January 2007
ISSN 1109-4028

Volume 8, Issue No 1
Pages 1 – 104

Chemistry Education Research and Practice

Published quarterly by The Royal Society of Chemistry

RSC | Advancing the Chemical Sciences

Chemistry Education Research and Practice

January 2007 | Volume 8, Issue no 1
ISSN 1109-4028 | Pages 1-104

Contents

Papers

Letter

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.

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

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.

Factors that influence pupil engagement with science simulations: the role of distraction, vividness, logic, instruction and prior knowledge

Susan Rodrigues

School of Education, University of Dundee, DD5 1NY, UK
e-mail: S.rodrigues@dundee.ac.uk

Received 3 October 2006, accepted 30 November 2006

Abstract: Constructivist perspectives advocate high quality visual and auditory multimedia to simulate complex and authentic situations. However, the influence of symbolic or representational learning materials on pupil engagement or learning outcomes is not clear. This paper reports on pupil engagement with two types of simulation commonly found in school science (to illustrate practical experiments or depict microscopic chemical interactions). The project pilot phase involved three 15-16 year old male pupils and a main phase involved twenty one 14-15 year old pupils. They were presented with a digital record of their 'think aloud' behaviour with the simulation and they were asked for retrospective comment. Pre and post surveys were also used. Distraction, vividness, logic, instruction and prior knowledge played a significant role in determining the nature of engagement and the outcome of engagement. E-assessment involving multimedia or symbolic representation in science education must take great care if it is to ensure that what it is assessing is the pupil's science capability and not information processing skills that rely on shared symbol identification or on the ability to follow the designers' logic of instructions. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 1-12]

Key words: Chemistry, Simulations. Secondary School Science simulations, practical experiments, microscopic representations, secondary school chemistry

Introduction

A review by Rodrigues (2004) documented the potential of various computer-based technologies for science education. Much of this potential is guided by a belief in their power to motivate pupils and is guided by a suggestion that goal orientation, interactivity and feedback empowers pupils and produces enhanced learning outcomes (Leach, and Moon, 2000; Rodrigues, Smith and Ainley, 2000; Wu, Krajcik and Soloway, 2001; Löhner, Van Jooligen and Savelsbergh, 2003; Mayer, 2003; Ardac and Akaygun, 2004; Lagowski, 2005). However, Cuban (2001) suggests that the potential of goal orientation, interactivity and feedback found in various multimedia has not been fully realised in science classrooms. This may be because, beyond macro design principles, little is known about the effects of micro design variables (Tabbers et al., 2004). Indeed, the influence of symbolic or representational learning materials is not clear (Betrancourt and Tversky, 2000). Given the trend for increasing reliance on computer-based teaching resources (Cuban, 2001; Lagowski, 2005), what is needed is a clearer understanding of what engages and sustains pupil interest when using multimedia in science.

Some authors suggest that what is known about the various design variables is sometimes ignored by software developers (Gee, 2003). But in reality, only a few studies in the area of formal education have investigated either pupils' perceptions of learning in computer technology-based classrooms (Deaney et al., 2003) or pupils' roles in determining the success or failure of computer based technology use (Levin and Wadmany, 2005). Though it should be noted that researchers have studied young people's engagement with computer games (see Kennedy, 2002).

Mayer (2001) suggests that pupils build mental representations of multimedia instructions and Sweller, van Merriënboer and Pass, (1998) have argued that changing presentation format can reduce unnecessary memory load. For example, Moreno and Mayer (1999) suggest that multimedia instructions are more effective, when verbal and visual information is presented close together (the contiguity principle). Mayer (2001) suggests instructions are effective when verbal information is presented aurally rather than visually (the modality principle). Experiments show that replacing written text with speech increases test scores (Kalyuga et al., 1999; Moreno and Mayer, 1999). Often Baddeley's (1992) two modality systems, (visual/spatial information and aural information) model of working memory is used to explain findings. Many studies demonstrating modality effects claim a reduction in extraneous load as a consequence of more efficient use of available memory resource. However, results could also be explained by a reduction in visual search (Tabbers et al, 2004).

The constructivist perspective advocates high quality visual and auditory interfaces to simulate complex and authentic situations. In essence the argument centres on the potential of multimedia to help learners construct their understandings by establishing links between what they know and familiar or not so familiar situations that appear realistic because of the high quality depiction available. However, analysis of static or interactive visualisations to help pupils integrate symbolic representations suggests that not all constructivist-influenced activity improves learning outcomes (Bodemer et al., 2004). Nevertheless, much of the perceived value of multimedia lies in the notion of interactivity (Prensky, 2005). The cybernetic literature on feedback provides many definitions for interactivity, but given the nature of multimedia commonly found in schools, the project reported in this paper uses Laurillard's (1993) definition of interactivity: users receiving "*intrinsic feedback on their actions that relate to the nature of the task goal*" (page 94). A project aim was to help understand how interactivity in this broad sense can arouse and sustain pupil engagement. By closely monitoring pupil engagement and by seeking pupils' perspectives on their engagement, the project hoped to provide insight into how multimedia design features function to either enhance or restrict learning and interest.

Simulations

Two types of simulations were used in the project. They were high quality simulations aimed at introducing college chemistry (United States of America, general chemistry) that could be downloaded free from the website, providing the source was acknowledged. http://www.chem.iastate.edu/group/Greenbowe/sections/projectfolder/animationsindex.htm

This website, rather than a software package, was chosen for two reasons. First, a website would provide readers with access to the simulations evaluated, and second, the website simulations matched the types used, and the concepts and learning outcome levels advocated for secondary school science in Scotland. For example, pupils usually aged between 14-16 years, can pursue Standard Grade Chemistry or Intermediate 1 and Intermediate 2 Chemistry. Course documents at this level include statements such as "*describe the formation of ions in terms of atoms loosing or gaining electrons*" or include activities such as audio visual material to illustrate lattice formation or "*prepare standard solutions of common acids and alkalis, titrate to check accuracy*". Two types of simulation, typical of many commercial products in terms of their appearance, their quality and their purpose, were used. Some of the simulations portray experiments; they provide opportunities to conduct experiments without incurring costs for equipment or resources, and they allow for multiple repeat experiments. From a teaching perspective the microscopic representations provide an opportunity to animate what has until recently only been shown as static models in textbooks, or as ball and stick models in classroom demonstrations. On the website, some of the simulations were accompanied with tutorial worksheets. However, for the purpose of this project, the simulations were used without accompanying worksheets as the intention was to determine what guides and influences pupil engagement when they have access to simulations and to see what pupils attend to, focus on, or continue to pursue. These aspects are important for two key reasons. First, the call for autonomous learning, suggests that pupils should take ownership for their learning. It is therefore important that we explore what it is that will engage pupils and encourage them to take ownership. Second, in Scotland the assessment portfolio is moving toward the inclusion of electronic forms of assessment. It is quite possible that simulations may be included in formal assessment practices for science education, and it is therefore important to design simulations that assess science appropriately.

The project used a simulation that depicts the microscopic level. These types of simulations usually involve the movement of coloured circles or spheres to represent the various ions, atoms, molecules or particles. Examples selected for the project included an illustration of a neutralisation reaction between sodium hydroxide and hydrochloric acid, an illustration of sodium chloride dissolving, and an illustration of the reaction between ozone and nitrogen oxide. In all three cases, various coloured circles moved around the screen. The neutralisation simulation contained text on screen prior to and during the simulation. The dissolving simulation provided text at the end of the action sequences and allowed the user to replay, zoom in, or select a particular viewing position. The ozone simulation involved the circles making contact, an equation remained on screen throughout, and there were tab navigation instructions on the top edge of the screenshot. It could be argued from a modelling perspective, that the spacing of the coloured circles in the simulations that were used probably reflect more accurately the spacing associated with particles in a gas and that this could influence the development of pupils' alternative conceptions. However, as the project aimed to investigate pupils' engagement with simulations commonly found in schools, none of the simulations were modified to address this spatiality issue.

The second type of simulation used was a representation of an experiment that included static equipment but required the user to complete text boxes, change quantities or move dials or scales. Examples of this type included experiments on titration, on the combustion of an organic gas and on the electrical conductivity of solutions. In all three, the user had to follow a particular sequence to make the simulation operational. In all three, error messages appeared if the user was incorrect. In one of the three, the user was provided with feedback showing how to correct the error made. In all three, there were tabs (with menus) to pullout, there were button type selection options, there were text boxes for users to key in information, there were sliders to change quantities, and there were submit buttons.

Methodology

There were two phases to the project. In both the pupils and their parents/guardians signed informed consent forms prior to participation. The pilot phase involved three male volunteers (private secondary school pupils aged 15-16 years). They used the two simulations individually and were interviewed individually. The interview questions, methodology and survey were refined during the pilot phase. However, the pilot study data provides some useful insight into the nature of pupil behaviour, and has therefore been included in this paper.

The main phase involved a well-regarded public secondary school and twenty-one volunteer, male and female, pupils (aged 14-15 years). Working in pairs, the pupils used two simulations for up to five minutes. In the main phase, pupils were asked to work in pairs (but due to the number of volunteers, there was also one group of three). The paired approach was an attempt to access pupils' thinking. While working together some pupils 'thought aloud' as they sought clarification from their peer or as they reacted to a screen shot. Working in pairs also reduced their perception of the exercise as a test. Their engagement was digitally recorded, to show the computer screen and to record their dialogue.

Directly after using the simulation, the pupils, in their pairs, were interviewed. The digital record was replayed. The pupils were asked to explain particular actions and behaviours (how they knew how to complete particular tasks, deployed particular protocols, and knew how or when to use particular actions). The retrospective interview account was initiated when the researcher replayed the digital record and asked the pupils to comment on their actions and engagement. When pupils' responses were particularly animated or they drew attention to aspects of their engagement, the researcher probed for clarification. In addition, during the pilot phase, the three pupils made comments about certain aspects; these were used as prompts in the main phase. For example, pupils in the pilot phase all mentioned design elements and were then asked for advice to a software designer. Consequently in the main phase, toward the end of the retrospective account-interview, all pupils were asked for advice to a software designer. In addition pupils' science concepts were explored briefly and informally.

This methodology is based on the work reported by Clarke (1998) in which retrospective accounts were used to access pupils' thinking and it uses a 'thinking aloud' strategy usually credited to Ericsson and Simon (1984), which is an approach that has often been used in cognitive psychology research, usually to explore pupils' problem solving strategies. In recent times research investigating human-computer interactions or involving usability testing have deployed thinking aloud protocols to evaluate new software (Crowther et al., 2004). The Ericsson and Simon (1984) method requires the researcher to provide initial general instructions and then withdraw, leaving the pupils to verbalise their thoughts whilst undertaking the task. The semi-structured retrospective interview technique included presenting pupils with a digital record of their pathway data documenting pupil behaviour and actions and asking pupils for retrospective comment. The digital record provided pupils with an artefact to support recall after the event and provided the researchers with opportunities to explore pupils' science understandings, and their engagement protocols, in a little more detail.A key project aim was to collect data on pupils' interactive decision making, taking into account the evaluation areas outlined by Milton and Lyons (2003): interface usability (how easy did the pupils find it to use?), content validity (what sense does it make to the pupils?) and educational utility (what do the pupils learn?)

Various perspectives can be used to analyse transcript episodes. In this project the pilot phase helped the researcher identify a set of circumstances. For example pre-determined questions, based on feedback collected during the pilot phase, were included during the retrospective interview accounts in the main phase. The researcher also identified what were considered to be intriguing pupil phenomenon (for example pupil repeated actions, returning to particular segments, or pressing particular buttons) and these were used as markers or cues when viewing the digital records of others). This fits in with conversation/discourse analysis methodology, as conversation/discourse analysis encourages research to be "data-driven" rather than pre-specified, and allows the study of pupil focus as it manifests in their interactions, resulting in a best fit heuristic (Hutchby and Wooffitt, 1998). Using common sense the researcher then interpreted episodes, elucidated the interpretation and provided some reasoning.

During explication, the researcher used personal knowledge (years of experience as a teacher, awareness of general pupil classroom behaviour, knowledge of curriculum content, research experience involving pupils use of various ICT, conversation analysis techniques, etc.) as a resource to study the pupil-computer-pupil interaction episodes. Clearly, the transcriber-researcher's interpretation of what pupil utterances imply is informed by the transcriber-researcher's perspective. However, validity can be promoted by ensuring that transcripts and interpretations are post-checked by others including, in this project, the pupils. Pupils were given their digital records and transcripts. Validity is also possible, because the pupil may, in later transcript episodes, refer to a previous interaction episode, and provide additional evidence to support or challenge the researcher's analysis. Also, episodes and their analysis can be compared, allowing an observation made in one interaction episode to be used as the trigger or starting point for further analysis of other episodes and engagements with the same or other pupils (Heath, 2004). The transcript excerpts provided are brief, but representative of the mechanisms and procedures in evidence in the wider data set. The selection of particular excerpts was primarily based on conversation thread length. The examples provided in this paper, are simply that, illustrative examples of the interaction episodes.

The digital recording of their behaviour was given to the pupils and the digital record was transcribed. The transcript excerpts provided in this paper are coded. DHS refers to the pilot school and the volunteers from the pilot are numbered. SMA refers to the main phase school, and the volunteers are represented with initials. Additional data was collected by asking the pupils to complete brief (one side) pre and post surveys. The pre survey had four questions on one side of A4, and collected information about pupils' perceptions of their experience and comfort in using a variety of ICT. One question asked about general computing ability, using a 5 point Likert type scale. Another question asked for their experience rating, the options were 'a lot of experience', 'some experience', 'a little experience', 'no experience'. A third question asked for their comfort level rating, the options were ''very comfortable', 'moderately comfortable', 'need some help', 'need lots of help'. The final question asked about computer use to perform particular tasks at home or school (send an email, played a computer game, posted messages on bulletin boards). These types of survey items are common in surveys on ICT use. The post survey had six questions on one side of A4. It included questions that sought pupils' views, and therefore included three open ended survey items asking pupils what background information and skills they needed in order to use the simulations they used, and what advice they would give a friend who was going to use the simulation. The post survey also asked the pupils to rate the technical demand, the science demand and the potential for confusion when using the simulation. A final statement allowed for any other comments. All surveys were piloted before being used in the main phase.

Findings

In the pre-survey, the twenty-one pupils in the main phase and the three in the pilot phase were asked to rate their experience level and comfort level for several applications. Table 1 documents their experience rating value. The pilot phase pupils' responses appear in brackets. Therefore, sixteen of the twenty-four pupils felt they had lots of experience of word processors, seventeen said they had lots of experience of computer games and seventeen said they had lots of experience of the internet, but only four said they had lots of experience with simulations. The Table 2 documents the pupils' perceived comfort level rating.

Table 1. Pupils' perceived experience for some ICT.

Activities using technology	a lot of experience	Some experience	A little experience	No experience
Word processing	13 (3)	6	2	0
Databases	3	7 (1)	11 (2)	0
Presentation software PowerPoint	3	10 (1)	6 (2)	2
Computer games	14 (3)	5	2	0
Internet search engines	15 (2)	6 (1)	0	0
Simulations	2 (2)	6	7 (1)	6
Desktop publishing	2	4	9 (2)	6(1)

Table 2. Pupils' comfort ratings for some ICT.

