected and corrected. The best 10% of the students (final mark in chemistry ≥ 27/30) solved, on average, 62 individual problems (ranging from 37 to 83). Almost all of them tried to solve problems in a creative way. The majority of the remaining students successfully solved at least one problem in a creative way. Several students verified their results according to the ASV method, but few solved all the problems using (the three formal steps of the method) ASV entirely. The time spent in verifying the result was quite valuable, and from the solutions in my collection there is proof that many students understood why and where they made errors and were able to correct themselves. None of the weak students verified the result of their solutions.

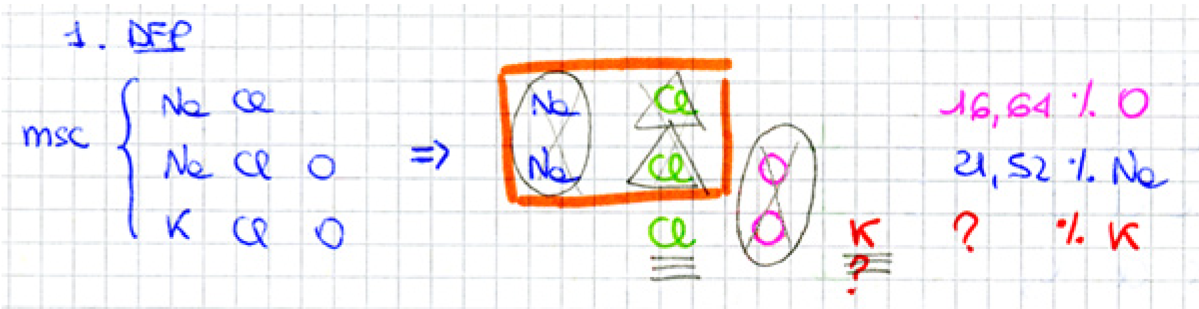
Some examples of creative solutions devised by students

Student 1

A mixture formed by NaCl, NaClO and KClO contains 16.64% of oxygen and 21.52% of Na. Calculate the percentage of K in the mixture (mxt).

1. Formal definition of the problem

She represented the problem by looking at the chemical formulas in this way:



From the total mass, subtracting the oxygen and the sodium, the mass of Cl and K will be obtained. In NaCl and NaClO, Na and Cl are in the same ratio: subtracting Na from (total) Cl, Cl in KClO is obtained because it is in stoichiometric relation with K. She considered 100.0 g of mixture (mxt).

2. Selection of appropriate information

Atomic masses: 16.00 g O/1 mol O; 22.99 g Na/1 mol Na; 35.45 g Cl/1 mol Cl; 39.10 g K/1 mol K.

100.0 g mxt – (21.52 g Na + 16.64 g O) = 61.84 g K + g Cl.

21.52 g Na ≡ 9.361×10^{-1} mol Na ≡ mol Cl in NaCl and NaClO.

3. Combine the various information

g Cl (in NaCl) + g Cl (in NaClO) = (9.361×10^{-1} mol Cl) x (35.45 g Cl/mol Cl) = 33.18 g Cl

(g K + g Cl) in KClO = 61.84 (g K + g Cl) – 33.18 g Cl = 28.66 g

74.55 g KCl : 39.10 g K = 28.66 g : x g K

x = 15.03 g K

Verification

74.55 g KCl : 90.55 g KClO = 28.66 g KCl : x g KClO

x = 34.86 g KClO

34.86 g KClO – 28.66 g KCl = 6.202 g O (in KClO)

16.64 g O – 6.202 g O = 10.44 g O (in NaClO)

16.00 g O : 22.99 g Na = 10.44 g O : y g Na

y = 15.00 g Na ≡ 48.57 g NaClO

21.52 g Na – 15.00 g Na = 6.52 g Na (in NaCl) ≡ 16.57 g NaCl

16.57 g NaCl + 48.57 g NaClO + 34.86 g KClO = 100.0 g mxt

Student 2

Here is another ingeneous solution (Cardellini and Felder, 2004):

22.99 g Na : 21.55 g Na = 35.45 g Cl : x g Cl; x = 33.18 g Cl (in NaCl and NaClO)

100.0 g mxt – 21.52 g Na – 33.23 g Cl – 16.64 g O = 28.66% KCl

74.55 g KCl : 28.66 g KCl = 39.10 g K : y g K; y = 15.03% K

Student 3

Another student developed this interesting and very sophisticated solution. This is a strategy that Marvin Levine (1994) called "*Look at the extremes*".

If the % of NaClO = 0, then 21.52 g Na are all from NaCl and 16.64 g O are from KClO. How many grams of NaCl there are in the mixture?

58.44 g NaCl : 22.99 g Na = x g NaCl : 21.52 g Na; x = 54.70 g NaCl

100.0 g mxt – 54.70 g NaCl = 45.30 g KClO

How many grams of oxygen there are in this mixture?

90.55 g KClO : 16.00 g O = 45.30 g KClO : y g O; y = 8.004 g O (the minimum content of O in the mixture).

If NaCl = 0, then 21.52 g Na are all from NaClO. How many grams of NaClO there are in the mixture?

74.44 g NaClO : 22.99 g Na = z g NaClO : 21.52 g Na; z = 69.68 g NaClO

100.0 g mxt – 69.68 g NaClO = 30.38 g KClO

How many grams of oxygen there are in this mixture?

[(69.68 g NaClO/74.44 g NaClO) + (30.38 g KClO/90.55 g KClO)] x (16,00 g O) = 20.34 g O (the maximum content of O in the mixture).

KClO can vary from 45.30 g to 30.38 g, and this make the variation of oxigen in the mixture from 8.004 g to 20.34 g.

Variation of KClO: 45.30 g – 30.38 g = 14.94 g

Variation of O: 20.34 g – 8.004 g = 12.34 g

(12.34 g O) x (1 g KClO)/(14.94 g KClO) = 8.260 x 10^{-1} g O

16.64 g O – 8.004 g = 8.636 g O

8.636 g O/8.260 x 10^{-1} g O = 10.46 (variation of KClO)

45.30 g KClO – 10.46 g KClO = 34.84 g KClO in the mixture

(34.84 g KClO)/(90.55 g KClO/mol KClO) = 3.848 x 10^{-1} mol KClO

(3.848 x 10^{-1} mol KClO) x (1 mol K/1 mol KClO) = 3.848 x 10^{-1} mol K

(3.848 x 10^{-1} mol K) x (39.10 g K/mol K) = 15.04 g K

The set of problems used in this study can be found in the appendix.

Another original solution for a similar problem, developed by two students who worked cooperatively, has already been reported (Cardellini, 2000b).

Discussion

Students are not obliged to solve the difficult problems that are presented during the course, but know that they can get a bonus if they are able to solve the problem in a way that are judged appropriate, original or new. Almost a third of students solved the problems by their reasoning to find a creative solution. This is an important result, since examples of creative problem solving are reported to be very rare (Treffinger and Ripple, 1971; Mayer, 1992). For the creative students, their GALT scores were equal to or greater than 18. Their disembedding scores were equal to or greater than 15. The creative problem solvers were clearly in the upper part of the distribution of formal reasoning skills and were also able to disembed relevent information from its surroundings in a problem. The majority of these creative students turned out to be very successful in the overall chemistry assessment as well as in their other subjects. Cognitive skills seem to have a role in successful problem solving, but personal motivations may be equally decisive (Hofstein and Kempa, 1985).

Conclusion

The approach to this problem solving experience is based on a constructivist view of learning which was summarized in a single statement by Bodner, "*Knowledge is constructed in the mind of the learner*." (Bodner, 1986) Students are actively engaged in finding the correct solution to the problem proposed and in making sense of the solution. They know that, at the end of the task, one of them will be called to the blackboard to explain the procedures used during that group's solution to the whole class and that a discussion will take place. The role of the teacher is mainly as a coach, a stimulus for reaching new goals. Students are asked to work in what is called the Zone of Potential Development (Brown and French, 1979). By working in a group they solve problems that they would be unlikely to solve alone. The method offers the students an environment in which they can learn and be motivated to learn as much as they can. This can be seen as an example of the experience enjoyed by the student who said, "*They didn't teach me anything, but I learned a lot.*" (Moore, 1996).

References

Bodner G.M., (1986), Constructivism: a theory of knowledge, *Journal of Chemical Education*, **63**, 873-878.

Bodner G.M., (2003), Problem solving: the difference between what we do and what we tell students to do, *University Chemistry Ed*ucation, **7**, 37

Bodner G.M. and McMillen T.L.B., (1986), Cognitive restructuring as an early stage in problem solving, *Journal of Research in Science Teaching*, **23**, 727-737.

Brown A.L. and French L.A,. (1979), The zone of potential development: implications for intelligence testing in the year 2000, *Intelligence*, **3**, 255-273.

Camacho M. and Good R., (1989), Problem solving and chemical equilibrium: Successful versus unsuccessful performance, *Journal of Research in Science Teaching*, **26**, 251-272.

Cardellini L., (1984), Il metodo ASV per risolvere i problemi chimici, *La Chimica nella Scuola*, **6**, 4-6.

Cardellini L., (1999), *Problemi Chimici*, Ragni, Ancona, Italy.