Activities using technology	Very comfortable	Moderately comfortable	Need some help	Need lots of help
Word processing	10 (3)	10	0	1
Databases	3	5 (1)	12 (1)	2 (1)
Presentation software/PowerPoint	5 (2)	5	9 (1)	2
Computer games	15 (3)	5	1	0
Internet search engines	16 (2)	4 (1)	1	0
Simulations	2 (2)	5	10 (1)	4
Desktop publishing	1	5 (2)	9	6 (1)

Twenty-three of the twenty-four felt that they were very or moderately comfortable with word processing and playing computer games, while only nine said they were very or moderately comfortable using simulations. These experience and comfort ratings are of interest for two reasons. First, it would appear that many pupils do not equate computer games with simulations, and second, their Internet use does not predominantly involve simulations.

Analysis of the pupils' digital recordings suggests that there were three key aspects that influenced the process of engagement. These were:

- distraction (led astray by interesting, but at the time redundant, segments) and vividness (items that stand out),
- logic (factors influencing ability to organise information) and instructions
- prior knowledge.

These factors played a significant role in determining the nature of engagement and the outcome of engagement as shown by brief excerpts from the transcripts.

Distraction and vividness

Some of the pupils highlighted particular elements within the simulation as aspects that determined their course of action. For example, the following is taken from the retrospective interview when two male pupils were watching their digital recording. The replay showedthat when the titration simulation appeared, the two boys started with instruction 5, "push the slider….."

Transcript 1

Pupil S	*It said to push the slider up to add volume. But when we tried to push the slider up it didn't seem to want to work.*
Researcher	*Why did you start with that instruction? Push slider to….*
Pupil S	*I don't know actually, because that is number 5. Eh?!* (chuckles)
Researcher	*But you were automatically drawn to that one?*

Pupil M *Yeah*
Pupil S *Yeah*
Researcher *And then you were clicking away at something.*
Pupil M *Yeah, the drop wise.*
Researcher *Oh the dropper. So why were you clicking on the dropper?*
Pupil M *It was the biggest thing that you can see.* (SMA MS)

Transcript 2

Researcher *Oh, ok. So what made you think you had to move that slider up?*
Pupil C *Because it was like, red and*
Pupil L *Like, it just looked like you are supposed to do it.* (SMA CL)

These two transcripts are representative of the feedback received from all the pupils with respect to the influence of these particular elements on the simulation. Notably, all those who tried the titration simulation were intuitively drawn toward using the slider and a red 'dropper' button. However, neither action resulted in any productivity as these actions were required during the experiment, and not at the start.

Logic and instructions

Analysis of the digital recordings of the pupils using the titration simulation also identified a similarity in their information processing skills. All the pupils had problems with the titration simulation, regardless of their familiarity with the science, or of their ability, age, and gender. Furthermore, unless the pupils followed the sequence of instructions, the simulation would not work. None of the pupils, who used the simulations representing 'practical experiments', followed the numbered instructions. In some cases the number went undetected and in other cases the number sequence was noticed, but their relevance was not identified. Some pupils accounted for their failure to follow the sequence by drawing attention to the placement of the instructions.

Transcript 3

Researcher *...And see that? It has got a sequence of numbers there, one, two...*
Pupil J *Yeah, I think the sequence could have been in order.*
Researcher *What do you mean by that?*
Pupil Lu *Well see its got like one, two (points to screen) and three is down there (points), it is like at the side, which is a bit...*
Pupil J *Then there is four (points to lower screen) and there's five* (points to upper screen)
Pupil Lu *And five is at the top and six is right in the middle which is a bit...*
Researcher *So the positions aren't in sequence?*
Pupil J *Yes.*
Pupil Lu *It should have like one, two, three. (points across screen)* (SMA LuJ)

Transcript 4

Pupil C *Step one button, but there was no step three so.*
Researcher *Ok so you were following the steps.*
Pupil L *Yes. I was trying to move it up, but I didn't see step three…*
Researcher *Oh yes, that instruction. Right. So you were then following the written instructions to try and get it to work and you were going across the screen.*
Pupil L *Yes.*
Pupil C *Instead of going down. It makes more sense to go across.* (SMA CL)

Transcript 5

Pupil 1 *I didn't really notice the number sequence, no.*
Researcher *So the order that you were doing it in was…*
Pupil 1 *Was largely by the spacing, by where they were positioned.* (DHS 1)

The sequence and position of instructions was noted by all who used the 'practical experiment' simulation. All the pupils signalled what they perceived to be the logical position of the instructions. The fact that the instructions were numbered did not register with the pupils, instead all focussed on the location of the instructions.

The following excerpt is indicative of a viewpoint shared by all the pupils. When they were asked what advice they would give a designer involved in constructing these simulations, two pupils provided the following:

Transcript 6

Pupil M	*To make the instructions more clear.*
Researcher	*And how could they do that? Because the instructions are all numbered.*
Pupil M	*Aye, but you put the instructions on the side no one reads.*
Pupil S	*See if you had them all in like in a row, because you have number five over here, and on the other side of some test tubes, so you might not even think about looking at the numbers and that, and your eyes are drawn to something. But if you have them all in like order going down or something.* (SMA MS)

It should also be noted that some of the pupils commented on the nature of the instructions.

Transcript 7

Pupil 1	*I would maybe make the tab for choosing the acids and alkalis a little more obvious and I would make it clear that you couldn't type in the boxes that you couldn't type in.* (DHS 1)

Though familiar with various forms of technology, they experienced some difficulty with the instructions.

Prior knowledge

All the pupils explained their progress by referring to their prior knowledge and what they considered to be the relevant field of science. Some even went as far as differentiating their experience in terms of being familiar with biology, chemistry or physics. Consequently some explained their difficulty in working the simulations, or understanding the simulations, in terms of labelling themselves as physicists or biologists and associating the simulations they were asked to do as chemistry. Those who did identify themselves as pupils with chemistry experience relied heavily on this prior knowledge to determine their process of engagement. Numbered sequences of instructions to guide progress were ignored as pupils relied on their prior real time experience for particular experiments, to determine the nature of engagement. In the following excerpt a pupil who has practical experience of titration experiments, draws on this to make sense of the titration simulation:

Transcript 8

Researcher	*So you went for weak acid and strong base.*
Pupil 3	*Yeah. And I think I went for base or was it acid to...*
Researcher	*Stick it in the burette? Any reason for that?*
Pupil 3	*In chemistry in class we always put the acid in the burette.*
Researcher	*In the burette. Ok, then what happens?*
Pupil 3	*I went for the indicator. I remember using bromothymol blue in chemistry so that is what I chose.* (DHS3)

Relying on their prior experience to help with the simulations depicting practical experiments was not as problematic as relying on prior experience to make sense of simulations showcasing microscopic interactions. Many made inaccurate associations that were not necessarily misconceptions, but simply mistaken microscopic identity when the pupils viewed the simulations of the microscopic interactions. For example, the next transcript is about a simulation representing acid-base neutralisation.

The centre of the simulation entry screen contains the following text "*The following is an acid-base neutralisation when aqueous NaOH is added to aqueous HCl*". At the top of the screen are tabs with 'replay' 'continue' 'pause' and 'one back'. In the simulation, water is represented by combined red and two white circles, the hydrogen chloride has dissociated leaving a chloride ion represented by a green circle with a negative charge beside it, and a hydroxonium ion, represented by a red circle attached to three white circles with a positive charge sign next to it. A few seconds into the simulation, a purple circle close to a red circle and white circle (meant to represent sodium hydroxide) enter the simulation at the top of the screen, just below the row of navigation tabs. During the retrospective interview, while watching their digital recording, the following conversation took place.

Transcript 9

Researcher *What sense did you make of it then?*
Pupil S *Basically that those lost their...*
Researcher *Which are those?*
Pupil J *Positons.*
Pupil S *Those (points to the* H_3O^+ *ions)*
Researcher *The red and white ones with the plus?*
Pupil J *Yes*
Pupil S *Those lost their protons, their plus when the purple ones came in.*
Pupil J *Yeah there's the minus ones there and they like smash up. (points to the chloride ion)*
Researcher *Ok. Oh ok. So what do you think was happening?*
Pupil S *They were joining together. Or something.*
Pupil J *Becoming neutral.*
Pupil S *Neutrons.*
Researcher *What was becoming neutral?*
Pupil J *The protons.*
Pupil S *and the minus ones.*
Pupil J *Electrons.*
Researcher *Ok. So what are the green things then?*
Pupil S *Neutrons. I mean electrons.* (SMA SJ)

These pupils did not hold misconceptions; they had simply confused the identity of the circles, and were guided by the signal of the charge written along side the circles. For the two pupils in the previous transcript, the clues provided in the simulation were missed. Thus, although the designers had inserted Cl^- beside the green circle, and H_2O beside the red circle attached to the two white circles, the pupils had not made those associations. These two pupils were not alone. Several of the pairs of pupils using this simulation initially made the same links. They related the reaction to subatomic particles. Most of the pupils modified their explanations when the icons and symbols were clarified. In some cases the pupils had debated whether the simulation was a representation of a subatomic particle reaction or a representation of ionic interaction.

Some of the pupils also explained their progress and their ability to engage with the simulations by referring to their prior knowledge of similar types of computer-based activity. For example two girls in the main phase of the project, had experience of the 'BBC Bitesize' revision online programme, they said:

Transcript 10

Pupil Ca *We knew instinctively that the button would change.*
Researcher *Did you, how did you know that?*
Pupil T *Because I use the 'Bitesize revision' so it is quite handy and they have got a lot of diagrams like this.* (SMA CaT)

Some of the pupils were familiar with technology button functions and that either helped or hindered their progress.

Transcript 11

Researcher *Ok. You went to acid. So what was your thinking when you were doing that?*

Pupil 2 *I wasn't very sure what I was meant to be doing. And at first I didn't actually know that there were tabs on the side. I usually associate those radio buttons with just selecting options.* (DHS 2)

Those with prior knowledge of various types of computer based presentations also experienced productive and unproductive moments. Some were able to draw on their prior experience to help make progress, while others found some of the technology to be at odds with what they had previously experienced.

In general, prior science knowledge was a key factor in deciding how to progress through a simulation. In all cases, prior science knowledge was recalled in order to engage with the simulation, and prior science knowledge was used to determine the course of action. In the simulations representing practical work, the pupils drew on their prior practical science knowledge to determine their course of action. In most cases this overrode the directions available in the form of numbered instructions. In the practical experiment simulations this was problematic, as the format of the simulation did not always allow for divergence to take into account the pupils' prior science knowledge. In the microscopic representation simulations this was problematic, as the prior science knowledge that pupils tapped into did not match the science knowledge presented by the designer. In addition, in some cases the role of prior science knowledge was problematic, because pupils discounted their capacity to undertake the simulation because they perceived themselves as not having experience in the relevant field of science.

Conclusion

The simulations used were of good quality and reflected the standard available to most schools. The simulations also addressed the levels identified in the various curricula documents. Analysis of the digital records for the pupils suggests that three key aspects influenced their ability to engage fruitfully with the simulations. These three aspects were:

- Distraction and vividness
- Logic and information
- Prior knowledge.

All the pupils commented on the logical placement and the nature of instructions, and this was seen to be instrumental in guiding progress. So, though the instructions in the titration simulation were assigned numbers to help identify the sequence for engagement, their location on the screen was not in a sequence identified by the pupils as logical. Earlier it was stated that Baddeley's (1992) two modality systems are often used to explain findings and that some studies demonstrating modality effects claim a reduction in cognitive load, as a consequence of more efficient use of available memory resource. However, these findings support the views of Tabbers et al. (2004), who suggested that a reduction in visual search may be the key factor. In this project, peripheral content appeared to have an effect on comprehension. As was seen, the third instruction in the titration simulation, which contained vital information and without which the burette could not be 'filled', was missed by all the pupils, because it was on the periphery and inserted as a vertical tab. The pupils appeared to attend to items that required minimum visual search. The pupils appeared to be drawn to items on screen that were vivid, for example the 'red dropper', or the slider instruction in the centre of the screen.

All the pupils referred, at some point, to their prior science knowledge, which they used to determine the course of action. In some cases this was problematic because pupils focussed on particular aspects of symbolic representations and mistakenly construed explanations based on these assumptions. For experts, or those very familiar with the use of symbolic representations of the microscopic level, these mistakes may appear to demonstrate a lack of understanding. However, I would suggest that this is not the case. When the symbols found in these symbolic representations were clarified, the majority of pupils were able to explain the designer intended science. Therefore their retrospective narrative, in which they identified particular microscopic particles erroneously, does not reflect their understanding of the science, but it does reflect their information processing capacity. Therefore, it is important that we ensure that symbolic representations used in multimedia software do not assume that the pupils will assign meanings to symbols as intended by the designers. As the pupils suggested, keys may help them clarify the identity of the icons used in the animations, and may then provide a more realistic opportunity to inspect their science understanding. Likewise, though most of the pupils were unable to carry out the titration because they relied on their prior knowledge and discounted the instructions available, those who had undertaken titrations in the past, were able to explain clearly the process, their intentions and their science understanding for the reaction.

In light of the fact that recent developments in assessment practices in education in Scotland have come to include the use of electronic formats, it is quite possible that e-assessment protocols will also be used in science education in the near future. The findings from this project suggest that e-assessment involving the use of multimedia or symbolic representation in science education will have to take great care if it is to ensure that what it is assessing is the pupils' science capability and not information processing skills that rely on shared symbol identification or on their ability to discern the designers' logic of instructions.

References

Ardac D. and Akaygun S., (2004), Effectiveness of multimedia-based instruction that emphasizes molecular representations on students' understanding of chemical change, *Journal of Research in Science Teaching*, **41**, 317-337.

Baddeley A. D., (1992), Working memory, *Science*, **255**, 556-559.

Betrancourt M. and Tversky B., (2000), Effect of computer-based simulation on user's performance: a review, *Le Travail Humain*, **63**, 311-329.

Bodemer D., Ploetsner R., Feuerlein I. and Spada H., (2004), The active integration of information during learning with dynamic and interactive visualisations, *Learning and Instruction*, **14**, 325-341.

Clarke D.J., (1998), Studying the classroom negotiation of meaning: complementary accounts methodology, In A. Teppo (Ed.) *Qualitative research methods in mathematics education*, Monograph No. 9 of the *Journal for Research in Mathematics Education*. Reston, VA: NCTM.

Crowther M.S., Keller C.C. and Waddoups G.L., (2004), Improving the quality and effectiveness of computer-mediated instruction through usability evaluations, *British Journal of Educational Technology*, **35**, 289–304.

Cuban L., (2001), *Oversold and underused: computers in schools 1980-2000,* Cambridge MA, Harvard University Press.

Deaney R., Ruthven K. and Hennessey S., (2003), Pupil perspectives on the contribution of ICT to teaching and learning the secondary school, *Research Papers in Education*, **18**, 141-165

Ericsson K.A. and Simon, H.A., (1984), *Protocol analysis: verbal reports as data* (Revised ed.), London, MIT Press.

Gee J.P., (2003), *What video games have to teach us about learning and literacy*, Gordonsville, VA; Palgrave Macmillan.

Heath C., (2004), Analysing face to face interaction: video, the visual and material, In David Silverman (Ed.), *Qualitative research: theory, method and practice* (second edition), London, Sage, pp. 266-282.

Hutchby I. and Wooffitt R., (1998), *Conversation analysis: principles, practices and applications*, Polity Press, Cambridge.

Kalyuga S., Chandler P. and Sweller J., (1999), Managing split attention and redundancy in multimedia instruction, *Applied Cognitive Psychology*, **13**, 351-37.

Kennedy H., (2002), Lara Croft: feminist icon or cyberbimbo? On the limits of textual analysis, *Games Studies*, **2**, 2. www.gamestudies.org/o2o2/kennedy/ Retrieved 7th July 2005.

Lagowski J.J., (2005), A chemical laboratory in a digital world, Proceedings of the 18th International Conference on Chemical Education, *Chemical Education International*, **6**, 1, 1-7. http://www.iupac.org/publications/cei/vol6/02_Lagowski.pdf Retrieved 26th March 2006.

Laurillard D., (1993), *Rethinking university teaching*, London, Routledge.

Leach J. and Moon J., (2000), Pedagogy, information and communications, *The Curriculum Journal,* **11**, 385-404.

Levin T. and Wadmany R.,(2005), Changes in educational beliefs and classroom practices of teachers and pupils in rich technology based classrooms, *Technology, Pedagogy and Education*, **14**, 281-307.

Löhner S., Van Jooligen W.R. and Savelsbergh, E.R., (2003), The effect of external representation on constructing computer models of complex phenomena, *Instructional Science,* **31**, 395-418.

Mayer R.E., (2001), *Multimedia learning,* Cambridge University Press, New York.

Mayer R.E., (2003), The promise of multimedia learning: using the same instructional design methods across different media, *Learning and Instruction* **13**, 125-140.

Milton J. and Lyons J., (2003), Evaluate to improve learning: reflecting on the role of teaching and learning models, *Higher Education Research and Development*, **22**, 297–312.

Moreno R. and Mayer R.E., (1999), Cognitive principles of multimedia learning: the role of modality and contiguity, *Journal of Educational Psychology*, **91,** 358-368.

Prensky, M., (2005), Engage me or enrage m': educating today's 'digital native' learners. Keynote. SETT Conference, SECC, Glasgow, Scotland. 21st September 2005.

Rodrigues S., (2004), Digital divides: e-literacy in science classrooms when using information communication technologies, *Science Education International*, **16**, 303-323.

Rodrigues S., Smith A. and Ainley M., (2001), Video clips and animation in chemistry CD-ROMs: student interest and preference, *Australian Science Teachers Journal,* **47**, 9-16.

Sweller J., van Merriënboer J.J.G. and Pass F., (1998), Cognitive architecture and instructional design, *Educational Psychology Review*, **10**, 251-296.

Tabbers H.K, Martens R.L. and van Merriënboer J.J.G., (2004), Multimedia instructions and cognitive load theory: effects of modality and cueing, *British Journal of Educational Psychology,* **74**, 71-81.

Wu H.K., Krajcik J.S. and Soloway E., (2001), Promoting understanding of chemical representations: students use of a visualisation tool in the classroom, *Journal of Research in Teaching*, **38**, 821-842.

Results of an interview study as basis for the development of stepped supporting tools for stoichiometric problems

Martin Fach*, Tanja de Boer and Ilka Parchmann

University of Oldenburg, Institute of Pure and Applied Chemistry, D-26111 Oldenburg, Germany
e-mail: martin.fach@uni-oldenburg.de

Received 28 July 2006, accepted 12 December 2006

Abstract: In recent years many research studies investigated students' misconceptions in stoichiometry, and problem solving strategies on stoichiometric problems. Additionally, alternative approaches for teaching this issue of chemistry were developed. However, among students and teachers this topic is still regarded as being difficult and unmotivating. Our approach is to combine (qualitative) investigations with the development and evaluation of specific teaching and learning material. To help students working on stoichiometric problems, we developed a set of stepped supporting tools (SST), based on the results of an interview study investigating the phases of the solution processes of German secondary school students (grade 9) on these problems. Resulting from students' difficulties detected in the interviews, which were in good agreement to those described in the literature, four different types of SST were developed, (1) giving general instructions on how to tackle (these) problems, (2) showing the steps of the solution process, (3) advising students how to carry these steps out and finally (4) providing them with a glossary of important terms. The method seems also to be applicable to other topics in chemistry, raising the prospect of a catalogue of SST for other problems, too. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 13-31]

Key words: stoichiometry, stepped supporting tools, problem solving, misconceptions, problem-centered interviews

Introduction

In research on education in the sciences there seem to be two main ways of working: many research groups focus on investigating certain aspects of science education, trying to give a detailed description of e.g. students' ways of thinking or the pros and cons of a certain teaching method. Other research groups try to develop teaching or learning materials to improve on classroom situations. However, the two aspects of research are seldom linked to each other. Although there are models that try to combine the two sides of the same coin [e.g. 'the model of educational reconstruction' (Duit, et al., 2005) or the model of 'developmental research' (Lijnse, 1995)] the 'investigating researchers' often only provide 'recommendations for teaching' without building on their results to produce teaching modules. Our approach tries to bridge this gap between investigation, development and evaluation. In a pilot project we tried to determine students' problem solving strategies on stoichiometric problems and to link the observed difficulties and underlying misconceptions to the steps of their solution in order to devise 'stepped supporting tools' (SST) for this type of problem. Using SST means that, while working on a (stoichiometric) problem, students can rely on a set of prepared cards, which give them supporting information for the solution of the problem without providing them with the full solution. It is widely agreed that providing every student solving a problem with the specific help he/she needs to overcome obstacles in the process of the solution is a very important factor in teaching. The method of using SST may be one way to approach this goal (cf. ForschergruppeKassel, 2006). To do this properly one has to know two things: (1) the way the student solves certain problems and (2) the problems and misconceptions that may obstruct this process. In recent decades many studies have been conducted to find out about students' misconceptions (cf. Griffiths, 1994; Barker, 2000) and about students' problem solving strategies in chemistry (cf. Gabel and Bunce, 1994). Although these data may be meaningless for any individual pupil, they give a broad database for determining which problems are most common and therefore need special attention.

The main purpose for the use of this kind of SST is not in an introductory stoichiometry course, but rather when stoichiometric problems are given to a class for practice. Within the project our goal is to extend our method to other chemical issues to generate a catalogue of SST for many problems in chemistry. This paper shows the method used to investigate students' strategies and to develop the SST. Another paper showing the specific design of the SST and the first evaluation results is in preparation.

Background

Much research has been done on stoichiometric problems in recent years (for reviews see e.g. Gabel and Bunce, 1994; Griffiths, 1994; Furió, et al., 2002). This is probably due to the fact that stoichiometry is a very basic and fundamental concept in chemistry. For example, students have to switch from thinking about concrete aspects of matter to more abstract thinking concerning aspects of particles, thus, they may enhance their conceptual understanding (cf. BouJaoude and Barakat, 2003). On the other hand, many authors agree that the concept is very difficult for students to grasp and therefore discouraging (e.g. Schmidt and Jignéus, 2003). Therefore, to close the gap between what is and what could be, research results will have to be implemented into school practice, providing teachers with specific teaching materials and thus combining fundamental research with day-to-day practice.

Research findings about learning difficulties in stoichiometry and proposals for alternative concepts for teaching

Many studies on stoichiometry dealt with students' misconceptions (e.g. Mitchell and Gunstone, 1984; Schmidt, 1990; Huddle and Pillay, 1996; BouJaoude and Barakat, 2000). The main findings were that students

a. equate the mass ratio of atoms in a molecule with the ratio of the number of these atoms, and the mass ratio with the molar mass ratio (Schmidt, 1990),
b. calculate the molar mass of a given substance by summing up the atomic masses and then multiplying or dividing this sum by the coefficient of the substance in the chemical equation; others do not understand the significance of the coefficients in a chemical equation at all (BouJaoude and Barakat, 2000),
c. confuse the concepts of conservation of atoms and possible non-conservation of molecules or do not take into account the conservation of atoms or mass at all (Mitchell and Gunstone, 1984),
d. cannot determine the 'limiting reagent' in a given problem, when one substance is added in excess (Huddle and Pillay, 1996),
e. confuse or do not know the definitions of and relationships between stoichiometric entities in general (e.g. Furió, et al., 2002).

Other studies took a closer look on students' approaches to solving stoichiometric problems. In a large scale study in the 1990s Schmidt (1994) found out that German students (n = 4181, grade 11-13, age 16-19) mostly used three different strategies to solve stoichiometric questions that can be solved without arithmetical calculations and without a calculator:

a. a strategy using explicit calculation of the amounts of substance (‘mole method’),
b. a strategy avoiding explicit calculation of the amounts of substance and instead using the ratio of molar masses (‘proportional method‘) and
c. a strategy using mere logical reasoning (‘logical method’).

Interestingly, the large majority of the students used the ‘logical method’ (c) to get to the right solution rather than the algorithmic and more abstract strategies (a) and (b). Schmidt concluded that this is due to the fact that the given problems were easy to calculate. This was corroborated by a more recent interview study with a small sample of Swedish students (n = 4, grade 12, age 17-18) of Schmidt and Jignéus (2003). They found that for easy-to-calculate problems the participating students also used the ‘logical method’ but switched to a mathematical strategy e.g. (a) or (b), when confronted with a more complicated task. However, a study conducted in Hungary produced totally different results (Tóth and Kiss, 2005). The Hungarian secondary school students participating in this study (n = 750, grade 7-11, aged 13-17) almost never used the logical method but the mole method (most often), or the proportional method. According to the authors this might be due to the fact that the mole method is the one taught in Hungarian schools most prominently. On the other hand, the German students participating in Schmidt’s study had more chemistry lessons (three to five per week) than the Hungarian students (two per week) and were older in general. So they can be seen as relative experts compared to the Hungarian students and thus developed a logical method more easily.

In another study BouJaoude and Barakat (2003) investigated the relationships of students’ problem solving strategies in stoichiometry to their conceptual understanding and to their learning approaches (e.g. ‘deep approach’ vs. ‘relating ideas‘ vs. ‘intrinsic motivation’, etc., as defined by Entwistle and Ramsden, 1983). Based on results that indicated a connection between sound conceptual and procedural knowledge and successful problem solving, they administered a learning approach questionnaire (LAQ) and a stoichiometry test, partially followed by unstructured interviews, to forty Lebanese students (grade 11, age 16-20). Through the stoichiometry tests and the unstructured interviews they found many of the misconceptions described in earlier studies. They derived three main strategy types from the tests and the interviews.

a. Correct strategies, which were subdivided into ‘algorithmic’, ‘efficient’ and ‘messy’ strategies,
b. Incorrect strategies, subdivided into ‘incorrect strategies-incorrect answer’ and ‘incorrect strategies-correct answer’ and
c. ‘Incomplete’ strategies.

As the authors state, the majority of students participating in this study used algorithmic problem solving “*even when they did not have adequate understanding of the relevant concepts*” (pp. 24-25). In contrast to results in the literature, they did not find a correlation between the factors ‘learning approach’ and ‘conceptual understanding’. Furthermore, they did not find any patterns in the problem solving strategies used by students with different learning approaches.

In another study conducted by Frazer and Servant (1986, 1987) on titration calculations, the authors investigated which one of four possible expert methods was used by students solving two titration calculation problems. Three of the four methods were similar to those reported by Schmidt (1994) if one transfers the methods to this kind of problem:

1. A method deriving the reaction stoichiometry from the balanced chemical equation and using direct calculation of amounts of substance.

2. A method avoiding calculation of amounts of substance and instead using a proportion equation.
3. A method immediately converting the reaction stoichiometry into the quantities given in the text and continuing by using the 'unitary method' (see also Williams, 1980).
4. A method using the 'quantity calculus' and selecting and rearranging equations (see also Packer, 1980).

In the same study of a sample of 244 students only 79 written answers given to the two titration problems were correct. Most of the students used the method (1) to solve the problems, but less than a quarter of these responses were correct. Method (2) was the second most popular, but with poorest success. Method (3) was only used by a few students on the first problem. Method (4) was not used by any of the students. Although the results were poor in general, the authors came to the conclusion that using the second strategy is least recommendable since students only have to "*fill in the blanks*". In contrast, using method (1) may lead to an interlinked understanding of the chemical concepts.

The authors of all these studies give general advice on how to overcome the learning difficulties in the field of stoichiometry. Furthermore, some authors have developed alternative approaches on how to introduce subjects of stoichiometry in school. In Germany, Rossa (1998) for example, suggested visualization of the complex and the abstract entity 'mole'. Kaminski et al. (1994) developed a unit for introducing the chemical formula, setting aside the whole mole-concept and providing students with a table listing how many atoms one milligram of an element contains instead of giving them the molar masses. According to the authors, introduction of the mole-concept should thus be postponed. In other countries similar proposals were made, too. These can be classified into those which focus attention on conceptual prerequisites; those which use new analogies; and those which emphasize applications (for a review see Furió, et al., 2002).

Aims of our project: Combining investigation and development

The above mentioned studies on students' misconceptions in stoichiometry and students' problem solving strategies on stoichiometric problems give a detailed and thorough description of the status-quo. Additionally, alternative approaches show possible ways of coping with some of these challenges. However, as mentioned above, all these studies and developmental approaches were conducted separately. All these researchers give general advice on teaching, but specific materials based on empirical data are rarely provided. This is corroborated by a questionnaire-study conducted in Germany, in which teachers were asked which topics of chemistry were difficult to teach and why (Fiebig and Melle, 2001). The issue of 'basic chemical laws' (including stoichiometry) turned out to be 'difficult to teach', partly because of the lack of suitable teaching methods. This shows that the above mentioned combination of investigation and development should gain more attention in research studies, and that specific teaching and learning material should be developed on the basis of investigation results and should be thoroughly evaluated afterwards.

Furthermore, the proposed teaching material is mostly for introductory courses. Only little teaching material has been developed to help students while having to revise the topic of stoichiometry, dealing with the difficulties of partly understood concepts and partly consolidated knowledge. Thus the aims of our study were

1. to investigate the problem solving strategies German pupils (grade 9, having already been taught the issue of stoichiometry) use when solving stoichiometric problems and to determine the stages of this process;
2. to investigate the problems and misconceptions that occur in the solution process and to identify the relevant ones hindering this process;

3. to develop specific stepped supporting tools from the results in order to use on further problems of this type, especially for revision purposes.

Additionally, since in Germany the issue of stoichiometry is being introduced in grades 8 or 9 (depending on the federal state), one can expect students' problem solving strategies to be different from those reported by Schmidt (1994), because the students have less experience in problem solving in general than older students do, and they also know less about the underlying chemical concepts, such as the particle theory or the chemical equation (see above, and Tóth and Kiss, 2005).

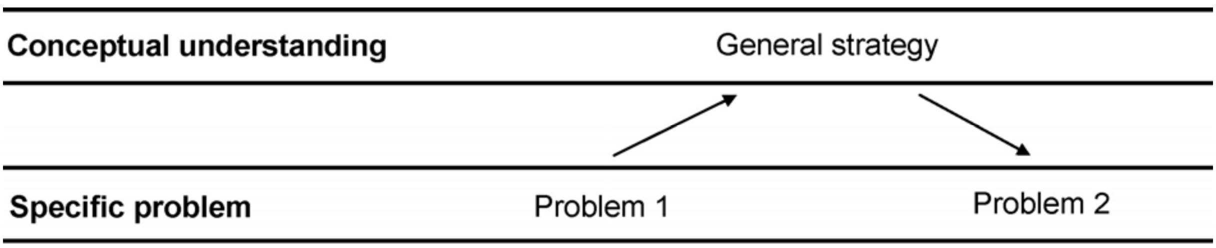
Method

Design of the interviews

One aim of our study is to take a close look at what students are thinking and how they are solving problems. To do this, interviews were conducted. Twenty students of five different classes of four secondary level schools (German Gymnasium) in Lower Saxony were interviewed. The issue of stoichiometry had been introduced in all five classes before. Four teachers participated in the study, one of them teaching two of the five classes. The students were chosen by their teachers. The teachers were told to choose two rather high-performing and two rather low-performing pupils, but since students had to volunteer to do the interviews, this was not always possible. The interviews were conducted by two of the authors (mostly T. d. B., partly M. F.) in rooms of the schools of the participating students and were audiotaped. Due to technical problems three audiotapes could not be used, so seventeen were eventually used for analysis (named student 1 to student 17 from hereon).