Cardellini L., (2000a), An interview with Alex H. Johnstone, *Journal of Chemical Education*, **77**, 1572.

Cardellini L., (2000b), Calcoli stechiometrici, *La Chimica nella Scuola*, **22**, 20-21.

Cardellini L. and Felder R.M., (2004), Cooperative Learning, *IS Informatica & Scuola*, **12**, 36-39.

Case R., (1975), Gearing the demands of instruction to the developmental capacities of the learner, *Review of Educational Research*, **45**, 59-87.

De Jong T. and Ferguson-Hessler M.G.M., (1986), Cognitive structures of good and poor novice problem solvers in physics, *Journal of Educational Psychology*, **78**, 279-288.

El-Banna H., (1987), The development of a predictive theory of science education based upon information processing theory, University of Glasgow: Ph.D. thesis.

Ennis R.H., (1987), A taxonomy of critical thinking dispositions and abilities. In Baron J.B. and Stemberg R.J. (eds.), *Teaching thinking skills: theory and practice*, W. H. Freeman and Co., New York, p. 10.

Entwistle N., (1988), *Styles of learning and teaching: an integrated outline of educational psychology for students, teachers, and lecturers*, David Fulton, London.

Felder R.M., (1996a), Matters of style, *ASEE Prism*, **6**, 18-23.

Felder R.M., (1996b), Active-inductive-cooperative learning: an instructional model for chemistry?, *Journal of Chemical Education*, **73**, 832-836.

Glaser R. and Chi M.T.H., (1988), Overview. In Chi M.T.H., Glaser R. and Farr, M.J. (eds.), *The Nature of Expertise*, Lawrence Erlbaum, Hillsdale, NJ, p. xv-xxviii.

Hayes J.R., (1988), *The complete problem solver (2nd ed),* Hillsdale, NJ, Lawrence Erlbaum.

Herron J.D., (1986), What can we do about sue: a case study of competence, *Journal of Chemical Education*, **63**, p. 530.

Herron J.D., (1990), Research in chemical education: results and directions. In Gardner, M. et al., (eds), *Toward a scientific practice of science education*, Erlbaum, Hillsdale, NJ, p. 35.

Hofstein A. and Kempa R.F., (1985), Motivating strategies in science education: Attempt at an analysis, *European Journal of Science Education*, **7**, 221-229.

Johnstone A.H., (2001), Can problem solving be taught?, *University Chemistry Education*, **5**, 69-73.

Johnstone A.H. and El-Banna H., (1986), Capacities, demands and processes – a predictive model for science education, *Education in Chemistry*, **23**, 80-84.

Johnstone A.H. and Su W.Y., (1994), Lectures – a learning experience?, *Education in Chemistry*, **31**, 75-79.

Johnstone A.H. and Otis K.H., (2006), Concept mapping in problem based learning: a cautionary tale, *Chemistry Education Research and Practice*, **7**, 84-95.

Johnson D.W., Johnson R.T. and Smith K.A., (1991), *Active Learning: Cooperation in the College Classroom*, Interaction Book Company, Edina, MN, p. 1:18-20.

Larkin J., McDermott J., Simon D.P. and Simon H.A., (1980), Expert and Novices Performance in Solving Physics Problems, *Science*, **208**, 1335-1342.

Lawrence G., (2000), *People types and tiger stripes*, Center for Application of Psychological Type, Gainesville, FL.

Lee K.-W.L., Tang W.-U., Goh, N.-K. and Chia, L.-S., (2001), The predicting role of cognitive variables in problem solving in mole concept, *CERAPIE*, **2**, 285-301. [http://www.uoi.gr/conf_sem/cerapie]

Levine M., (1994), *Effective problem solving*, 2nd Ed., Prentice Hall, Englewood Cliffs, NJ, Sec. 7.

Mayer R.E., (1975), Informations processing variables in learning to solve problems, *Review of Educational Research*, **45**, 525-541.

Mayer R.E., (1992), *Thinking, problem solving, cognition*, Freeman, New York, p. 383.

Moore J.W., (1996), Editorial: they didn’t teach me anything, but I learned a lot, *Journal of Chemical Education*, **73**, A291.

Niaz M., (1987), Relation between M-space of students and M-demand of different items of general chemistry and its interpretation based upon the neo-Piagetian theory of Pascual-Leone, *Journal of Chemical Education*, **64**, 502-505.

Niaz M. and Logie R.H., (1993), Working memory, mental capacity and science education: towards an understanding of the ‘working memory overload hypothesis’, *Oxford Review of Education*, **19**, 511-525.

Novak J.D. and Gowin D.B., (1984), *Learning how to learn*, Cambridge University Press, New York.

Novak J.D., (1998), *Learning, creating, and using knowledge: concept maps™ as facilitative tools in schools and corporations*, Erlbaum, Mahwah, NJ.

Reid N. and Yang M.-J., (2002), The solving of problems in chemistry: the more open-ended problems, *Research in Science & Technological Education*, **20**, 83-98.

Roadrangka V., Yeany R.H. and Padilla M.J., (1983), *The construction and validation of Group Assessment of Logical Thinking (GALT)*, Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Dallas.

Schoenfeld A.H., (1980), Teaching problem-solving skills, *American Mathematical Montly*, **87**, 794-805, p.795.

Stamovlasis D. and Tsaparlis G., (2000), Non-linear analysis of the effect of working memory capacity on organic synthesis problem solving, *Chemistry Education Research and Practice*, **1**, 375-380, [http://www.uoi.gr/conf_sem/cerapie].

Treffinger D.J. and Ripple R.E., (1971), Programmed instruction in creative problem solving, *Educational Leadership*, **4**, 667-675.

Tsaparlis G., Kousathama M. and Niaz M., (1998), Molecular-equilibrium problems: manipulation of logical structure and M-demand, and their effect on student performance, *Science Education*, **82**, 437-454.

Whitehead A.N., (1953), The aims of education and other essays. in Northrop, F.S.C. and Gros, M.W. (eds), *Alfred North Whitehead*, Macmillan, New York.

Witkin H.A., (1974), *Psychological differentiation: studies of development*, Wiley, New York.
Witkin H.A. and Goodenough D.R., (1981), *Cognitive styles: essence and origins. Field dependence and field independence*, International Universities Press, New York.
Wood C., (2006), The development of creative problem solving in chemistry in Glasgow, *Chemistry Education Research and Practice*, **7**, 96-113.
Zoller U., (1993), Are lecture and learning compatible?, *Journal of Chemical Education*, **70,** 195-197.

Appendix

The following problems are ordered from the easiest to the most difficult. Many students solved problem no.1; few solved problem 4 and almost no one tried problem 5.

Relative atomic mass: H=1.008; C=12.01; O=16.00; Na=22.99; S=32.07; Cl=35.45; K=39.10; Mn=54.94.

1. A mixture formed by NaCl, $NaClO_3$ and $KClO_3$ contains 33.40% of oxygen and 16.00% of Na. Calculate the percentage of K in the mixture (Cardellini, 1999; p. 68).

2. A mixture of NaCl, $NaClO_3$ e Na_2SO_4 contains 25.454% of Cl and 36.060% of O. What is the percentage of $NaClO_3$ in the mixture?

3. A mixture of Na_2S, Na_2SO_4 and CH_3COONa weighing 25.19 g, contains 24.29% of S and 21.51% of O. How many grams of CH_3COONa are contained in the mixture?

4. A mixture of CH_4O, C_6H_6 and C_7H_6O weighting 44.37 g gives the elemental analysis: C = 68.74%; H = 8.905%; O = 22.355%. How many grams of C_6H_6 are contained in the mixture? (Cardellini, 1999; p. 90).
5. A mixture contains KClO, $MnCl_2$ and $KMnO_4$. What must the percentage of potassium be to have the chlorine percentage equal to the percentage of oxygen?

Research into practice: visualisation of the molecular world using animations

Roy Tasker and Rebecca Dalton

School of Natural Sciences, University of Western Sydney, Locked Bag 1797, Penrith South DC, NSW, 1797, Australia
e-mail: r.tasker@uws.edu.au; r.dalton@uws.edu.au

Received 21 December 2005, accepted 30 January 2006

Abstract: Most chemistry teaching operates at the macro (or laboratory) level and the symbolic level, but we know that many misconceptions in chemistry stem from an inability to visualise structures and processes at the sub-micro (or molecular) level. However, one cannot change a student's mental model of this level by simply showing them a different, albeit better, model in an animation. Molecular-level animations can be compelling and effective learning resources, but they must be designed and presented with great care to encourage students to focus on the intended 'key features', and to avoid generating or reinforcing misconceptions. One misconception often generated is the perception of 'directed intent' in processes at the molecular level, resulting from the technical imperative to minimise file size for web delivery of animations. An audiovisual information-processing model – based on a combination of evidence-based models developed by Johnstone and Mayer, cognitive load theory, and dual-coding theory – has been used to inform teaching practice with animations, and seed questions for research on student attributes affecting development of mental models using animations. Based on this model, the constructivist *VisChem Learning Design* probes students' mental models of a substance or reaction at the molecular level before showing animations portraying the phenomenon. Opportunities to apply their refined models to new situations are critical. [*Chem. Educ. Res. Pract.*, 2006, **7** (2), 141-159]

Keywords: molecular visualisation; audiovisual information-processing model; animations

Introduction

Chemistry involves interpreting *observable* changes in matter (eg. colour changes, smells, bubbles) at the concrete *macroscopic* or *laboratory level* in terms of *imperceptible* changes in structure and processes at the imaginary *sub-micro* or *molecular level***.** These changes are then represented at an abstract *symbolic level* in two ways: qualitatively, using specialized notation, language, diagrams, and symbolism; and quantitatively, using mathematics (equations and graphs).

Figure 1 illustrates these three levels for an iron(III) thiocyanate equilibrium. The apparently unchanging solution in the beaker at the laboratory level can be linked to an animation portraying the dynamic, invisible processes at the molecular level. The equation represents the equilibrium at the symbolic level. Our initial hypothesis, later revised from our studies described below, was that simply showing an animation of this equilibrium at the molecular level might help students build a better conceptual understanding of what it means for a system to be at equilibrium, and to interpret the real meaning of the double arrows.

Figure 1. Chemical equilibrium presented at the three thinking levels.

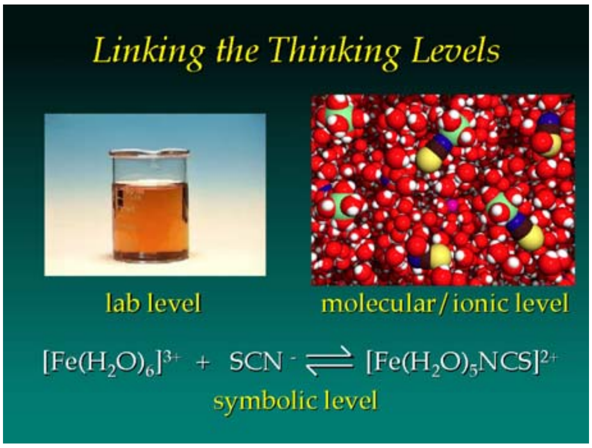
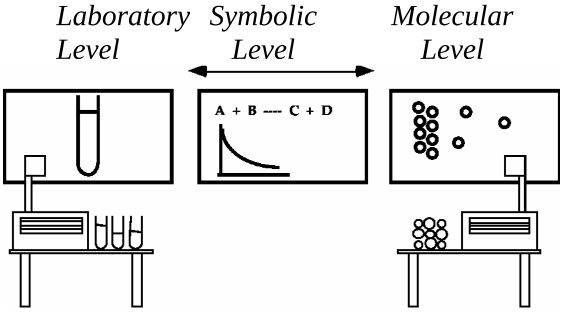


Figure 2. Using the three thinking-level approach in the lecture theatre. This approach is reinforced in the laboratory notes, tutorials and assessment.



The need to be able to move seamlessly between these three 'thinking levels', first described by Johnstone (1982, 1991), is a major challenge for students learning chemistry (Kozma, 1997). One of the authors (RT) first used these levels *explicitly* as a teaching strategy in the late 1980s (Tasker, 1992), allocating each part of the lecture stage to a thinking level (Figure 2), rewriting laboratory manuals specifying when each level is relevant, and designing exam questions to probe a student's ability to integrate laboratory work and theory at each level. Other researchers have recommended teaching with these levels in an explicit way, and helping students to draw links between them (Tasker, 1996, Russell, 1997; Hinton, 1999). Now almost every general chemistry textbook (e.g., Bell, 2005) mentions this presentation strategy in early chapters, whilst few reinforce the idea throughout the text.

Need to develop mental models of the molecular level

Since the mid-1970s there has been convincing evidence (Kleinman, 1987; Lijnse, 1990, and references therein) that many student difficulties and misconceptions in chemistry result from inadequate or inaccurate mental models at the molecular level. Moreover, many of the misconceptions are common to students all over the world, and at different educational levels, and even amongst students who were performing well in formal examinations (Nurrenbern, 1987; Nakhleh, 1992, 1993a; Nakhleh, 1993b; Niaz, 1995). The most important finding was that many misconceptions were extraordinarily resistant to change, despite targeted teaching interventions.

Until the early 1990s there was a shortage of convincing resources that portrayed the dynamic molecular level with sufficient accuracy to help students to construct useful mental models of structures and processes at this level. Most teaching was restricted to the laboratory and symbolic levels, in the hope that students' models of the molecular world would 'develop naturally'. Students were left to construct their models from the static, often oversimplified, two-dimensional diagrams in textbooks; confusing ball-and-stick models; and their own imaginative interpretation of chemical notation - for example, does "NaCl(aq)" mean that ionic solutions contained dissolved 'NaCl molecules'?

The purposes of this paper are to:

- show how the Johnstone three "thinking-level" model acted as the seed for the *VisChem* project to assist students to construct useful mental models at the molecular level
- describe research on the effectiveness of the *VisChem* animations, and the need to embed them within a 'learning design'
- present an audiovisual information-processing model, based on work by Johnstone, Mayer, Paivio, and Sweller, to inform the development of learning designs
- illustrate one such design – the *VisChem Learning Design* – with an example for developing a student's molecular-level model for ions in aqueous ionic solutions.

The VisChem project – visualising the molecular level with animations

Motivated by a frustration with the lack of resources in the early 1990s depicting Johnstone's sub-micro level, the *VisChem* project was funded to produce a suite of molecular animations depicting the structures of substances and selected chemical and physical changes (Tasker et al., 1996; also see vischem.cadre.com.au for availability). The animations were produced as useful models at this level, with careful attention to the often-competing demands of scientific accuracy (e.g., close proximity of adjacent molecules in the liquid state; internal molecular bond vibrations; and the diffuse nature of electron clouds), the 'artistic license' required for clear communication (e.g., reduced speed of molecules in the gaseous state; less crowding in the liquid state; absence of internal molecular bond vibrations; and use of shiny boundary surfaces), and technical computing constraints (close-up view to limit the number of moving 3D objects to be rendered; and the directed depiction of portrayed events to minimise the number of animation frames, and file size). Resources were then developed to link these animations to the macro and symbolic levels.

What kinds of messages can be communicated in molecular-level animations?

The molecular world is multi-particulate, dynamic and, in the liquid state, crowded; and the interactions are often subtle (e.g., electron transfer) and complicated. Animations can be effective for helping students to construct and apply useful mental models of this world. However, as we will see below, effective use of these resources for meaningful learning should be based on a learning theory that is evidence-based, and able to inform teaching practice. Some of the *VisChem* animations described below can be downloaded (Tasker, 2002b – http://www.learningdesigns.uow.edu.au/exemplars/info/LD9/index.html, and go to the molecular construction tool and animations 'Crosslink').

In contrast to textbook illustrations, animations can show the dynamic, interactive, and multi-particulate nature of chemical reactions explicitly. For example, the laboratory-level observation of silver crystals growing on the surface of copper metal shown in Figure 3 is hardly consistent with the misleading diagram, often written on a whiteboard, of one copper atom donating an electron to each of two silver ions. An animation (Figure 4) can show reduction of *many* silver ions on the copper surface, with concomitant release of *half as many* copper(II) ions from the metal lattice. This is a much better explanation for the 2:1 stoichiometric ratio in this reaction.

Figure 3. When copper metal is covered with silver nitrate solution, silver crystals form on the surface of copper metal; some copper 'dissolves', and the solution gradually turns blue.

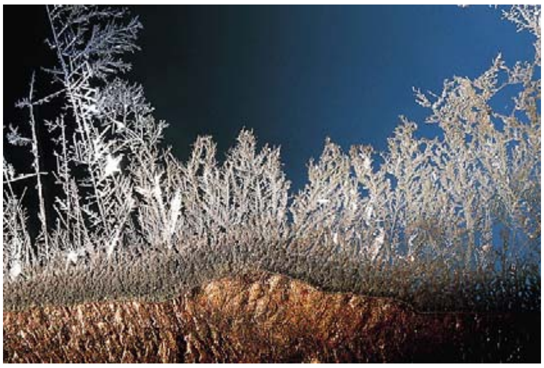
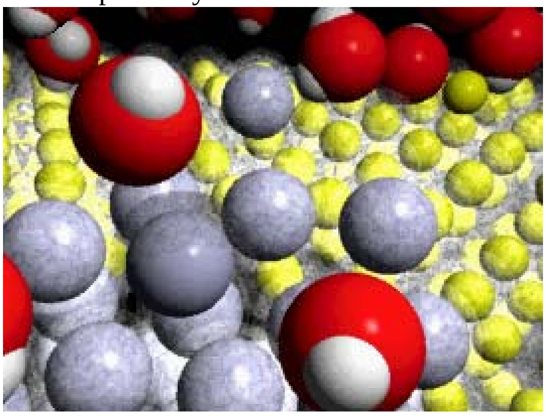


Figure 4. Frame from a *VisChem* animation showing reduction of silver ions to silver atoms on a growing crystal; with concomitant release of copper(II) ions, in a two to one ratio respectively.



Animations of the molecular world can stimulate the imagination, bringing a new dimension to learning chemistry. What could it be like inside a bubble of boiling water, or at the surface of silver chloride as it precipitates, as depicted in Figures 5 and 6 respectively?

Figure 5. A frame of the *VisChem* animation that attempts to visualise gaseous water molecules 'pushing back' the walls of a bubble in boiling water.