The interviews were planned according to the 'problem-centered interview' of Witzel (2000). According to this method, social data of the students was established with the help of a short questionnaire, which also included some general questions on chemistry in school that were referred to as the warming-up phase, and which prepared them for the main phase of the interview. These three parts are crucial to the 'problem-centered interview' method. In the main phase students were asked to solve two specific stoichiometric tasks, which can be classified as 'calculations based on reaction equations'. The two tasks were chosen as being typical for revision problems in the field of stoichiometry. Therefore, they are rather complex compared with the easy-to-solve problems Schmidt suggested for an introductory stoichiometry course. Furthermore, the students were to tell the interviewer their general strategy of tackling problems of this type in between the two solutions (cf. Figure 1). This setting was chosen to see whether the students were able to detach their process from that specific situation and whether the strategy is consistent with the one used for the second task.

Figure 1: Setting of the interview (modified after Parchmann, et al., 2006).



Another aspect of the setting deals with the phrasing of the tasks. As a consequence of the TIMS- and PISA-studies it is said that problems dealt with in German schools are phrased too often one-dimensionally and in a too formal and abstract way, since 'real-world problems' are usually more multi-dimensional and information has to be extracted from the text (cf. Ralle, 2001). In order to investigate whether a more realistic, therefore more complex phrasing of a problem has an impact on students' success of solving it, the first task was constructed in two different versions with the same chemical content, but with different texts: one was very formal (formal version), the other was put into a short context story (context version) (cf. Figure 2 and Figure 3).

Figure 2. The formal version of the first stoichiometric task (translation into English by the authors).

> Ammonia (NH_3) reacts with hydrogen chloride (HCl) to form ammonium chloride (NH_4Cl). Calculate what masses of ammonia and of hydrogen chloride are needed to produce 4 kg of ammonium chloride.

Figure 3. The context version of the first stoichiometric task (translation into English by the authors).

> "Kids and grown-ups love it so, the happy world of HARIBO!" Who does not like the dark brown liquorice and its lovely taste?! Tasty above all is the liquorice-plant. As another ingredient "ammonium chloratum" very often is used as an acidifier. This name is apothecary's jargon and the old chemical name of ammonium chloride (NH_4Cl).
> Ammonium chloride, also known as "salt of salmiak", is allowed up to 20 g/kg in liquorice by the nutrition laws.
> Ammonium chloride can be produced from ammonia (NH_3) and hydrogen chloride (HCl).
> You are a student trainee with HARIBO's. Suddenly the machine in which the ammonia and the hydrogen chloride are combined and added to the liquorice sounds the alarm, but your boss is in his office on an urgent telephone call and does not want to be interrupted. The valves controlling the doses of ammonia and hydrogen chloride have to be adjusted. You have to act quickly. It is very important not to overdose the ammonium chloride, because its amount in the liquorice is controlled regularly by an independent supervision agency.
> The machine produces ammonium chloride for a thousand bags of 200 g liquorice each.

Of the seventeen students, ten worked on the formal version and seven on the context version. The second task (cf. Figure 4), which was used to determine the stability of the strategy, was the same for all students.

Figure 4. The second stoichiometric task (translation into English by the authors).

> In industrialized states every tenth adult suffers daily, and every third occasionally, from heartburn. In Germany almost ten million people are affected by this disease. Helpful medicines against heartburn are the so-called 'antacids'.
> Antacids bound acids against an over-acidification of the stomach. The stomach acid contains hydrochloric acid (HCl, dissolved in water), which reacts with the antacid to form water, carbon dioxide and the corresponding salt. Generally, antacids are carbonates. Very common is 'Bullrichs salt', which uses sodium hydrogen carbonate as active ingredient. A single dose of an antacid is to be chosen in a way that it will react with two grams of hydrochloric acid. You are to formulate the instruction leaflet for a pharmaceutical company, which sells calcium carbonate ($CaCO_3$) as an antacid. The dosing instruction says that in case of heartburn, one tablet with sufficient water is to be ingested.
> But the specification, how many grams of calcium carbonate one tablet contains, is still missing…

While solving the two tasks students could rely on a preliminary set of SST, which were constructed on the basis of the procedure in chemistry textbooks. These tools mostly gave supporting information on stoichiometric entities and connections between them. In addition to the preliminary SST, the interviewer was allowed to ask questions regarding the solution, and thus was able to let the interviewee become aware of mistakes. All participants were provided with paper, pencil and a pocket calculator. The notes students took while solving the problems were collected for further analysis. The periodic table of elements (PSE) was handed out on demand, to see if students knew where the information for determining molar masses is given.

Students were asked to explain their process aloud while working on the solution ('think-aloud' technique). The explanations were audio taped, transcribed and analysed.

Development and selection of categories for analysis

The coding of the interviews was done in three steps. Firstly, the phases of the student's solution process were marked and the parts of the interview, in which students referred to their pre-knowledge, to the material given to them such as the task sheet, the PSE or the prepared preliminary SST, or to the interviewer's help or explanations and questions respectively, were coded. Maps were produced of these codes, showing the solution process and specifying the parts in which the interviewer had intervened (cf. Figures 5, 6 and 7). The maps were used to visualise the solution process and were used for interpretation alongside the coded transcripts. Secondly, the quality of the interviewees' knowledge was evaluated according to a coding scheme based on the facets of performance formulated by Duit et al. (2001, p. 172; cf. to the steps of 'scientific literacy', Bybee, 1997). In line with our research questions, we adopted the following five categories:

Recognition of facts and information in the material: Is the interviewee able to recognise necessary information given in the material or is he/she not?

Knowledge of facts, definitions, etc.: Which facts and definitions does the interviewee know?

Explanation of connections: Is the interviewee able to explain the connections between stoichiometric definitions and concepts or is he/she not?

Strategy: Does the interviewee know which steps to carry out to solve the task?

Process: Does he/she know how to do it; is he/she able to do it appropriately and by him-/herself?

In a third step, the students' problems and misconceptions were coded. The category 'problem' was used to indicate smaller problems in comprehension or with definitions. The category 'misconception' was used, when the interviewee showed severe misunderstanding of a concept even after the interviewer's inquiry.

After having extracted the strategies, we analysed these according to BouJaoude and Barakat (2003). The categories Frazer and Servant (1986) and Schmidt (1994) formulated were found to be only partly applicable, because due to the complexity-level of the two problems, none of the participants solved the two tasks by mere logical reasoning (cf. Tóth and Kiss, 2005). Therefore, these categories were used only on a side line (see *calculation strategy*). Furthermore, the problems and misconceptions were correlated with the problem-solving strategies to see at which step problems occurred and how they could be categorized.

The general strategy that the interviewees were asked to explain was coded according to its steps and eventually visualised in a map, too. His/her general (theoretical) strategy was compared firstly to his/her applied strategy, and secondly to the strategy of an expert. This was done for each student.

Results

Students' strategies

Within one map, the strategy, explanations of certain steps, gaps in the strategy with help activities and the problems and misconceptions of a student can be seen. These maps were first analysed independently, and afterwards compared to each other to see if general features can be seen. They have to be read in columns (cf. Figures 5 and 6). The middle column reflects the steps in the solution process used by the student. His/Her ability to explain the process or the strategy on his/her own account is marked by boxes placed on the left hand column. All interventions and help of the interviewer and the material are marked by boxes placed on the right hand column. Therefore, one can see at one glance, whether a solution strategy is made up by the student him-/herself or whether he/she needed help while solving the task. The strategy itself can also be seen in this representation of the interview data. The white boxes (middle column) show the solution strategy of each student applied to the problem and can easily be compared to the student's general strategy. Therefore, one can see whether the general strategy is different from the applied strategy in the two tasks. The yellow (correct) and green (incorrect) boxes mark the steps in the solution in which explanations are given by the student. The red boxes in the maps show those parts of the solution process in which the interviewer or the material helped the interviewee by showing him/her the mistakes and telling him/her how to continue. Additionally, the problems and misconceptions are linked to the solution process. These are shown by the blue boxes.

Figure 5. Solution process of student 1, task 1. Colour code: white boxes = strategy steps, yellow boxes = student is able to explain connections of chemical concepts, olive green boxes = student is able to explain the strategy, light blue boxes = problems with certain content.

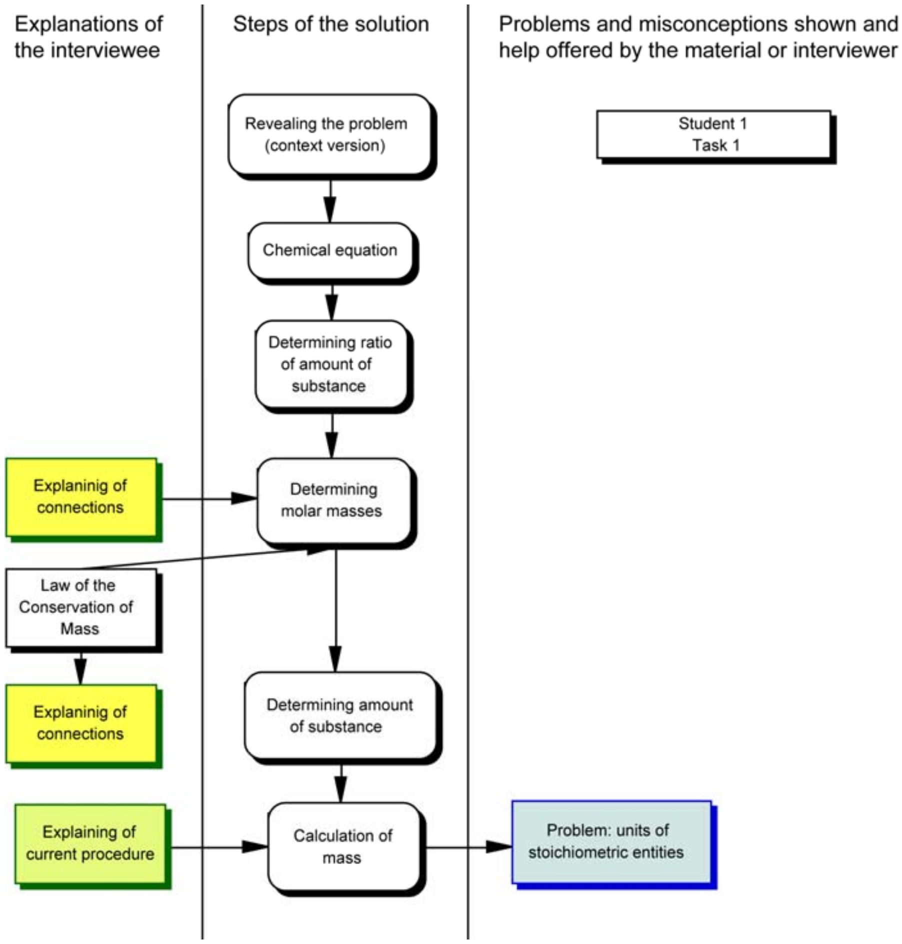
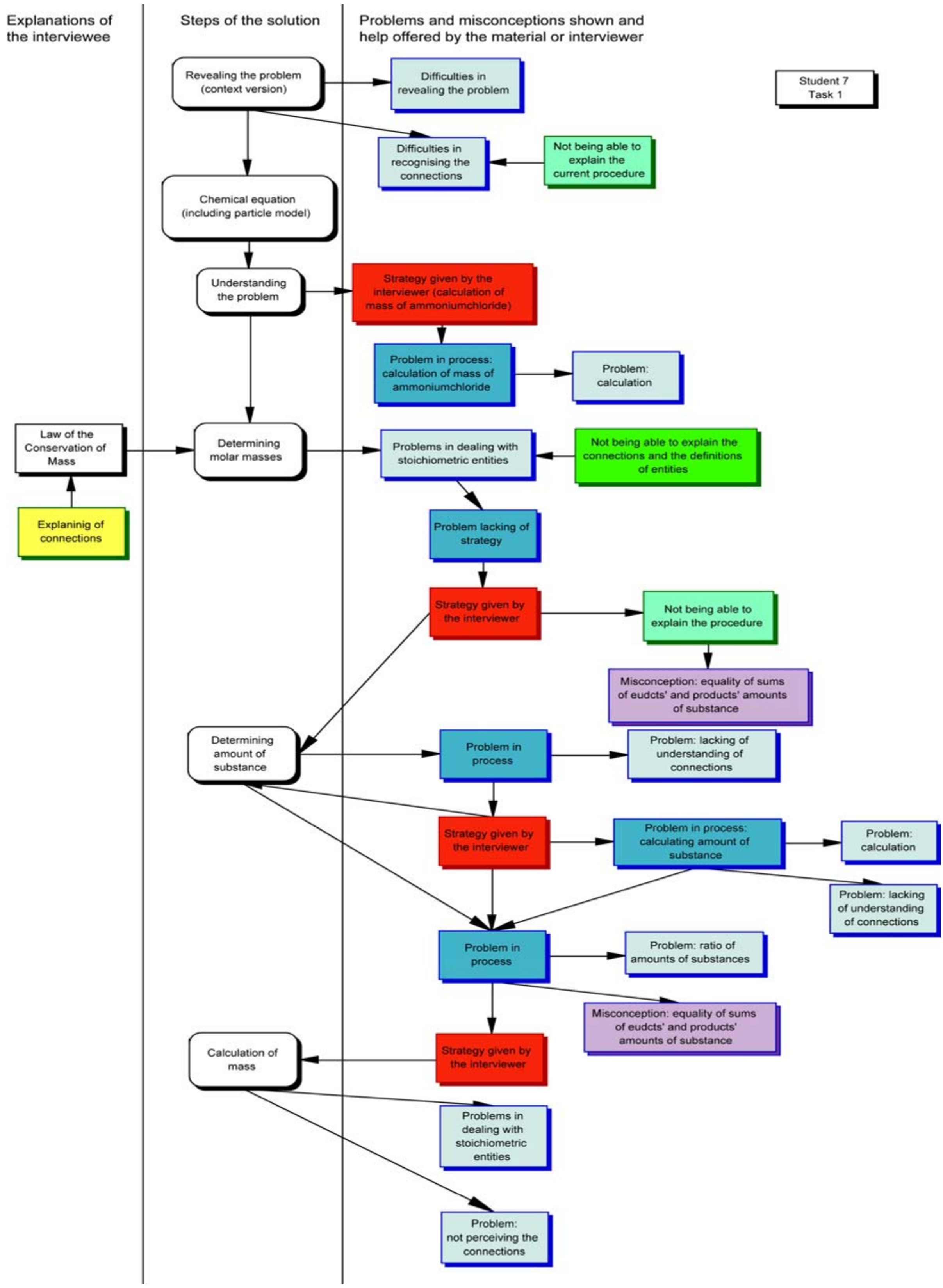


Figure 6. Solution process of student 7, task 1. Colour code: white boxes = strategy steps, yellow boxes = student is able to explain connections of chemical concepts, green boxes = student is not able to explain connections of chemical concepts, light blue boxes = problems with certain content, red boxes = strategy given by the interviewer, violet boxes = misconceptions, darker blue boxes = general problems in the solution process, turquoise green boxes = student is not able to explain the strategy.