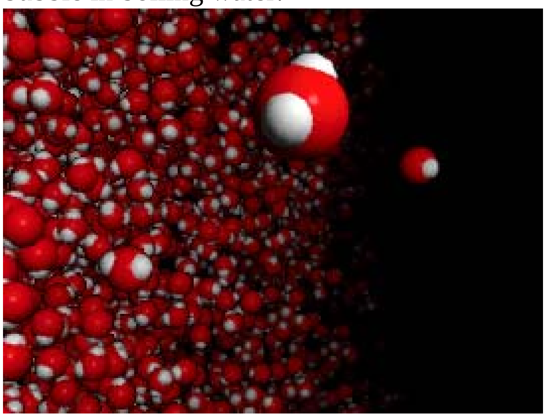
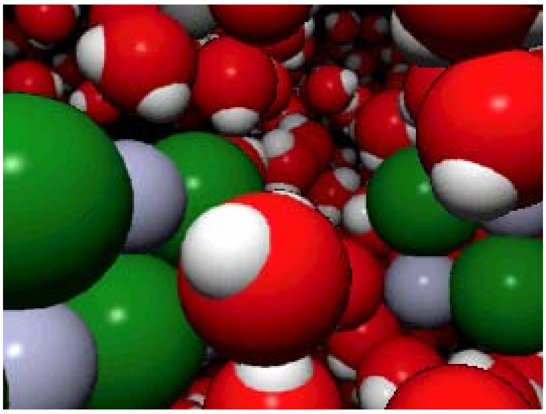


Figure 6. A frame from another animation that depicts the precipitation of silver chloride at the molecular level.



Most molecular-level processes involve competition. Examples include the competition for a proton between an iron(III)-bound hydroxide and a solvent water molecule (Figure 7); and between lattice forces and ion-dipole interactions when sodium chloride dissolves in water (Figure 8).

Figure 7. Frame from a *VisChem* animation showing the 'tug-of-war' between an iron(III)-bound hydroxide and a solvent water molecule for the proton.

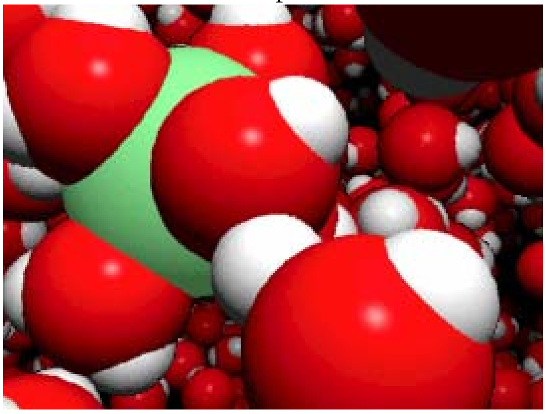
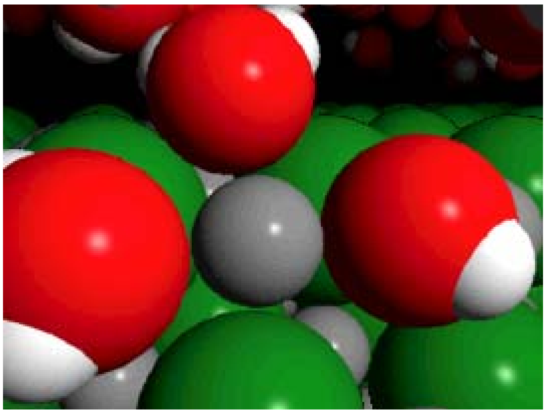


Figure 8. Frame from a *VisChem* animation showing the hydration of a sodium ion on the surface of sodium chloride despite strong attractive forces from the rest of the lattice.



Research on the effectiveness of VisChem animations for constructing mental models

We have conducted research into factors that affect a student's ability to form scientifically acceptable mental models of chemical substances and processes at the molecular level after exposure to *VisChem* animations. Our study examined the changes in mental models of first-year chemistry students (N = 48) following a semester of teaching that emphasized molecular visualisation using the animations (Dalton, 2003). The study used a pre-test/post-test design with follow-up interviews. A transfer-test was also administered after the post-test, and prior to interviews. The animations were presented on the basis of recommendations in the literature (Milheim, 1993) and practical experience over five years of using the animations in lectures (see vischem.cadre.com.au, and go to Educational Support/Resources).

This study demonstrated that showing animations to students, with opportunities for them to practise drawing representations of the molecular world, significantly increased the number of scientifically acceptable 'key features' in students' representations of chemical phenomena at the end of the semester (Table 1). Students developed more vivid mental imagery of these phenomena (Table 2) and had greater confidence in their images (Table 2). Evidence from interviews with fourteen students revealed, without prompting, that changes were largely attributed to having viewed *VisChem* animations. There was also an indication that some students had been able to transfer their ideas from animations to new situations, as evidenced by a statistically-significant correlation between the post-test and transfer test (n = 35, r = 0.69, p = 0.01), and comments in interviews.

Table 1. Means and standard deviations for corresponding sections on the pre-test and post-test (N = 48).

	Pre-test		**Post-test**	
Section	**M**	**SD**	**M**	**SD**
Molecular Substances: General Features	6.2	2.4	8.0*	1.8
Molecular Substances: Specific Features of Water	1.7	0.9	2.2*	0.8
Ionic Solid	2.5	1.5	3.8*	1.6
Ionic Solution	2.0	1.9	5.7*	1.5
Test Total	12.4	5.1	19.8*	4.0

* $p \leq 0.001$, one-tailed paired t-test

Table 2. Means and standard deviations for confidence and imagery vividness scales in the pre-test and post-test.

	Pre-test		**Post-test**	
Scale	**M**	**SD**	**M**	**SD**
Confidence (N = 30)	3.5	1.3	4.8*	0.7
Imagery Vividness (N = 42)	3.4	1.4	4.8*	0.9

* $p < 0.001$, Wilcoxon matched-pairs signed-ranks test

A longitudinal study of student reflections on VisChem animations

A study with third-year university chemistry students (N = 30) provided evidence that these benefits appeared to persist throughout a chemistry degree. These students demonstrated long-term recall of *VisChem* animations, and some felt that exposure to these resources helped them with various concepts, topics and subjects throughout their degrees. Other benefits identified by these students are outlined in Table 3.

Table 3. Possible benefits of instruction with *VisChem* animations.

Benefit (No. of Students)	**Quote**
Visualisation (14)	*"….. helped in developing visual images"*
Movement and interactions (13)	*"They helped me visualise interactions of different molecules."* *"It is a constantly moving environment."*
Understanding (10)	*"…. increased my ability to understand concepts."*
Made learning easier (8)	*"…. visualising made chemistry easier."*
Interpretation of laboratory-level phenomena (8)	*"You can imagine the molecules, atoms, structures, on a molecular level instead of just a macro level."*
Aroused interest/curiosity (6)	*"….. increased interest in chemistry."*
Improved thinking in 3D (4)	*"Animations helped me think in a 3-dimensional way with all molecules."*
Good foundation for future learning (4)	*"Understanding of first-year chemistry ideas and concepts made a good foundation for years that followed."*
Application to new situations (2)	*"The animations give a guide that can be applied in other situations, it's the visualising and thinking about them which made them most useful."*

These students also revealed some of the possible limitations of *VisChem* animations. For example, two students said that because

"The visual communication (animations) was not supplemented in further years …." they were not as beneficial as they could have been.

Overall, our results indicated that *VisChem* animations can encourage and aid students to develop mental pictures of the molecular level that are multi-particulate, dynamic, interactive and three-dimensional (Dalton, 2003).

Need for simulations to complement and supplement animations

In criticising the *VisChem* animations, one astute student indicated that they were misleading because they appeared to portray chemical reactions as mechanical and deterministic processes, lacking the element of randomness:

"This animation [portraying precipitation of AgCl] ... *shows water molecules ... sort of carrying this structure* [AgCl ion pair] *along … like a bunch of little robots … The animation depicts something that … I think really happens by chance, as a very deliberate and deterministic sort of process and I think that's slightly misleading … Surely it must be possible to make it look less deliberate, less mechanical, maybe by showing ... the odd one or two going into the structure but not all of them."*

This student pointed out an important limitation of most chemistry animations. Technical constraints to reduce rendering times, and minimising file size to enable rapid delivery over the web, have resulted in animations that convey the clear perception of 'directed intent' in molecular-level processes, instead of a more scientifically-accurate, probabilistic model.

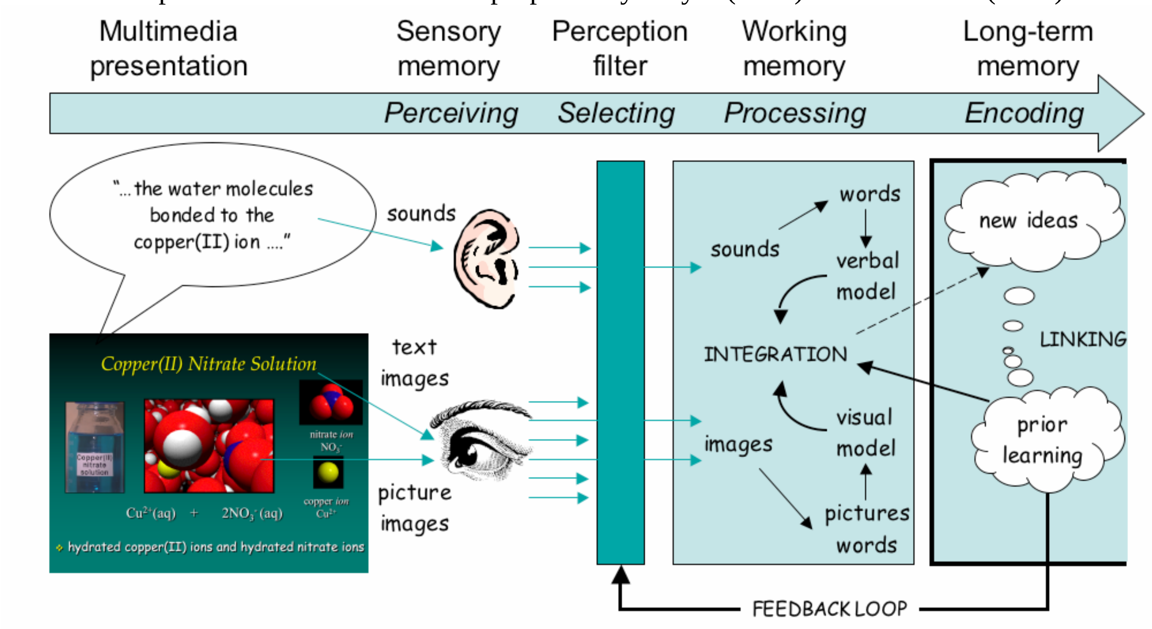
In contrast to choreographed animations, theory-driven simulations (e.g., *Odyssey* by Wavefunction, Inc.; see wavefun.com) offer a more accurate depiction of structures and processes at the molecular level. However, a limitation of simulations is that they often do not show key features of molecular events clearly because they occur rarely (sometimes taking years in the slowed-down timescale used), at random, and usually with intervening solvent molecules blocking the view! Clearly simulations and animations can complement one another.

Even with judicious use of animations and simulations, some students continue to retain poorly formed ideas and harbour misconceptions. We wanted to know how student attributes contributed to whether or not students were able to develop their mental models via the use of visualisation tools. In order to pose useful research questions we used an audiovisual information-processing model based on the wealth of research on how the brain attends to, processes, stores, and retrieves audiovisual information.

An audiovisual information-processing model

The model we used (Figure 9) is a composite of the established models developed by Johnstone (1986) and Mayer (1997), together with ideas from dual-coding theory (Paivio, 1990), and cognitive load theory (Sweller, 1988; 1994). Johnstone has used his model to inform all forms of chemistry teaching (e.g., Johnstone, 1994), and Mayer's model has been used successfully to derive instructional design principles for multimedia explanations (Moreno, 2000).

Figure 9. A multimedia information-processing model for learning from audiovisual information. This is a composite of theoretical models proposed by Mayer (1997) and Johnstone (1986).



The model describes learning in terms of an audiovisual information processing system that involves perceiving verbal and visual stimuli in separate parts of the sensory memory; selection through a filter; integration and processing of the verbal and visual information within a working space of limited capacity; and storage of this information in the long-term memory (LTM), for efficient retrieval and transfer to new situations.

Just as Johnstone's information-processing model has implications for good teaching practice, this embellished model has implications specifically for presentation of audiovisual information, and some of these have been supported with experimental data (Sweller 1994; Moreno 2000).

For example:

- students should be given manual control over the pace and content in animations (e.g., by dragging the Play bar back and forth, and pausing where appropriate, rather than just clicking on the Play button. This reduces the rate of information load presented, and provides time for cause-and-effect reflection
- verbal and visual information should co*mplement* one another, not *supplement* one another, as this risks overloading the working memory space
- text should be presented within graphics rather than separately as captions, and animations presented with simultaneous, rather than separate narration

A significant result from this research is that the working memory capacity can be expanded slightly by mixing the senses used to present information. That is, it is easier to process information when some is presented visually, and the remainder is presented auditorily, than it is when all the information is presented through a single sense – either all visually or all auditorily (Sweller, 1994). This provides a strong argument for the use of narration in animations, rather than using text captions. The latter practice is often done to minimise the animation file size for web delivery, but there is a cost in effectiveness for learning.

This model, with all its implications for teaching practice with animations, provided us with the basis to identify factors that should influence learning with animations.

Research to identify student attributes influencing effectiveness of VisChem animations

A preliminary study was conducted to examine the student attributes affecting the development of students' mental models when *VisChem* animations were used. Factors to examine were selected on the basis of the different stages in the audiovisual information-processing model (Figure 9).

Ability to perceive details in visual displays (disembedding ability) was measured using the *Group Embedded Figures Test* (Witkin, 1971). Visuospatial working memory capacity was measured using the *Figural Intersection Test* (Pascual-Leone, 1969; Johnson, 1982). Relevant prior knowledge held in LTM that might influence perception, was identified using a pre-test. An idea of the extent to which students attempted to relate new information to old, and structure information in LTM, was determined using a modified version of the *Study Processes Questionnaire* (SPQ; Prosser, 2000). This version includes surface and deep factors, but not the achieving factor from the original SPQ. The ability to retrieve key features of molecular structures and processes from LTM was measured using a post-test (equivalent to the pre-test).

Interpretation of the results of a multiple regression analysis (N = 22) suggested that prior knowledge, disembedding ability and deep and surface learning had a significant effect on the development (gain from pre-test to post-test) and the sophistication (post-test score) of students' mental models (Table 4 and Table 5). Note that prior knowledge correlated negatively with gains from pre-test to post-test, suggesting that students with low prior knowledge in fact learnt more from the instruction than those with high prior knowledge. This is perhaps not surprising considering that students with low prior knowledge had less-developed images before instruction and, therefore, more potential for progress.

In addition, visuospatial working-memory capacity was shown to correlate significantly with post-knowledge (r = 0.59, p = 0.05, N = 13). A follow-up study, replicating aspects of the original study, confirmed the role of prior knowledge, and to a lesser extent disembedding ability, in the sophistication and development of students' mental models. The other factors (surface and deep learning styles) were not examined in the second study.

Table 4. Correlation between pre-/post-test gain and student attributes.

Independent Variable: Pre-/Post-Test Gain (N = 22)			
Variable	Standardised coefficient (β)	Significance (p)	Adjusted R^2
Prior knowledge	-0.64	< 0.005	0.45
Disembedding ability	0.64*	< 0.005	
Surface learning	-0.74	< 0.05	
Deep learning	0.93	< 0.005	

* Disembedding ability data were transformed (reflect and $\log_{10}$) to approach normality. Reflecting the scores reverses their order; hence the sign of ß has been changed to reflect the true directionality of the relationship between the variables.

Table 5. Correlation between post-test score and student attributes.

Independent Variable: Post-test (N = 22)			
Variable	Standardised coefficient (ß)	Significance (p)	Adjusted R^2
Prior knowledge	0.58	< 0.0001	0.76
Disembedding ability	0.43*	< 0.005	
Surface learning	– 0.49	< 0.05	
Deep learning	0.62	< 0.005	

* Disembedding ability data were transformed (reflect and $\log_{10}$) to approach normality. Reflecting the scores reverses their order; hence the sign of ß has been changed to reflect the true directionality of the relationship between the variables.

In summary, the highest post-test scores were obtained by students with high prior knowledge, high disembedding ability and high visuospatial working-memory capacity, who adopted deep-learning strategies and limited their use of surface learning strategies. Greatest gains (from pre-test to post-test) were achieved by students with low prior knowledge who had high disembedding ability and used deep-learning strategies not surface learning strategies.

In terms of the audiovisual information-processing model, we propose the following interpretation of our results:

- animations encourage a student with low prior knowledge to develop new ideas in LTM to create their mental models
- high prior knowledge in the LTM allows a student to *perceive* subtle but relevant features in an animation enabling development of more sophisticated mental models
- high prior knowledge also enables comparison of an image created in working memory from viewing an animation, with an existing mental model in LTM, leading to confirmation or modification of the existing mental model
- high disembedding ability allows a student to perceive the desired key features in a ‘busy’ animation
- high working-memory capacity ensures a student is able to manage the information from complex animations effectively, and construct and manipulate mental models of the phenomena
- adoption of deep-learning strategies and not surface learning approaches enables a student to relate ‘key features’ in animations to models in the LTM for deep understanding.

Unfortunately the sample size for this study was too small to extract convincing statistical significance (N = 22 for multiple regression, N = 13 for visuospatial working memory capacity correlation), and for this reason the results cannot be generalized. However, the fact that these factors have been reported in the literature as having an influence on other aspects of student learning adds weight to our findings. In general, our results support the value of the audiovisual information-processing model for predicting factors influencing student learning, and the importance of considering each aspect of the model when constructing learning designs.

This research implies that:

- prior knowledge should be revealed so that animations can be presented at a level appropriate to build on this knowledge
- key features in animations must be highlighted in some manner to ensure that students are able to extract the visual information
- students should be encouraged to adopt a deep approach to learning in order to make sense of the features in animations

- animations should be designed and presented so as to minimise extraneous cognitive load.