The transcripts of the interviews and the maps show great individuality of the students' solution processes. At first sight, the map and thus the solution process seems to be unique for each student. All participants showed a variety of ideas and strategies, many of which led to correct answers. However, many misconceptions and problems reported in the literature were also found, which prevented some students from solving the tasks. In these cases the material and the interviewer showed the students the right way, since the goal of our study was to investigate which parts of the solution are most complicated and need special attention. Figure 5 and Figure 6 show the maps of two students for the first task, student 1 showing a self-contained and advanced strategy, whereas student 2 shows an unreflected, incomplete and chaotic strategy, which would not have succeeded without the help of the interviewer.

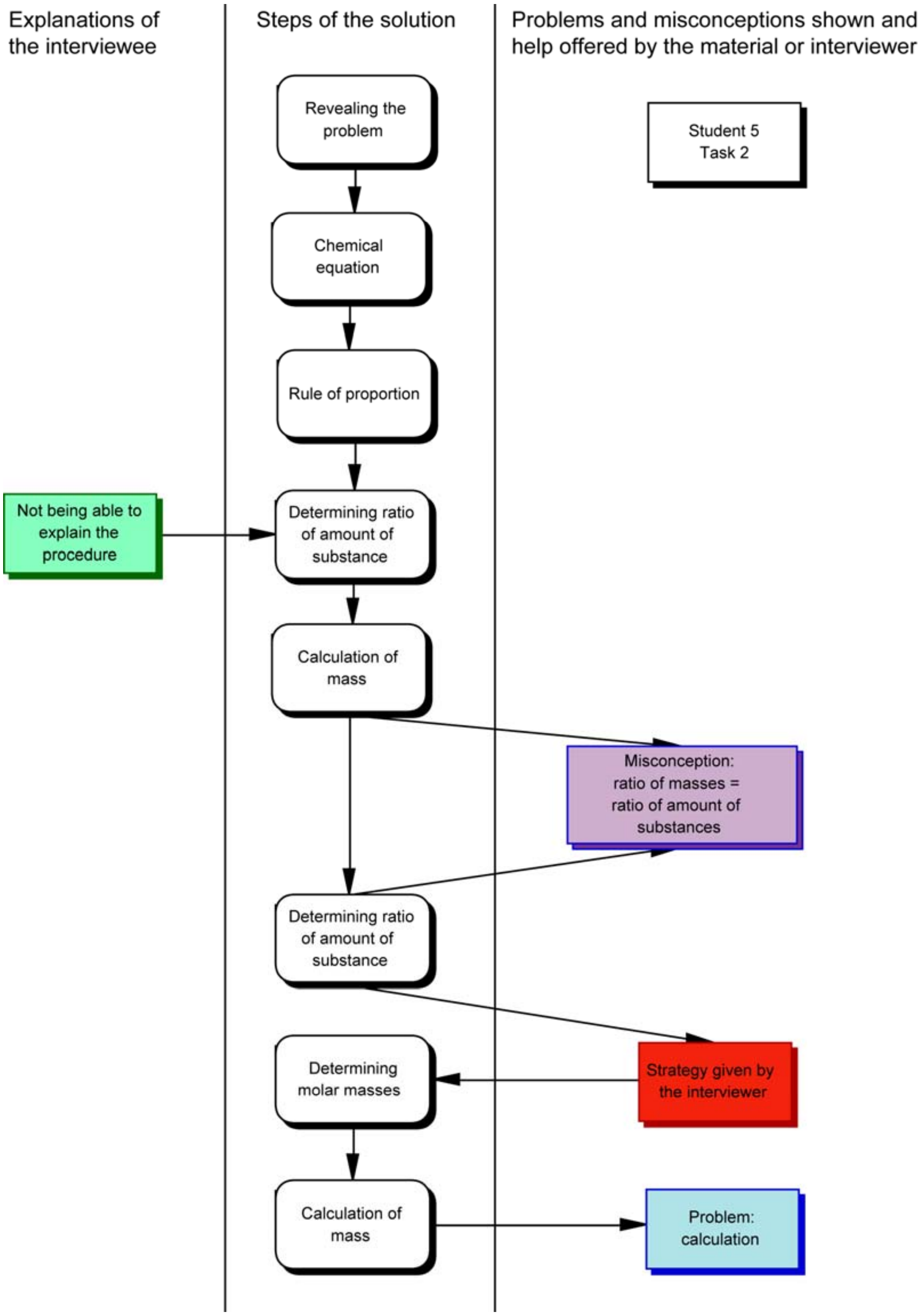
At second sight, one can see that all strategies contain up to six steps, which were sometimes combined with each other. These were:

- extracting the problem from the text given
- formulating the chemical equation
- calculating (the necessary) molar masses
- calculating the amount of substance
- considering the ratio of amount of substance
- calculating the mass.

Sometimes the step 'extracting the problem from the text given' was not explicitly mentioned by the students. One group of students combined the steps 'formulate the chemical equation', 'calculating molar masses', 'calculating the amount of substance' and 'considering the ratio of amount of substance' in a table which provided an overview of these entities. The order in which these steps were conducted sometimes varied between the interviewees. Since the step 'considering the ratio of amount of substance' did not need to be the last but one step, this was done right after formulating the chemical equation by some students. Another reason for varying the order of this scheme was that some steps were simply forgotten by the students, because they did not realise that all these steps are crucial when solving stoichiometric problems. Interestingly, students' not knowing *that* a certain step had to be done to solve the problem did not necessarily imply they did not know *how* to do this. Some students could perform steps correctly after having been told that this would lead to the correct solution by the interviewer or the preliminary SST. An example of a students' strategy showing this is given in Figure 7. The red box marks the step in which the interviewer said which step had to come next. The student, however, could then perform this step on his own.

From the maps it can be seen that it had no measurable impact on the solution process of the second task, whether the first task had been put in the formal version or the longer context-version. Generally, most students could cope better with the second task, but some still had severe problems. Only one of the students had more problems solving task two than she had had solving task one. Therefore, revision of stoichiometric terms and concepts seems to be an important factor when solving these problems repetitively.

Figure 7. Solution process of student 5, task 2. Colour code: white boxes = strategy steps, light blue boxes = problems with certain content, red boxes = strategy given by the interviewer, violet boxes = misconceptions, turquoise green boxes = student is not able to explain the strategy.



The nature of students' strategies

As mentioned above, the solution strategies were categorized according to BouJaoude and Barakat (2003). Since in our study students had the possibility to ask for help and thus mistakes were corrected during the interviews, which was both not the case in BouJaoude's and Barakat's study, the category 'incorrect answer' was not included. Rather, the nature of the help needed was categorized together with the solution strategy. Results are shown in Table 1.

Table 1. Categorization of students' problem solving strategies, based on BouJaoude and Barakat (2003).

Student	*Quality of the general strategy used by and quality of help needed by the student*
1	Algorithmic and reflective (efficient according to the stage of learning), needs little help on the second task.
2	Algorithmic, needs help during the solution of the first task.
3	Algorithmic, needs help during the solution of both tasks.
4	Algorithmic, needs help during the solution of both tasks.
5	Incomplete and messy, but on the way to an algorithmic strategy, needs help on both tasks.
6	Incomplete and chaotic, needs extensive help on both tasks.
7	Incomplete and chaotic, needs extensive help during the whole solution process.
8	Messy on first problem, needs extensive help, algorithmic on the second tasks.
9	Algorithmic, needs a key-word on the first task.
10	Algorithmic, needs a little help on the first task.
11	Algorithmic and reflective (efficient according to the stage of learning).
12	Incomplete, needs help on the first task, algorithmic on the second task.
13	Incomplete, mathematical-algorithmic, needs extensive help on both tasks.
14	Incomplete and chaotic, needs extensive help during the whole solution process.
15	Messy, but on the way to an algorithmic strategy, needs help on the first problem.
16	Algorithmic and reflective (efficient according to the stage of learning).
17	Algorithmic, needs help on the second task, difficulties with more complex tasks.

Table 1 shows that most students participating in our study either used an algorithmic and correct strategy, or an incomplete strategy in solving the problems, which meant that the latter students needed help while solving the tasks. Only three students showed a strategy which can be classified as 'reflective' or 'efficient to their state of learning'. For example, they calculated the mass of the second reagent in task 1 (cf. Figures 5 and 6) by using the law of conservation of masses and not by using the stoichiometric equations. However, while solving task 2 (cf. Figure 7), students only had to calculate the molar masses of hydrogen chloride and calcium carbonate to answer the question correctly, but every student started to calculate the molar masses of all three products, as well. Here they failed to apply an efficient strategy but used an algorithm ('calculate molar masses for all given substances') instead.

With most of the students, the general strategy and the applied strategies were almost the same, including the important steps and the order in which these were carried out. However, some of the students were only able to determine a strategy which reviewed the solution process of the first task, and thus was not independent from a specific given problem. Few others were unable to describe their general strategy after having solved the first task. These students were asked again to determine the strategy after they had solved the second task. If the general strategy obviously differed from the applied strategy on the first task, the interviewer asked further questions. In consequence, most students looked more carefully at their solution, identifying further steps. These were especially marked in the transcripts as being helped by the interviewer. The fact that some of the students were not able to describe a general strategy indicates that they are not able to abstract from a specific given task. This has to be discussed in the context of metacognition, but since it was only a subgoal of the study to investigate if the students' general and the applied strategies were the same, we shall not take it any further.

Calculation strategy – different kinds with different success

The strategies were then analysed further. The students were grouped according to their teachers, and the strategies of the students of one group were compared to each other both within groups and across groups. In this part of the analysis, we used the categories Frazer and Servant (1986) and Schmidt (1994) established. The maps show that the students used two different kinds of calculation strategies, some solving the tasks by using the stoichiometric equation

(1) $m = n \cdot M$,
(where m: mass, n: amount of substance, M: molar mass)

to calculate the amount of substance and then using the ratio of amount of substances for solving the two tasks. This is analogous to the 'mole method' as reported by Schmidt (1994). Others used the 'rule of proportion' or the 'rule of three' to calculate directly the required mass and thus avoiding the use of the concept 'amount of substance':

(2) $$\frac{M_1 \cdot n_1}{M_2 \cdot n_2} = \frac{m_1}{m_2}.$$

This is an analogous procedure to using the 'proportional method' as reported by e.g. Schmidt (1994). These students often multiplied the coefficients given in the chemical equation with the molar masses of the substances. Doing this on the one side leads to a correct answer, but on the other side shows an incorrect understanding of the concept of molar masses (cf. BouJaoude and Barakat, 2000; Frazer and Servant, 1986). Interestingly, the 'rule of proportion' method caused more trouble during the calculation, because students often mixed up numerators and denominators which led to false results. Table 2 shows the kind of calculation strategy used by the students and the problems that occurred.

Table 2. Students' strategies used for calculating the tasks and the nature of their difficulties.

Student	*Type of calculation strategy*
1	Equation, no problems during calculation.
2	Equation, problems with ratio of amount of substance during calculation.
3	Equation, problems with ratio of amount of substance during calculation.
4	Rule of proportion, problems with calculating masses and with amount of substance.
5	Rule of proportion, problems with calculating masses and with amount of substance.
6	Starting with rule of proportion, problems with calculating masses, after being helped switched to use the equation.
7	Mixture of rule of proportion and equation, massive problems with calculating.
8	Starting with rule of proportion, problems with calculating masses, after being helped switched to use the equation.
9	Rule of proportion, problems with calculating masses on the first task.
10	Rule of proportion, problems with calculating masses on the first task.
11	Equation, no problems during calculation.
12	Mentions the rule of proportion in the general strategy, but uses the equation in the applied strategies, problems with calculating masses and with ratio of amount of substances.
13	Rule of proportion, problems with entities and with ratio of amount of substances.
14	Starting with rule of proportion, problems with calculating masses, after being helped switched to use the equation.
15	Equation, problems with calculating masses and with amount of substance.
16	Equation, no problems during calculation.
17	Equation, problems with calculating masses and with ratio of amount of substances on the second task.

It can be seen that seven of the seventeen students used the equation to solve the tasks, while five used the rule of proportion. Often students of this latter group had problems when mixing up the given and wanted values, which may have been due to a simple mistake, but it could also have been due to a lack of understanding of stoichiometric entities like molar masses and amounts of substances. Another frequently observed problem concerned the ratio of amount of substance. Here both the students using the equation and the students using the rule of proportion had problems in applying the ratio correctly to their calculation. A third group, consisting of three students, started their calculations using the rule of proportion, but switched to using the equation after meeting problems doing the calculation, and being helped by the interviewer. The interviewer presented them the equation within one card of the preliminary SST. Interestingly, all these students coped better with the equation than with the rule of proportion. One said at the end of the interview that he will use the equation on all stoichiometric problems from now on instead of using the rule of proportion, "*because it all becomes easier*."[1] Of the remaining two students one mentioned the rule of proportion in his general strategy but actually used the equation while doing the calculation, another used a mixture of rule of proportion and equation. Both of them had the above mentioned problems in doing the calculations.

Influence of the teacher

An interesting observation can be made, when assigning the results of Table 2 to the groups formed by 'teacher-affiliation'. As mentioned above, the seventeen students were taught by four teachers. Students 1 and 2 were taught by teacher A, students 3-10 by teacher B, students 11-14 by teacher C and students 15-17 by teacher D. As can be seen in Table 2, the students of teachers A and D used the same mathematical strategy, that is, they used the equation to solve stoichiometric problems. In contrast, students of teacher B mostly used the rule of proportion, only one of them used the equation. Two of them switched to using the equation while doing the calculation. On these thirteen students it can be seen that they generally use a solution strategy, which was most likely the one taught by their teachers, and only occasionally applied a different strategy to these tasks. The four students of teacher C showed a somewhat different behaviour in carrying out quite diverse strategies, but this may be explained with the fact that their classes had been regrouped changing from grade 8 to grade 9, as one of those students said after the interview. Therefore, those students had probably been introduced to stoichiometry in different ways, because their previous teachers all showed them different methods of solving stoichiometric problems. This was confirmed by one student, saying that he was highly confused on how to best solve those problems, because the teacher in grade 8 had taught them differently from the teacher in grade 9. As expected, the way the teacher actually teaches stoichiometry seems to have a great impact on how young students solve stoichiometric problems and teaching should therefore be prepared with great care (cf. Tóth and Kiss, 2005).

Problems and misconceptions

Another aim of the study was to investigate which problems and misconceptions occurred within the solution process. Table 3 gives an overwiew of the problems and misconceptions found in the interviews.

[1] All students' quotations were translated from German into English by the authors.

Table 3. Students' problems (p) and misconceptions (m) in the different steps of the solution process. The numbers say how many students had these problems and misconceptions

Part of solution process	*Problems (p) and misconceptions (m)*	*Number of students*
chemical equation	p: balancing chemical equation	4
	p: confusing 'atoms', 'molecules' and 'particles'	1
stoichiometric entities		
- non-specific	p: dealing with entities in general	9
	p: dealing with units	2
- molar mass	p: molar mass and coefficients/indices in chemical equation	8
	p: differentiating between independent entities (constants, e.g. molar mass) and dependent entities (e.g. mass)	5
	p: handling of PSE	7
- amount of substance	m: "*amount of substance cannot be less than one mole*"	3
ratio of amount of substance	p: realizing the ratio of amounts of substance needs to be taken into account	8
	m: "*sums of amounts of starting materials and products have to be the same*"	3
	m: "*ratio of masses = ratio of amounts of substance*"	4
mass	p: calculation	2
	p: two unknowns in one equation	3
	p: differentiating between independent entities (constants, e.g. molar mass) and dependent entities (e.g. mass)	3

As said above minor difficulties were classified as 'problems' and the category 'misconception' was only used when students showed severe misunderstandings of a concept. From Table 3 it can be seen that most problems occurred with regard to stoichiometric entities. Many students knew the definitions of these entities, but did not show a profound understanding, i.e. could not tell the connections between them. Some students mixed up stoichiometric concepts, revealing poor understanding, e.g.: "*The mole-number is, how many particles are contained in ... one u or in one atom ... something like that (student 17)*". These problems in understanding are described in previous studies, as well (cf. Furió, et al., 2002). Additionally, some problems occurred regarding the chemical equation. Most students knew that formulating the chemical equation is necessary for solving the task, but many did not know what information is provided in the equation, i.e. the ratio of amounts of substance. Therefore, they failed when the ratio had to be taken into account. One student put up the equation only with the names of the substances involved in the reaction and not with their chemical symbols. Being asked wherefrom he knew the ratio of the amounts of substances, he said: "*I don't know – gut feeling*" (student 3). Another great difficulty presented the 'ratio of amount of substance', which many students did not take into account or, as can be seen in the above quote, had no idea where to derive it from. Interestingly, only minor problems occurred doing the actual calculations.

Of the misconceptions described in the literature, many were found in our study as well. For example, four students equalled the ratio of masses with the ratio of amount of substance, as reported by Schmidt (1990). Another group of three students said that in a chemical equation the sums of the coefficients of starting materials and products and therefore their amounts of substance have to equal each other, as it is in a mathematical equation (cf. also

Figure 6, in which this misconception is shown). This indicates a misconception similar to the 'confusion of conservation of atoms and possible non-conservation of molecules', as reported by Mitchell and Gunstone (1984). One misconception that we found, however, has not been reported in literature, yet: "Amount of substance cannot be less than one mole". Thus,in task 2, the amount of substance of the hydrochloric acid is 2 g / 36.5 g/mol = 0.055 mol; three students stumbled over this; they wondered how an amount of substance could be less than one mole, as can be seen in the following two quotations (translation into English is by the authors).

S 6: 0.054794 ... uh, no, that's not possible...
I: Why not?
S 6: Because it's not even one mole. That's ... less than one mole.
I: And you're wondering about that?
S 6: Yes, I am.
[Pause for thinking – student 3 is thinking how to calculate the amount of substance of the hydrochloric acid]
S 3: OK, then I don't have a whole ... no, that isn't possible.
I: What do you want to calculate?
S 3: How many moles are in two grams ... of hydrochloric acid. But that has to be less than one mole.
I: Yes.
S 3: Is that correct?
I: If one mole is 36.5 g...
S 3: Yeah, but I mean is it correct that it's less than one mole?
I: Think about it for yourself. If one mole weighs 36.5 g ... and now you only have two grams...
S 3: I see that it's less than one mole. But I mean is it possible that it's like that?

A possible reason for the occurrence of this misconception is that in German chemistry lessons the mole is often introduced as a particle number rather than as an amount of substance. These students seem to believe that one mole *means the same as* one particle. Furthermore, the information that one mole is the amount of substance which represents the number of $6x10^{23}$ particles seemed to be meaningless for them, although they did know it. This is corroborated by other parts of the interviews in which the definition of one mole was repeated by all these three students. Nevertheless, they were not able to see that their calculation of the amount of substance of the hydrochloric acid was correct.

Discussion and implications – development of SST as teaching materials

If the students participating in this study solved stoichiometric problems correctly, they did it algorithmically. The fact that we chose contextualised problems with a rather complex setting did not show any effect on the strategy of solving the two tasks or the success in doing this. Solving stoichiometric tasks in this case included formulating the chemical equation, determining molar masses, calculating amount of substance, determining the ratio of amounts of substances and, finally, calculating the mass; therefore, these tasks can be classified as 'calculations based on the reaction equation'. The last three steps were either done in three separate steps by the students, using the stoichiometric equation, or combined in one step by calculating via the rule of proportion. The order in which the steps were carried out differed from student to student.

Nearly all students taught by the same teacher used a similar strategy, so the scheme the teachers used when introducing stoichiometry seems to have a great impact on students' solving strategies. This is in agreement with the results of Tóth and Kiss (2005), but in contrast to the results of Schmidt (1994). One possible reason for this is that the students participating in our study – like the Hungarian students participating in Tóth's and Kiss's study – were novices in the field of stoichiometry compared to the students participating in Schmidt's study. Beginners in solving these specific stoichiometric problems seem to tend to rely on a given structure, which may be due to the relative complexity of the solution of those problems, consisting of five single steps, rather than applying their own strategy which can be more efficient and thus save time. Therefore, an introductory stoichiometry course should provide students with a (self-developed?) scheme which includes all the necessary steps. From the fact that every student participating in our study used such a scheme, it may be concluded that this already happens in many classes and stoichiometry courses. With regard to the content of this scheme the data indicate that using equation (1) and thus explicitly using the term 'amount of substance' may be preferable to using the proportional equation (2) (see *calculation strategy*). This is corroborated by the results of Frazer and Servant (1986).

However, quite a lot of the students had only incomplete strategies for solving the tasks. If these students are to develop a viable strategy and thus are able to solve stoichiometric problems, they will have to be supported in their learning process, e.g. by SST. Three different kinds of incomplete strategies were found: To one group of students certain steps were unknown, and after being told they still did not know how to carry them out. A second group did know all the necessary steps but had problems in conducting single steps properly. Interestingly, a third group of students forgot that certain steps had to be carried out, but after being told, could do them correctly. In addition, some students had problems in working out the actual task from the given text, i.e. deciding which information is crucial for solving the tasks. Reviewing the calculation result for correctness was done only by very few students. From this, we derive four categories for SST. For this kind of stoichiometric problems given to a class for repetitive purpose, these SST may be:

1. A set of '(learning-) strategic SST'. These include general hints of how to work out the actual task from a longer text, and how to review the result and thus prove its correctness.
2. A set of 'content-strategic SST'. Here the steps of a possible solution are given, but only little information is provided on how these are carried out.
3. A set of 'content-related SST'. For every step of this solution, a detailed description is provided, which says how to do it.
4. A general glossary of important terms. Taking into account the difficulties many students had in explaining or defining stoichiometric entities correctly, an additional glossary of stoichiometric terms should be provided, giving background information and showing the connections between these entities.

The difficulties and misconceptions that were found in this study mostly occur with regard to the chemical content of the problems, i.e. the chemical equations and stoichiometric entities. Many difficulties and misconceptions described in the literature were found here too, but also an additional one that has not been described before. The actual calculation, i.e. doing the necessary mathematics, caused only minor problems.

Since quite a lot of new terms are being introduced within the topic of stoichiometry, which often sound similar to each other or include related concepts (e.g. the 'mole', 'molar mass', 'amount of substance', 'number of particles', etc.), beginners in stoichiometry should be given a chance to review these definitions while practising stoichiometric problems. This seems also appropriate because many misconceptions are likely to arise when definitions and connections of these terms and concepts are misunderstood.

The method used in this study works; however, it must also be seen critically. Since the interviewer was allowed to intervene in the solution process depending on the questions and problems of the interviewees, these interventions were highly individual and thus could not be planned a priori. The preliminary set of SST proved only to be partly applicable in most situations, but since it was the aim of the study to develop a sophisticated set of SST, this was assumed before starting the study. The interviewer has to act very carefully in the intervention phase so as not to push the students in a certain direction. Additionally, analysis has to be carried out thoroughly. Since the results of this study are corroborated by those of other studies, we think that the method is valid and can be transferred to other topics of science. For example, the first three sets of SST are likely to be appropriate for chemical problems in general. This is corroborated by the findings of studies on problem solving in chemistry (cf. Gabel and Bunce, 1994). A recent German study (ForschergruppeKassel, 2004) on physical and chemical problems on density, force and solubility found a combination of learning-strategic and content-related SST to be more helpful than either of them alone. For the more complex stoichiometric problems, the additional set of content-strategic SST should be added to help those students who have forgotten one of the several steps to be conducted.

The publication of the specific design of the SST is in preparation. Further evaluative research must show if this design is successful in helping students to solve stoichiometric problems, especially due to the fact that only seventeen students participated in this study. If it proves to be so, these categories for SST will be extended to other chemical topics. Thus, a catalogue of SST for chemical problems can be assembled and given to students to generally improve their achievement in chemistry.

Acknowledgements

We would like to thank Dr. Verena Reineke, whose comments were very helpful in improving the manuscript.

References

Barker V., (2000), *Beyond appearances: students' misconceptions about basic chemical ideas: a report prepared for the Royal Society of Chemistry*, Education Division, Royal Society of Chemistry London.

BouJaoude S. and Barakat H., (2000), Secondary school students' difficulties with stoichiometry, *School Science Review,* **81***,* No. 296, 91-98.

BouJaoude S. and Barakat H., (2003), Students' problem solving strategies in stoichiometry and their relationships to conceptual understanding and learning approaches, *Electronic Journal of Science Education,* **7***,* No.3, online journal, http://unr.edu/homepage/jcannon/ejse/ejse.html.

Bybee R.W., (1997), *Achieving scientific literacy: from purposes to practices*, Portsmouth, NH: Heinemann Publishing.

Duit R., Häußler P. and Prenzel M., (2001), Schulleistungen im Bereich der naturwissenschaftlichen Bildung [School performances in the field of science education], In Weinert, F.E. (ed.), *Leistungsmessungen in Schule [Assessment in school]*, Weinheim, Basel: Beltz Verlag.

Duit R., Gropengießer H. and Kattmann U., (2005), Towards science education research that is relevant for improving practice: the model of educational reconstruction, In Fischer, H.E. (ed.), *Developing standards in research on science education*, London: Taylor and Francis.

Entwistle N. and Ramsden P., (1983), *Understanding student learning*, New York: Nichols Publishing Company.

Fiebig S. and Melle I., (2001), "Problemthemen" des Chemieunterrichts in der Sekundarstufe I - gegenwärtige Situation und fachdidaktische Konsequenzen ["Problematic issues" in secondary school chemistry lessons - present situation and educational consequences], *CHEMKon,* **8***,* No. 4, 199-202.

ForschergruppeKassel, (2004), Aufgaben mit gestuften Lernhilfen [Stepped supporting tools for problems in school science], *Lernchancen,* **7***,* No. 41, 38-39.

ForschergruppeKassel, (2006), Archimedes und die Sache mit der Badewanne – Gestufte Hilfen im naturwissenschaftlichen Unterricht [Archimedes and the question about the bathtub – stepped supporting tools in science education], *Friedrich Jahresheft - Diagnostizieren und Fördern, No. XXIV*, 84-88.

Frazer M.J. and Servant D., (1986), Aspects of stoichiometry titration calculations, *Education in Chemistry,* **23***,* 54-56.

Frazer M.J. and Servant D., (1987), Aspects of stoichiometry – where do students go wrong? *Education in Chemistry,* **24,** 73-75.

Furió C., Azcona R. and Guisasola J., (2002), The learning and teaching of the concepts 'amount of substance' and 'mole': a review of the literature, *Chemistry Education Research and Practice,* **3***,* 277-292.

Gabel D.L. and Bunce D.M., (1994), Research on problem solving: chemistry, In Gabel, D.L. (ed.), *Handbook of research on science teaching and learning*, New York: Mac Millan.

Griffiths A.K., (1994), A critical analysis and synthesis of research on students' chemistry misconceptions, In Schmidt, H.-J. (ed.), *Problem solving and misconceptions in chemistry and physics*, Hong Kong: ICASE.

Huddle P.A. and Pillay A.E., (1996), An in-depth study of misconceptions in stoichiometry and chemical equilibrium at a South African university, *Journal of Research in Science Teaching,* **33***,* 65-77.

Kaminski B., Jansen W. and Flint A., (1994), Chemische Formeln im Anfangsunterricht [Chemical formulae in freshmen chemistry courses], *CHEMKon,* **1***,* No. 4, 183-188.

Lijnse P.L., (1995), 'Developmental research' as a way to empirically based 'didactical structure' of science, *Science Education,* **79***,* 189-199.

Mitchell I.J. and Gunstone R.F., (1984), Some student conceptions brought to the study of stoichiometry, *Research in Science Education,* **14***,* 78-88.

Packer J.E., (1980), Letters to the editor – titrimetric calculations, *Education in Chemistry,* **17***,* 154.

Parchmann I., Gräsel C., Baer A., Nentwig P., Demuth R., Ralle, B. and the ChiK Project Group (2006), 'Chemie im Kontext': A symbiotic implementation of a context-based teaching and learning approach, *International Journal of Science Education,* **28***,* 1041-1062.

Ralle B., (2001), Eine veränderte Aufgabenkultur als Herausforderung [A different culture of using tasks as a challenge], *Der Mathematisch-Naturwissenschaftliche Unterricht (MNU),* **54,** 387.

Rossa E., (1998), Das Mol im Bild [The mole in a picture], *Chemie in der Schule,* **45***,* 8-13.

Schmidt H.-J., (1990), Secondary school students' strategies in stoichiometry, *International Journal of Science Education,* **12***,* 457-471.

Schmidt, H.-J. (1994). Stoichiometric problem solving in high school chemistry, *International Journal of Science Education,* **16***,* 191-200

Schmidt H.-J. and Jignéus C., (2003), Students' strategies in solving algorithmic stoichiometry problems, *Chemistry Education: Research and Practice,* **4***,* 305-317.

Tóth Z. and Kiss E., (2005), Hungarian secondary school students' strategies in solving stoichiometric problems, *Journal of Science Education - Revista de educación en ciencias,* **6***,* 47-49.

Williams D.G., (1980), Letters to the Editor – titrimetric calculations, *Education in Chemistry,* **17***,* 182.

Witzel A., (2000), The problem-centered interview, *Forum: Qualitative Social Research,* **1**, No. 1, online-journal, http://qualitative-research.net/fqs.

Predicting at-risk students in general chemistry: comparing formal thought to a general achievement measure

Scott E. Lewis and Jennifer E. Lewis*,

Department of Chemistry, University of South Florida, FL, 33620, USA
e-mail: jlewis@cas.usf.edu

Received 10 August 2006, accepted 1 January 2007

Abstract: This study is an investigation into the ability of pre-assessment measures of formal thought ability and general achievement to predict students at-risk of poor performance in college-level general chemistry. Over a three year period, data on formal thought ability (as measured by the Test of Logical Thinking, or TOLT) and/or general achievement (as measured by the Scholastic Aptitude Test, or SAT) was collected from over 3000 students as they entered a general chemistry course. The outcome measure was an American Chemical Society general chemistry exam at the end of the course. Findings indicate that both the formal thought and the general achievement measure can successfully identify at-risk students in this setting, with neither measure being superior in doing so. The presence of distinct groups of students correctly predicted to be at-risk by only one of the measures demonstrates that formal thought ability and general achievement each represent an independent hindrance to success in chemistry. Therefore, efforts to help at-risk students should include a focus on the development of formal thought as well as a content review. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 32-51]

Keywords: Assessment, formal thought, at-risk students, performance predictors, college chemistry

Introduction

All too often, substantial numbers of students in college fail to demonstrate sufficient understanding of chemistry to proceed beyond the introductory course, general chemistry. This circumstance hinders not only the individual student but also the field of chemistry. While the costs to the individual are immediate and obvious (not only the regrettable lack of knowledge of chemistry but also a closed door to any major field of study requiring that knowledge), the costs to chemistry are also significant. With each year this trend continues, chemistry loses numerous individuals who now will not contribute to the growth of the discipline. Indeed the ramifications stretch beyond chemistry, as other science curricula require general chemistry prior to course work within their program (Tai et al., 2005). Students who cannot muster an acceptable understanding of general chemistry are prevented from contributing to many science fields. On a more systemic level, the inability of students to continue in science-oriented courses because of low performance in general chemistry represents a major setback in efforts to create a scientifically-informed populace and a technically-proficient workforce. For these reasons, unsatisfactory student performance in college-level general chemistry remains a critical area of concern.