These recommendations were implemented within commercial constraints in a series of interactive multimedia projects associated with various chemistry textbooks that used the *VisChem* animations (Tasker 1999, 2001, 2002a, 2003, 2004).

Research into teaching practice: the *VisChem* Learning Design

In the growing field of interactive multimedia, a 'learning design' is a research-based sequence of 'learning activities', each involving one or more 'learning objects'. Learning objects are digital assets (e.g., an animation, photograph) in a context (provided by a narration, caption), designed usually with interactivity. The audiovisual information-processing model, with its implications for good practice, and Johnstone's three 'thinking-level' model, informed the development of the constructivist *VisChem Learning Design* (Tasker, 2002b). The design is described below, and then illustrated with an example. More details can be found on the web site for the design (Tasker, 2002b). This is one of a collection of exemplary ICT-based learning designs selected by a panel of Australian university educators to facilitate the uptake of innovative teaching and learning approaches in Australian universities (Harper, 2001).

The *VisChem Learning Design* can be used for any chemistry topic that requires a scientifically acceptable mental model of the molecular world. A typical learning experience in a face-to-face lecture context would involve students:

- *observing* a chemical phenomenon (chemical reaction or property of a substance) as a lecture demonstration, lab activity, or audiovisual presentation; and *documenting* their observations in words and/or diagrams
- *describing* in words, and *drawing* a representation of what is occurring at the molecular level to account for the observations; with the lecturer explaining the need for drawing conventions (eg. to indicate relative size, movement, number, and crowding of molecules)
- *discussing* their representation with a peer, with the aid of the lecturer's advice to focus on the key features of the representation that explain the observations
- *viewing* an animation portraying the phenomenon at the molecular level, first without, then with narration by the lecturer, and looking for key features that might explain the observations
- *reflecting* with the peer on any similarities and discrepancies between their own representations and the animation, and then discussing these with the lecturer
- *relating* the molecular-level perspective to the symbolic (eg. equations, formulas) and mathematical language used to represent the phenomenon
- *adapting* their mental model to explain a similar phenomenon with an analogous substance or reaction

The key criteria for the success of this design to promote visualisation as a learning strategy are the:

- constructivist approach that encourages the student to articulate prior understanding, and focus attention on key features of the prior mental model at the molecular level, *before* seeing the animations
- opportunity to discuss ideas and difficulties with peers
- practice and application of the visualisation skills developed, with the explicit expectation that these skills are valued and would be assessed

The learning outcomes are to assist students to:

- construct scientifically acceptable mental models of substances and reactions at the molecular level
- relate these models to the laboratory and symbolic levels in chemistry
- apply their models to new substances and reactions
- use their models to understand new chemistry concepts that require a molecular-level perspective
- address common misconceptions identified in the research literature
- improve their confidence in explaining phenomena at the molecular level
- enhance their *enjoyment of chemistry by empowering them to use their imagination to explain phenomena*, instead of just rote-learning terms and concepts, and solving problems algorithmically.

An example of the VisChem Learning Design: visualisation of an aqueous ionic solution

The learning design is an attempt to make each stage of the audiovisual information-processing model—perceiving, selecting, processing and encoding—as efficient as possible. In the following example, we will assume that students have had previous experience with visualising simple substances—ionic compounds and water—using *VisChem* animations, and are familiar with graphic conventions for representing molecules and ions.

One important learning outcome of this example is for the student to visualise an ionic solution in terms of moving *hydrated* ions that occasionally form transient ion pairs. The most common misconceptions are that the ions do not interact with the solvent and, more seriously, are clustered together in their 'ionic formula units'. These misconceptions pose significant problems for students understanding solution stoichiometry (e.g., in a solution of 0.1M Na_2SO_4 the $[Na^+]$ is 0.2M) and other related concepts such as colligative properties.

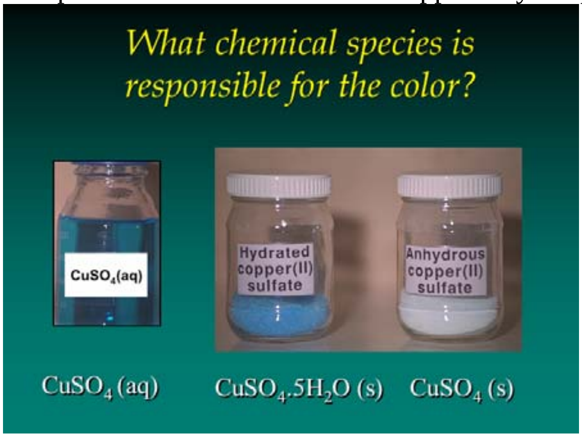
The learning design starts with a simple, but interesting observation.

Step 1. Observing a phenomenon

In the first step of the design students write observations for a laboratory-level chemical phenomenon such as a physical property of a substance (e.g. a metal conducts electricity), or a reaction between substances (e.g. precipitation of an ionic compound). One can present this phenomenon as a live demonstration, or with video, but one must ensure that *all relevant observations* are contributed by students.

Ideally, the phenomenon should be unusual, or counter-intuitive. For example, solid hydrated copper(II) sulfate and aqueous copper(II) sulfate solution are both light blue, but solid anhydrous copper(II) sulfate is white (Figure 10). The question is – what is/are the chemical species responsible for the blue colour?

Figure 10: An example of an observation with an apparently simple explanation.



At this point the instructor should allow students to think about and discuss this observation before rushing in with an explanation. An immediate, but incorrect suggestion might be that, since all bottles contain copper(II) ions and sulfate ions, and only the bottles containing water are blue, perhaps water alone is responsible for the blue colour. A moment's thought tells the student that this volume of pure water is colourless, so the answer must be more interesting!

The aim of this step in the design is to capture attention with an engaging context, and to generate a 'need to know'. In terms of the audiovisual information-processing model, the attention centres in the brain are being activated to select relevant aspects of visual and verbal information from the eyes and ears.

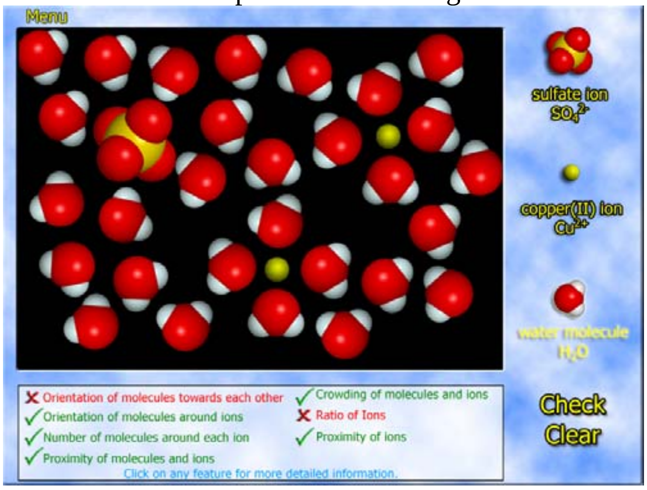
Step 2. Describing and drawing a molecular-level representation

In this step students attempt to explain their observations by drawing *labelled molecular-level representations* of the substance or reaction, and also describe their ideas in words. One needs to develop the 'drawing literacy' of the students by discussing conventions (e.g., representing relative sizes of atoms and ions, using space-filling or ball & stick models), and point out that they will have to do such drawings as part of formal assessment. This is a signal that communicating the details of one's model of the molecular level is a skill worth developing.

At this point ask the students to represent their mental models of the chemical samples in all three bottles to their peers. This should be done in *both* words *and* diagrams to cater for students with a preference for expressing their ideas verbally or visually. With respect to the blue colour, perhaps there is an interaction between the ions and the water molecules in the blue solid and solution?

An alternative to drawing a representation of copper(II) sulfate solution is to use the *Molecular Construction Tool* (a free, downloadable program from the *VisChem Learning Design* web site – Tasker, 2002). The advantage of the tool is the progressive feedback available on the student's representation at any stage of the construction process (Figure 11).

Figure 11. A sample screen from the *VisChem Molecular-Level Construction Tool* (Tasker, 2002) showing feedback on seven key features. In this student's construction the feedback generated shows the ion ratio is incorrect, and too many water molecules are oriented incorrectly for optimum H-bonding.



In terms of the audiovisual information-processing model, the aim of this step in the design is to recall prior knowledge to prime the students' perception filters to focus on the key features of their own mental models.

Step 3. Discussing with peers

Following the advice to students to identify *key features* that explain the observations, they should receive initial feedback on their representations by discussion with peers (or from the *Molecular Construction Tool*). One should not identify correct or incorrect key features at this stage.

The feedback in the Tool is not designed to replace this discussion, but to focus attention on the seven key features of the representations that relate to crowding, proximity of molecules and ions, and ion hydration. At this point student attention will have hopefully been drawn to the key features, now priming the perception filter for selecting relevant verbal and visual information from the molecular-level animations that follow.

Step 4. Viewing animations and simulations

Animations and simulations can depict the dynamic molecular world more effectively than static pictures and words because students are spared the cognitive load of having to 'mentally animate' the content. However, animations are *only effective if they are presented in a way that takes account of the limitations and processing constraints of the working memory.*

In this example, two animations depicting copper(II) nitrate solution would be presented. One animation shows all the water molecules and hydrated Cu^{2+} and NO_3^- ions (Figure 12), the other shows only the hydrated ions (Figure 13). Both animations are 'busy' and, without prior experience with similar animations, the cognitive load on the working memory would be too high. However, since student attention should be focused at this point on searching for something new, they should *perceive* the hydration around the ions.