Since basic constructivism indicates that the prior knowledge and skills with which students enter a course play a role in success (or its absence), it is both possible and valuable to identify students who are at-risk of not succeeding in a course at the point when they first enter the course. To do so provides the opportunity for assisting these students early on, while success is still possible. Further, knowledge about the factors contributing to low (at-risk) performance can inform the design of interventions aimed toward reducing the challenges faced by these students. The first task is identifying the at-risk population with reasonable accuracy, and the second is suggesting potential interventions. Ideally, the measure used for identification contains within itself implications for a potential remedy. This paper compares the accuracy, degree of overlap, and implications for potential interventions of two measures that can be used to identify students at-risk of not succeeding in general chemistry. It therefore joins a long history of 'predictor papers' but is unique in its combination of generalizability, a focus on at-risk students, and consideration for the implications of choosing a particular predictor.

The need for more work with predictors

Extensive work has been done on the ability to predict success in college chemistry. Past studies of college chemistry have examined the ability of SAT (Pederson, 1975; Pickering, 1975; Bender and Milakofsky, 1982; Craney and Armstrong, 1985; Nordstrom, 1990; Bunce and Hutchinson, 1993; Spencer, 1996), ACT (Carmichael et al., 1986; Nordstrom, 1990 House, 1995), high school GPA (Carmichael et al., 1986), high school chemistry grade (Ozsogomonyan and Loftus, 1979; Craney and Armstrong, 1985; Nordstrom, 1990), personality characteristics (House, 1995) and Piagetian tasks (Bender and Milakofsky, 1982; Bunce and Hutchinson, 1993) to predict final chemistry course grade. In all these studies, however, the use of chemistry grade as an outcome variable relies on the ability of chemistry grade to approximate chemistry understanding. The extent to which this approximation is valid depends on several decisions peculiar to the course, the instructor and the institution. Decisions such as grading on a curve or an absolute scale, grading based completely on exam performance versus consideration of student homework, the allowance of extra credit, and even the method by which each exam was created, can all alter the extent to which chemistry grades reflect true student understanding of chemistry. As a result, the generalizability of the above studies depends on whether all of these factors are handled the same way at other institutions. Of the studies presented above, the work by Bender is the only one to provide detailed evidence of the grading procedures employed so that a replication could be attempted at another institution.

A more replicable option for an outcome variable is the use of a single exam as a measure of students' chemistry understanding. The exam questions can readily be made available for scrutiny in order to provide a clear picture of what constitutes success in chemistry, with none of the ambiguity surrounding course grade. In addition, the scoring of a single exam lends itself readily to the statistical procedures commonly used with predictors. One example of such a procedure is present in Yager et al.'s examination of the effects of taking high school chemistry (Yager et al., 1988). In this study, students were measured on a standard exam, a course final exam, and by a final course grade to provide multiple measures of success in chemistry. In particular, the use of a standard exam allows for a ready assessment of generalizability.

Finally, considerable work has gone into the development of chemistry-based diagnostic exams for course placement and prediction of performance (Russell, 1994; McFate and Olmsted III, 1999). These instruments tend to incorporate both math and chemistry questions and can be said to measure chemistry ability rather than incoming chemistry-specific knowledge. Such instruments seem to have a reasonably high success rate in predicting chemistry grades (McFate and Olmsted III, 1999; Legg et al., 2001; Wagner et al., 2002), but leave open the question of what can be done to assist the students who score low on such measures. Some suggestions put forth include recommending increased study time or remedial coursework for such students. Similar suggestions are presented in Yager's study**.**

But these suggestions do not necessarily lead to specific remedies: for example, what should be the design and intent of remedial coursework? How should the increased study time be spent? Even in the case of chemistry-focused exams that can identify specific deficiencies indicating that students did not achieve a sufficient understanding of chemistry via their earlier chemistry courses, how should we construct a second attempt to teach these concepts so that it will be successful? These questions are of particular importance, since recent research has suggested that remedial coursework may offer only marginal improvements in chemistry success (Bentley and Gellene, 2005; Jones and Gellene, 2005).

Our dissatisfaction with the remedies that can be offered on the basis of chemistry diagnostic exams or high school chemistry GPAs led us to the most important facet of our study. In particular, we wanted to compare two potential methods of identifying at-risk students that would suggest slightly different remedies in order to consider whether either, neither, or both remedies are tenable. Our study therefore compares two predictors, one with a long history of success at predicting course grades (SAT score), and another with a sound theoretical underpinning (formal thought ability) but with less information available as to its efficacy as a predictor in a college chemistry setting. As will be discussed in the next section, formal thought ability has a theoretical link to specific chemistry topics, and a research base aimed at improving formal thought performance (Lawson and Nordland, 1976; Adey and Shayer, 1990; Shayer and Adey, 1992a, 1992b, 1993) means that interventions to improve formal thought could be readily applied. In a similar fashion, the role of Math SAT in predicting performance implies that math skills are responsible for success, which would lead to specific suggestions of additional math course work or tutorials.

Although our study joins a long history of predictor papers, no one has yet offered a replicable predictor study that contains within itself clear guidelines for the construction of remedies. Further, our focus is predicting students at-risk of performing poorly in general chemistry, since it is for these students that interventions are needed. This means we look specifically at how well the predictors in our study identify students who perform at the lower end of our outcome measure, something which few previous studies have done. (Notable exceptions are Legg (Legg et al. 2001) and Wagner (Wagner et al., 2002)). For our outcome measure, as a result of considering the limitations of previous studies based on course grades (discussed above), we chose a standard exam designed to measure student understanding of chemistry. Our results are therefore generalizable to the extent that the content of this exam matches the desired outcomes at other institutions. The exam is available to the public, so this determination can be made (Examinations Institute of the American Chemical Society, 1997). Further, our study allows us to see whether, in the specific case of college chemistry performance, a simple paper and pencil measure of formal thought ability, the Test of Logical Thinking (TOLT), can stand up against the SAT's successful history.

Formal thought and science achievement

With the intent of identifying at-risk students in a way that would inherently suggest a particular remedy, our predictor selections had to be focused on measures that have the potential to describe a large hindrance for students. Because of its basis in a well-described learning theory, the construct of formal thought offers the ability to suggest specific difficulties students face, leading to specific remedies. Formal thought has been described as one of a series of factors necessary for a successful performance (Lawson, 1979, 1983; Chandran et al., 1987), so the *absence* of formal thought would definitely be expected to lead to a poor performance – exactly what is necessary for a good predictor of at-risk status.

Formal operational thought is the last stage of cognitive development as described by Piaget, in which 'deduction no longer refers directly to perceived reality but to hypothetical statements' (Inhelder and Piaget, 1958). In formal thought, possibilities are regarded as hypothetical at first, and then verified by empirical evidence: in short, deductive reasoning. Contextually, this leads to the meaningful manipulation of empirical results, as well as a familiarity with the abstract. Also taken from Piaget's work is a series of reasoning patterns that would describe formal thought operations. Adey and Shayer (1994) grouped the reasoning patterns into three main categories. The first category, the handling of variables, includes the control and exclusion of variables, the recognition of multiple classification schemes, and the description of combinatorial possibilities. The second category, relationships between variables, includes the use of ratios, and proportion (comparing of two ratios), as well as compensation (use of inverse relationships), correlation and probability. The final group, formal models, describes the creation of an abstract representation of complex behaviours. Also included in this last group is the use of logical reasoning. Within Piagetian theory, the onset of formal thought would be characterized by the development of all the cognitive operations at about the same time, a postulate that has been supported by empirical evidence (Lawson and Renner, 1975; Lawson and Nordland, 1976; Lawson et al., 1978).

Certain aspects of formal thought have been suggested as explaining the difficulty some students face in chemistry. The second category, relationships between variables, for example could explain an inability to relate mathematical formulas to underlying concepts, a task frequently required in chemistry. In keeping with this idea, some researchers have hypothesized which chemistry concepts require formal thought (Herron, 1975), and others have investigated links between formal reasoning ability and conceptual understanding of specific topics in chemistry (Abraham and Williamson, 1994; Demerouti et al., 2004). Neo-Piagetian theories of learning still incorporate formal thought ability as one of several critical cognitive factors important for problem-solving in chemistry (Niaz, 1987, 1996; Tsaparlis 2005). Tsaparlis et al. investigated the effects of several cognitive variables on student performance on several types of molecular equilibrium problems and found that developmental level in terms of formal thought ability was the most important predictor of success; however, additional work led to the identification of developmental level as a potentially confounding factor in studies using chemistry problems with complex logical structures to investigate the importance of working memory capacity (Tsaparlis et al., 1998; Tsaparlis and Angelopoulos, 2000).

Formal thought has been postulated as a necessary condition, either directly or indirectly, for conceptual change to occur (Oliva, 2003). Thus, in addition to describing students' incoming abilities, formal thought may play a role in whether and how students actively incorporate new information presented in the course. One early example is the work of Lawson and Renner (1975), who showed that students at the concrete operational stage are unable to develop an understanding of formal concepts, and that students at the formal operational stage demonstrate an understanding of both formal and concrete concepts. Lawson (1982, 1985) pointed out that such results could be interpreted largely as a spurious correlation, describing what might be a more general intelligence measure underlying the success seen on both measures. In the 1982 study, a partial correlation between formal thought and biology achievement while controlling for fluid intelligence revealed a significant relation, illustrating that it was the formal thought measure that better corresponded to this biology achievement measure.

We continue this line of investigation by examining whether formal thought features a unique relationship to overall achievement in college chemistry. While Lawson demonstrated that controlling for a general intelligence measure did not remove the relationship between formal thought and biology achievement, no one has investigated whether a general achievement measure, such as SAT, may be at the heart of that relationship. The potential overlap between general achievement and formal thought has important classroom implications for assisting at-risk students. If formal thought and general achievement have a high degree of commonality in relating to course performance, then formal thought maps onto a broader range of general abilities and any potential remedies should consider this. For example, efforts to promote formal thought alone may have limited utility, as other factors such as math ability would still hamper success. However, if students with low formal thought are hindered in the course regardless of scores on the general achievement measure, then formal thought still represents a series of specific traits that are independent of more general measures. In this case, interventions targeted solely toward the development of formal thought would have significant potential to assist at-risk students, whereas those that focus solely on developing math skills (such as algebraic manipulation) would not.

Although the contrast between a general achievement measure and a formal thought measure provides theoretical interest, the primary goal of this study is to produce a generalizable model for identifying at-risk students that will be useful for recommending specific interventions leading to success in general chemistry. It is in this frame that we have discussed the two potential predictors and the implications arising from their comparison. Our goal therefore leads to a series of research questions:

- Which predictor, SAT or a formal thought measure, is better able to identify at-risk students?
- Are the at-risk students identified by each predictor distinct groups, which may lead to more specific interventions geared for each group of students?
- Can a combination of SAT and formal thought measures provide an advantage in identifying at-risk students?
- And, to what extent are all at-risk students identified by this set of predictors?

Methods

Instruments: predictor and outcome variables

SAT

The SAT is a college entrance exam common in the U.S., typically administered in a student's final year of high school (Educational Testing Service, 2006). When the SAT data were obtained, the mathematics portion of this multiple-choice exam covered basic topics in mathematics, including algebra, geometry, data analysis, and probability and statistics, while the verbal portion involved reading comprehension and vocabulary skills. SAT sub-scores were obtained from the university's registrar as they were reported from the Educational Testing Service. SAT sub-scores have been found to have reliability coefficients exceeding 0.9 and a large body of research has demonstrated predictive validity towards college grades, convergent validity with other predictors used in admissions, and construct validity by panel reviews and item analysis (Cohen and Cronbach, 1985).

Test of Logical Thinking (TOLT)

Several measures of formal thought have been developed, validated and utilized in the research literature. What these measures share is an attempt to approximate the original Piagetian interviews. Emulating a Piagetian interview is problematic, especially with large numbers of students, due to the time-intensive nature of the interview procedure. As a result, written exams, in particular, have been constructed to take the place of these interviews. Perhaps the closest approximation to the interview procedure is Shayer and Adey's Science Reasoning Tasks (Shayer and Adey, 1981), in which students are asked to make written predictions before they witness demonstrations and then are asked to explain what they saw in each case. Depending on the task, questions may be free-response or require students to select from a set of responses.

However, for the present study, with class sizes approaching 200 students, we were concerned about the timing for student responses and doubtful that all students would be able to witness a demonstration adequately. As a result, we elected to choose a completely written exam. Among the possibilities are the Inventory of Piagetian Developmental Tasks, IPDT (Bender and Milakofsky, 1982); the Group Assessment of Logical Thinking, GALT (Roadrangka et al, 1983); the Test of Logical Thinking, TOLT (Tobin and Capie, 1981); and the Piagetian Logical Operations Test, PLOT (Staver and Gabel, 1979). Of these choices, the TOLT was selected because of its ease of administration (normally taking 40 minutes of class time); two-tiered question design, which reduces the possibility of students' guessing the correct answer (Treagust, 1988); published validity (Tobin and Capie, 1981), and use in the research literature (Haidar and Abraham, 1991; Yarroch, 1991; Williamson and Abraham, 1995; BouJaoude et al., 2004). Additionally, a Spanish language TOLT has been developed and validated, making the instrument available to a larger audience (Acevedo and Oliva, 1995). Because of the ease of administration and bilingual availability, the TOLT may be seen as a preferential predictor to the SAT from an international perspective. For this case and others in which SAT scores for entering students are not widely available, the TOLT can be given in less than one class period.

The TOLT was developed and validated by Tobin and Capie to measure what they termed formal reasoning ability. In order to do so items previously used by Lawson (Lawson, 1978; Lawson et al., 1979) were selected so that the test comprised two items for each of five modes of formal reasoning: controlling variables, proportional reasoning, probabilistic reasoning, correlational reasoning and combinatorial reasoning. To receive a correct score for each item, students need to select the correct answer from up to 5 choices and select the correct reason for the answer from 5 possible reasons. The only exceptions are the combinatorial reasoning questions, where students are required to list all the correct combinatorial possibilities without any replication. The validation of TOLT was done by relating student scores on the TOLT with student performance via interviews, for students ranging from grade six to college (Tobin and Capie, 1981).

ACS 'Special' Exam

As noted, a large variety of predictor papers rely on student grades as an outcome variable. As a research base, the results of these studies are generalizable only to the extent one can assume that student grades at the research institution match the desired student outcomes of other locales. More importantly, without extensive detail, this assumption becomes impossible to assess. With the desire to produce a generalizable model for identifying at-risk students, we selected an exam produced by the American Chemical Society (ACS), whose Division of Chemical Education features an Examinations Institute, which provides exams to chemistry teachers and administrators in high schools, colleges and universities. This exam is copyrighted and kept secure, so it can be given to candidates year after year, making it easy to make valid comparisons of student scores from different years and institutions.