Figure 12. Frame from the *VisChem* animation portraying the hydrated ions in copper(II) nitrate solution.

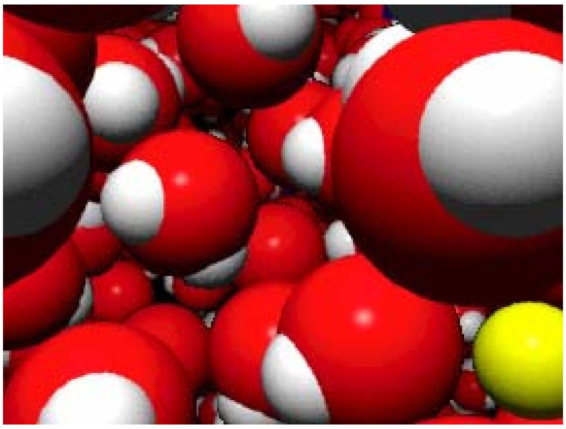
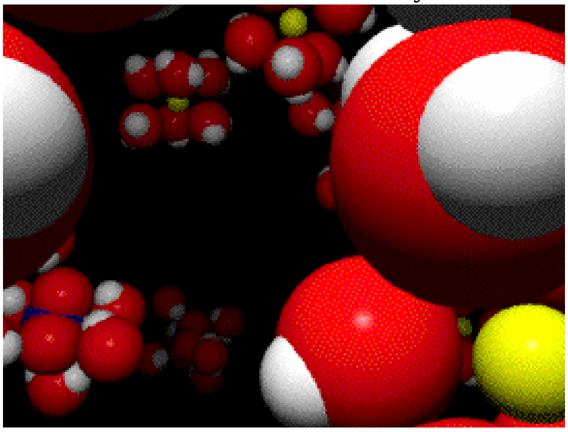


Figure 13. Frame from another *VisChem* animation portraying the hydrated ions in copper(II) nitrate solution, with the solvent water molecules removed for clarity.



Time permitting; each animation should be presented three times:

- First, without commentary, with students encouraged to look for key features they had, or did not have in their own representations.
- Second, in animation stages ('chunks' to reduce the load on working memory), each with narration by the lecturer drawing attention to the important key features, and with responses to any questions from students.
- Third, in its entirety again, with repeated, simultaneous narration.

Step 5. Reflecting on any differences with prior conceptions

In this step students reflect on differences between key features in the animations and in their own representations; *amending their drawings accordingly, if necessary.* Student drawings and descriptions of their conceptions of structures and processes at the molecular level often reveal misconceptions not detectable in conventional equation-writing questions. This activity in the learning design provides the opportunity for students to identify these misconceptions in their own representations, or those of their peers. *Experience shows this is more effective than having the lecturer simply listing common misconceptions.*

Step 6. Relating to other thinking levels

In this step one should encourage student discussion to link the key features of the molecular-level animations to the other two thinking levels. In this example, the following questions would be useful:

Laboratory Level

- Can you see a relationship between the blue colour and hydration of the copper(II) ions? If so, are the copper(II) ions bonded to water molecules in solid hydrated copper(II) sulfate?
- Symbolic Level
- Calculate the ratio of (Cu^{2+} ions : NO_3^- ions : H_2O molecules) in a 1 M copper(II) nitrate solution. This enables students to visualise the term 'concentration' in terms of 'crowding', and to give some meaning to '1 M' compared to the concentration of pure water (1000 g/L = 55.6 M). The answer is 1 : 2 : 56.
- How many water molecules, on average, are there between the ions in a 1 M copper(II) nitrate solution? This requires students to think of about 56 water molecules in a cube including one hydrated Cu^{2+} ion and two hydrated NO_3^- ions. The answer is about two or three water molecules.

In terms of the audiovisual information-processing model, we are trying to link their new insight from the animations to their prior knowledge.

Step 7. Adapting to new situations

In order to extend the links within the LTM, students are asked to draw a molecular-level representation for an analogous substance or reaction shown at the laboratory level. This establishes whether the students can transfer their ideas to a new example.

We have found that if visualisation is to be taken seriously by students as a learning strategy, it is essential that they are encouraged to practise their new skills with new situations, and assess their visualisation skills in one's formal assessment. In addition to questions that probe qualitative and quantitative understanding of concepts at the symbolic level, we need to design questions that require students to articulate their mental models of molecular-level structures and processes.

One reason student misconceptions at the molecular level are not detected at college level is that questions rarely probe this level of understanding explicitly. A good example of how one can probe deep understanding of difficult chemistry concepts by thinking at the molecular level is illustrated in Figure 14. This question probes whether the student has a molecular-level perspective of the difference between acid strength, acid concentration, and acidity (indicated by pH), in contrast to an algorithmically rehearsed expression in terms of mathematical functions.

Conclusion

The need for a chemistry student to move seamlessly between Johnstone's three 'thinking-levels' is a challenge, particularly for the novice. Our work in the *VisChem* project indicates that animations and simulations can communicate many key features about the molecular level effectively, and these ideas can link the laboratory level to the symbolic level. However, we have also shown that new misconceptions can be generated.

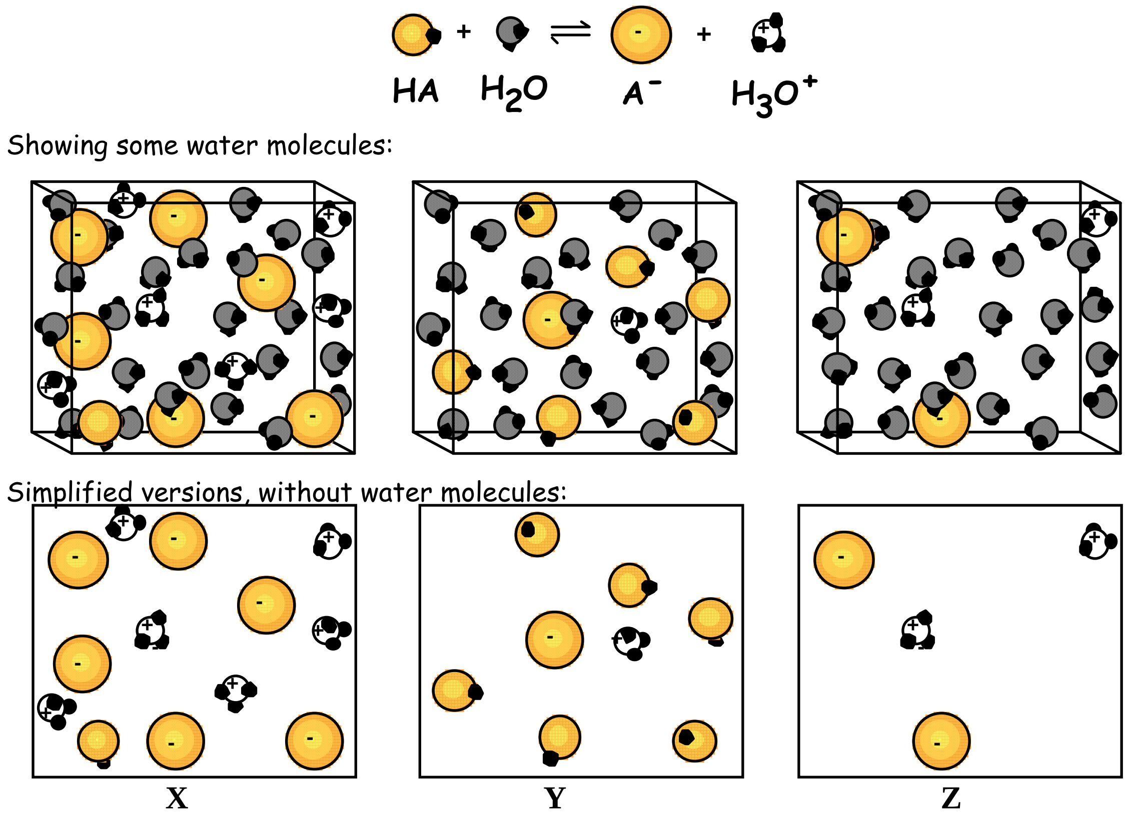
To use animations effectively, we need to direct our students' attention to their key features, avoid overloading working memory, and promote meaningful integration with prior knowledge. We can do this by using constructivist learning designs that exploit our knowledge of how students learn. The audiovisual information-processing model in this paper, based on the work of Mayer and Johnstone, can guide us in developing effective learning designs for this purpose.

'Scarring' misconceptions are those that inhibit further conceptual growth. To identify these misconceptions we need a strategic approach to assist our students to visualise the molecular level, and assess their deep understanding of structures and processes at this level.

Figure 14. A question to probe whether a student understands the difference between acid strength, acid concentration, and acidity at the molecular level.

Compare the diagrams (**X**, **Y**, and **Z**) below and match each diagram to an acid solution (**A**, **B**, or **C**) described in the following table. *Explain your reasoning*.

Solution	Acid	Concentration	K_a
A	trichloroacetic acid, CCl_3COOH	0.010 M	3.0×10^{-1}
B	chlorous acid, $HClO_2$	0.035 M	1.0×10^{-2}
C	benzoic acid, C_6H_5COOH	0.035 M	6.5×10^{-5}



References

Bell J., (2005), *Chemistry – a project of the American Chemical Society,* WH Freeman & Co., New York.

Dalton R.M., (2003), *The development of students' mental models of chemical substances and processes at the molecular level,* PhD Thesis, University of Western Sydney.

Harper B., O'Donoghue J., Oliver R. and Lockyer L., (2001), New designs for Web-Based Learning environments. In C. Montgomerie, & J. Viteli (Eds.), *Proceedings of ED-MEDIA 2001, World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 674-675), Association for the Advancement of Computing in Education, Tampere, Finland.

Hinton M.E. and Nakhleh M.B., (1999), Students' microscopic, macroscopic, and symbolic representations of chemical reactions, *The Chemical Educator*, **4**, 1-29.

Johnson J., (1982), Figural intersection test (fit): a measure of mental (m) attentional energy, *Unpublished manuscript*, York University, Ontario.

Johnstone A.H. and El-Banna H., (1986), Capacities, demands and processes – a predictive model for science education, *Education in Chemistry*, **23**, 80-84.

Johnstone A.H., (1982), Macro and microchemistry, *School Science Review*, **64**, 377-379.

Johnstone A.H., (1991), Why is science difficult to learn? Things are seldom what they seem, *Journal of Computer-Assisted Learning*, **7**, 701-703

Johnstone A.H., Sleet R.J. and Vianna J.F., (1994), An information processing model of learning: its application to an undergraduate laboratory course in chemistry, *Studies in Higher Education,* **19**, 77-87.

Kleinman R.W., Griffin H.C. and Kerner N.K., (1987), Images in chemistry, *Journal of Chemical Education,* **64**, 766-770.

Kozma R.B. and Russell J., (1997), Multimedia and understanding: expert and novice responses to different representations of chemical phenomena, *Journal of Research in Science Teaching,* **34**, 949-968.

Kozma R.B., Russell J., Jones T., Marx N. and Davis J., (1996), The use of multiple, linked representations to facilitate science understanding, In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl. (Eds.), *International Perspectives on the Design of Technology-Supported Learning Environments* (pp 41-60), Lawrence Erlbaum Associates, New Jersey.

Lijnse P.L., Licht P., Waarlo A.J. and de Vos W. (Eds.), (1990), Relating macroscopic phenomena to microscopic particles, *Proceedings of Conference at Utrecht Centre for Science and Mathematics Education,* University of Utrecht, and references therein.

Mayer R.E., (1997), Multimedia learning: are we asking the right questions? *Educational Psychologist*, **32**, 1-19.

Milheim W.D., (1993), How to use animation in computer assisted learning, *British Journal of Educational Technology*, **24**, 171-178.

Moreno R. and Mayer R.E., (2000), A learner-centered approach to multimedia explanations: deriving instructional design principles from cognitive theory, *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*. **2**.

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

Nakhleh M.B., (1993), Are our students conceptual thinkers or algorithmic problem solvers? *Journal of Chemical Education,* **70**, 52-55.

Nakhleh M.B. and Mitchell R.C., (1993), Conceptual learning vs problem solving, *Journal of Chemical Education,* **70**, 190-192.

Niaz M., (1995), Relationship between student performance on conceptual and computational problems of chemical equilibrium, *International Journal of Science Education*, **17**, 343-355.

Nurrenbern S.C. and Pickering M., (1987), Concept learning versus problem solving: is there a difference? *Journal of Chemical Education*, **64**, 508-510.

Paivio A., (1990), *Mental representations: a dual coding approach,* Oxford University Press, New York.

Pascual-Leone J., (1969), *Figural intersection test,* York University, Ontario.

Prosser M., Trigwell K., Hazel E. and Waterhouse F., (2000), Students' experiences of studying physics concepts: the effects of disintegrated perceptions and approaches, *European Journal of Psychology of Education*, **15**, 61-74.

Russell J.W., Kozma R.B., Jones T., Wykoff J., Marx N. and Davis J., (1997), Use of simultaneous-synchronised macroscopic, microscopic, and symbolic representations to enhance the teaching and learning of chemical concepts, *Journal of Chemical Education*, **74**, 330-334.

Sweller J., (1988), Cognitive load during problem solving: effects on learning, *Cognitive Science*, **12**, 257-285.

Sweller J., (1994), Cognitive load theory, learning difficulty and instructional design, *Learning and Instruction*, **4**, 295-312.

Taber K.S., (1998), The sharing-out of nuclear attraction: or “I can't think about physics in chemistry”, *International Journal of Science Education*, **20**, 1001-1014.

Tasker R., (1992), Presenting a chemistry youth lecture, *Chemistry in Australia*, **59**, 108-110.

Tasker R., Bucat R., Sleet R. and Chia W., (1996), The VisChem project: visualising chemistry with multimedia, *Chemistry in Australia*, **63**, 395-397; and *Chemistry in New Zealand*, **60**, 42-45.

Tasker R., (1999), CD for *Chemistry: molecules, matter and change*, Jones L. and Atkins P. 4th Ed. WH Freeman & Co., New York.

Tasker R., (2001), CD and Web Site for *ChemCom — Chemistry in the Community*. An American Chemical Society Project. Heikkinen H (ed.) 4th Ed. WH Freeman & Co., New York. See http://www.whfreeman.com/chemcom and go to ‘Interactive ChemCom Media for Instructors and Students’, Last accessed on December 18, 2005.

Tasker R., Dalton R., Sleet R., Bucat B., Chia W. and Corrigan D., (2002), VisChem: visualising chemical structures and reactions at the molecular level to develop a deep understanding of chemistry concepts, Last accessed on December 18, 2005, from the Learning Designs Web site: http://www.learningdesigns.uow.edu.au/exemplars/info/LD9/index.html

Tasker R. and Jones L., (2002), CD, workbook, and Web Site. *Bridging to the lab: media connecting chemistry concepts with practice*, WH Freeman & Co., New York.
See http://bcs.whfreeman.com/bridgingtothelab/ and view the free sample, last accessed on December 18, 2005.

Tasker R., Bell J., and Cooper M., (2003), Web Site for *General Chemistry - an American Chemical Society Project*. Bell J. (ed). WH Freeman & Co., New York.
See http://www.whfreeman.com/acsgenchemhome/, click on ‘Web Companion’, last accessed on December 18, 2005.

Tasker R., (2004), Web Site for *Chemical principles, the quest for insight*, Atkins, P., & Jones, L. 3rd Ed. WH Freeman & Co., New York. See http://www.whfreeman.com/chemicalprinciples/ and go to ‘Animations’, last accessed on December 18, 2005.

Witkin H.A., Oltman P.K., Raskin E. and Karp S.A., (1971), *Manual for embedded figures test, children’s embedded figures test and group embedded figures test*, Consulting Psychologists Press, Palo Alto.

CHEMISTRY EDUCATION RESEARCH AND PRACTICE
(*CERP*)

Themed Issue on

EXPERIMENTS AND THE LABORATORY IN CHEMISTRY EDUCATION

Scheduled for publication in April 2007

Guest Editors: Avi Hofstein and **Rachel Mamlok-Naaman,**
The Weizmann Institute of Science, Department of Science Teaching, Israel

CALL FOR PAPERS

Contributions are invited for a themed, peer-reviewed issue of CERP on

Experiments and the laboratory in chemistry education

The contributions will be of two kinds:

(a) research-based papers;
(b) effective practice-oriented papers.*

Possible subjects for contributions include:

- Learning chemistry in the school laboratory (conditions for effective learning, learning in and from the chemistry laboratory, social interaction).
- Students' attitude towards and interest in chemistry laboratory work.
- Students' perceptions of the chemistry laboratory learning environment.
- Assessing students' performance, progress and achievement using different modes of presentations in the chemistry laboratory (written evidence, practical examination(s), continuous assessment, development and implementation of assessment tools).
- Matching goals for learning with laboratory practice.
- Incorporating inquiry technologies in the laboratory.
- Simulation and the laboratory.
- The role and effectiveness of demonstrations of chemical experiments.
- Teacher education and professional development in connection with laboratory instruction.

Papers could refer to one or more of the elementary, secondary (high) school or university levels. The list is intended to suggest the scope of possible contributions, but it is not exclusive.

* Please note that papers that describe innovative laboratory experiments or demonstrations without providing some evidence about their actual effectiveness on learning and/or student motivation and interest (this is what is meant by 'effective practice') will not be given consideration.

- **Submissions of manuscripts** (in the format required by the journal – see guidelines on the journal homepage at http://www.rsc.org/Education/CERP/guidelines.asp) or enquiries concerning the suitability of possible contributions should be sent directly by e-mail to: **Avi Hofstein,** The Weizmann Institute of Science, Department of Science Teaching, Israel, avi.hofstein@weizmann.ac.il . Please copy your correspondence also to cerp@rsc.org .

- **IMPORTANT DATES:** Submission of manuscripts by: **October 31, 2006**. Potential contributions will be subject to the journal's usual peer-review process. Where revisions are required, these must be submitted by **February 15, 2007**.