The Examinations Institute offers more than fifty exams covering general chemistry, organic chemistry, analytical chemistry, physical chemistry, inorganic chemistry, biochemistry, polymer chemistry, and high school chemistry (American Chemical Society, 2006). The first semester general chemistry exams include various lengths of a conventional exam and a special examination (SP97A) meant to combine conceptual knowledge questions with the conventional (algorithmic) type questions. Given the recent push towards conceptual understanding of chemistry in the research literature (Pickering, 1990; Sawrey, 1990; Nakhleh, 1993; Nakhleh et al., 1996) our view is that both conceptual and conventional assessment methods play an important role in the objectives of most general chemistry courses. As a result we selected the ACS special examination as the outcome variable for this study. Though this exam played a large role in determining student grades (it served as the final exam for the course, 25% of student grades), there were other factors that also contributed to student grades. Thus, it would be possible, though unlikely, for a student to complete the course successfully despite a poor performance on the exam.

Participants

The TOLT was administered during the first week of classes in 22 classes of the first semester of general chemistry at a large southeastern public urban research university over the course of three academic years. Students were given 45 minutes to complete the TOLT. Taking the TOLT comprised a small portion of the students' grades, and students were not graded based on their performance on the TOLT. These administration procedures resulted in TOLT scores for 3798 students out of an estimated 4180 students enrolled in the 22 classes. Of the 3798 students, 56.0% of the students were in their first year in college, 62.1% were female and 74.1% reported having at least one full year of high school chemistry.[a] A majority of the students described their major or intended major as pre-med or allied health professions. Finally, when asked about the grade they expected to receive in the course, 97.2% responded with either an A or a B. Only one student reported anticipating failing the course.

At the end of the course, students took the ACS exam as a final exam to measure student academic achievement. Of the 3798 students, ACS exam scores were available for 2871 students (75.6%). Since completing the ACS exam was a course requirement, the likely reason for not obtaining ACS exam scores was students not completing the course. Finally, student SAT scores were obtained from institutional records. Among the 2871 students that took the ACS exam, SAT scores were available for 2284 students. The most likely causes for missing SAT scores were students taking the ACT in place of the SAT or students enrolling in the course after SAT records were pulled. The focus of the analysis was the 2284 students for whom complete data was available. The decision to omit missing data will be revisited in a later section, particularly since the missing data may disproportionately represent at-risk students.

Examining the data

Among the complete data, steps were taken to determine if there were any outliers in the data, so that no single data point would have an unusually large effect on the results of the analysis. Outliers were determined by evaluation of the standardized residuals for a multiple regression model that included both SAT sub-scores and TOLT. An examination for any standardized residuals greater than 3 (Stevens, 1999), revealed nine students found to be inconsistent with the general pattern, and these were omitted from future analysis, resulting in 2275 students.

Prior to examining the trends between variables, descriptive statistics were evaluated and are presented in Table 1.

Table 1. Descriptive statistics for measures used.

	TOLT (0-10)	Math SAT (200-800)	Verbal SAT (200-800)	ACS Exam (0-100)
Mean	6.80	559.14	540.58	52.02
St. Dev.	2.613	83.505	82.648	16.638
Skewness (Std. error = 0.048)	-0.664	-0.048	0.070	0.240
Kurtosis (Std. error = 0.097)	-0.452	-0.166	-0.115	-0.690

The normality tests indicate that the TOLT scores feature a significant negative skew, indicating the scores were more heavily distributed at the higher values. This may be a result of the setting of the study, since the TOLT was designed for grades 6 through college, while the sample consists entirely of college students. While most statistical tests rely on a normality assumption, the tests employed are also very robust to violations of normality (Cohen et al., 2003).

Analysis procedures

As described previously, several research questions guided the nature of this investigation. To investigate each question, inferential and descriptive statistics were used. Inferential statistics established the utility of the predictors by relating the predictors to performance and assisted in interpretation of the descriptive statistics. Where possible, effect sizes were reported as a standardized measure of the differences seen, and operationalized using Cohen's qualitative terms: small, medium and large effects. As Cohen describes them, small effects are where the effect is small relative to the effect of uncontrollable extraneous variables (noise), medium effects are thought to be large enough to be visible to the naked eye and large effects are described as clearly visible (Cohen et al., 2003). Descriptive and inferential statistics were used to relate the ability of the models to identify at-risk students.

The first step in identifying at-risk students is to classify what would constitute an at-risk student. Typically, a grade of 'C' is meant to denote an average performance, and students whose scores fall substantially below an average performance can be considered at-risk. For this study, substantially below average was considered to be scoring below the 30^{th} percentile on the ACS exam. For this sample, students who scored below 43.3% correct on the ACS exam (i.e. more than 0.525 standard deviations below the mean), represent the at-risk group. 802 out of the 2275 students (35.3%) in the sample scored below this cut-off. Since it would have been conceivable to make the decision regarding the at-risk cut-off differently, the effects of choosing different cut-offs are discussed in a later portion of this paper.

Results

Which predictor, SAT or a formal thought measure, is better able to identify at-risk students?

First the extent to which TOLT and SAT sub-scores have a linear relationship with academic performance was determined via correlation analysis. The results from the relevant correlations are presented in Table 2.

Table 2. Comparison of correlation coefficients.

	TOLT	VSAT	MSAT	ACS Exam
TOLT	---			
VSAT	0.492	---		
MSAT	0.654	0.625	---	
ACS Exam	0.510	0.527	0.608	---

all coefficients $p < 0.001$

The presence of significant positive correlation coefficients is indicative of a relationship with academic performance among all the predictors. Correlation coefficients also provide an indication of the strength of relationship between the predictors and the outcome variables. Using Cohen's effect size operation, each of the predictors features a large effect size with the outcome variable, and a large effect size between each predictor. Thus each predictor is believed to be a reasonable construct in explaining ACS exam score, and consideration will need to be given to the possibility of the predictors over-lapping. This is consistent with our goal to determine whether formal thought and general achievement can be thought of as distinctly different in terms of their bearing on college chemistry achievement.

In order to determine the ability of the predictors to identify at-risk students, two linear regression models were used. The first model relates TOLT to students' scores on the ACS exam, and the second relates the SAT sub-scores to the ACS exam. The combination of both SAT sub-scores in one model was chosen to represent the practical option for data available to instructors. The results from the two regression models are shown in Tables 3 and 4 below:

Table 3. TOLT model results.

Coefficient	Slope	Std. Error	t-value	p
Constant	29.936	0.837	35.748	<0.001
TOLT	3.245	0.115	28.249	<0.001

$R^2 = 0.260$ Model F = 798.006

Table 4. SAT model results.

Coefficient	Slope	Std. Error	t-value	p
Constant	-25.196	1.995	-12.630	<0.001
SATV	0.04864	0.004173	11.657	<0.001
SATM	0.09107	0.004130	22.050	<0.001

$R^2 = 0.405$ Model F = 774.514

Using each model, it is possible to depict which students would be identified as at-risk by each set of predictors. For the TOLT model, TOLT scores of 4 or less are predicted to be below the cut-off. By this criterion, the TOLT model identifies 471 students in the sample to be at-risk. Of those 471 students, 332 students had an actual ACS exam score below the cut-off, indicating 70.5% of those predicted were correctly classified (see Table 5). Of the 139 incorrectly classified, 75 scored below average on the ACS exam.

The SAT model predicts students to be below the cut-off via a variety of SAT score combinations, so it is not appropriate to name a single set of SAT cut-offs. However, as rule of thumb, scores below 500 on both the math and verbal portion would qualify as at-risk in this context, although a score below 500 on one portion could potentially be offset by a higher score on the other. Application of the SAT model led to a classification of 451 students as at-risk based on the combination of SAT sub-scores, slightly lower than the number of students the TOLT model predicted. Of the 451 students, 327 were correctly classified, a 72.5% success rate, a rate slightly higher than the TOLT model (see Table 5). Of the 124 incorrectly classified in this group, 69 scored below average.

Table 5. The model predictions: at-risk students.

	Predicted At-risk	Actually At-risk	% correct predictions
TOLT model	471	332	70.5%
SAT model	451	327	72.5%

The similar success rates in identifying at-risk students is curious, given the lower R^2 of the TOLT model compared to the SAT model. As a measure of goodness of fit, it is expected that the higher the R^2 value, the better the model would be at predicting scores. This expectation would hold true if predictions for the entire sample were considered. However, in looking only at the at-risk students, a subset of the sample is being examined. Why the TOLT model is able to identify at-risk students better than expected based on its R^2 value may be understood by considering the following scatter-plots, in which the relationship between ACS exam score and TOLT can be compared with the relationship between ACS exam score and Math SAT. Because of the large number of data points, a 20% random sample of the data is used in these plots.

Figure 1. Low TOLT identifies at-risk.

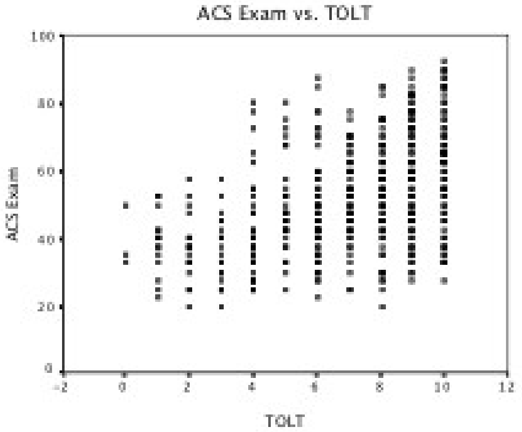
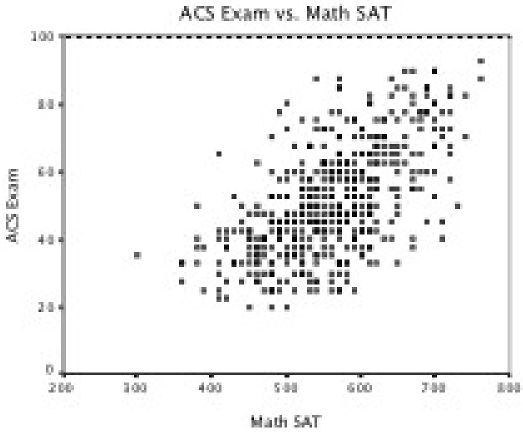


Figure 2. SAT predicts entire range.



The TOLT plot (Figure 1) demonstrates that the variability of ACS exam scores is low for the low TOLT scores, while the ACS exam scores span almost the entire range for the high TOLT scores. This broad distribution at the high end is the likely cause for the lower R^2 for the TOLT model as compared to the SAT model. In the Math SAT plot (Figure 2), a more linear trend is observed: high Math SAT scores correspond to higher ACS exam scores, while lower Math SAT scores correspond to lower ACS exam scores, which would lead to a higher R^2. (Verbal SAT has a similar distribution, as does a combination of the two SAT sub-scores using the weighting found in the regression model; only one of these three possible plots is shown for simplicity). For this reason, SAT would be better suited for identifying successful students than TOLT, while the at-risk students in this sample are comparably identified by each model.

Are the at-risk students identified by each predictor a distinct group, which may lead to more specific interventions geared at each group of students?

Since all three predictors (TOLT, Math SAT, and Verbal SAT) used in the two models feature strong correlations with each other (Table 2), it is tempting to hypothesize that poor performance on any one of the three is indicative of poor performance on all, so that a student predicted to be at-risk by one model would also be predicted to be at-risk by the other. However, this turns out to hold true for just over one-half of the cases predicted to be at-risk: of the 471 students predicted to be at-risk by the TOLT model and the 451 students predicted to be at-risk by the SAT model, only 266 of the students were classified as at-risk by both models. It is useful to consider three exclusive categories: students predicted to be at-risk by both models, at-risk by only the TOLT model, and at-risk by only the SAT model. Table 6 shows the number of students that fall into each category (top row), and the resulting performance on the ACS exam for each category (middle two rows). The bottom row of the table presents the rates of correct prediction in each category for comparison.

Table 6. The overlap between models.

Model Predictions of At-risk Status	Only TOLT At-risk (n=205)	Only SAT At-risk (n=185)	Both models At-risk (n=266)
Correct (scored Below Cut-off)	113	108	219
Incorrect (scored Above Cut-off)	92	77	47
% Correct	55.1%	58.4%	82.3%

From Table 6 it appears that each model, TOLT and SAT, describes a distinct trait that hinders success in chemistry. There is a distinct group of 113 students that performed poorly on the TOLT and on the ACS final exam while performing satisfactorily on the SAT measure. A similar situation occurs for 108 students who performed poorly on the SAT and on the ACS final exam while performing reasonably well on the TOLT. These two cases demonstrate that the two models identify different groups of students as being at-risk, even though 219 students were correctly predicted by both models to be at-risk. It should also be noted that neither of the models identifies all students who are at-risk: out of the 1619 students not predicted to be at-risk by either model, only 1257 (77.6%) in fact performed above the cut-off. This will always be the case: the models attempt to identify factors that are necessary for success in general chemistry, but necessary does not mean sufficient.

Statistical comparisons between percent correct predictions employed an arcsine transformation to stabilize variances (Cohen, 1988). The highest percent correct is for those who would be classified as at-risk by both the TOLT model and by the SAT model, demonstrating that a combination of low scores on both measures leads to a greater chance of students performing poorly on the ACS final exam. The differences in correct prediction rate between this ‘both’ category and each of the two ‘only’ categories were significant with a medium effect size. No evidence supporting a significant difference in percent correct between the only TOLT category and the only SAT category was found, indicating that neither model isolates a distinct group of at-risk students better than the other.

Can a combination of SAT and formal thought measures provide an advantage in identifying at-risk students?

The previous discussion has shown that, if both the TOLT model and the SAT model predict a student to be at-risk, that is very likely to be the case! However, this post-hoc combination of the predictions of two different models may be too conservative, identifying only a relatively small number of at-risk students. It may be possible to construct a single model using both sets of predictors that will retain a high success rate and identify a larger number of at-risk students. To investigate this possibility, a model (shown in Table 7) was constructed to use both SAT sub-scores and TOLT scores:

Table 7. Combined TOLT and SAT model.

Coefficient	Slope	Std. Error	t-value	p
Constant	-19.477	2.106	-9.250	<0.001
SATV	0.04410	0.004163	10.594	<0.001
SATM	0.07253	0.004738	15.310	<0.001
TOLT	1.0440	0.13574	7.691	<0.001

$R^2 = 0.420$ Model F = 549.274

Each predictor entered the model significantly, indicating that, even when controlling for the variability that comes from the other predictors, both SAT sub-scores and TOLT still feature a significant relation with the ACS exam, which is consistent with the interpretation of the data in Table 6 that TOLT and SAT map onto performance in distinct ways. (Note that the incorporation of interaction terms TOLT*MSAT and TOLT*VSAT into the model adds only 0.009 to R^2; therefore, these terms were not retained in the model.) Similar to the SAT model, this new model that combines both SAT sub-scores and TOLT scores has many combinations of predictor scores that would result in an at-risk prediction. The combined model predicts 489 students to be at-risk, higher than the 266 predicted by the overlap of the two individual models. Of those 489 students, 354 scored below the ACS final exam cut-off and thus were correctly classified. This leads to a success rate of 72.4%, which is only slightly higher than the 70.5% seen with the TOLT model, and essentially equivalent to the 72.5% rate of the SAT model. Further, of the 354 students correctly classified by this combined model, 351 had been identified by one of the two previous models. Thus the combination of both predictors in a single regression model fails to provide an improved way to identify at-risk students, since only three additional students were correctly found by combining the two sets of predictors. Of the 135 misclassified, 71 scored below average on the exam.

To what extent are all at-risk students identified by this set of predictors?

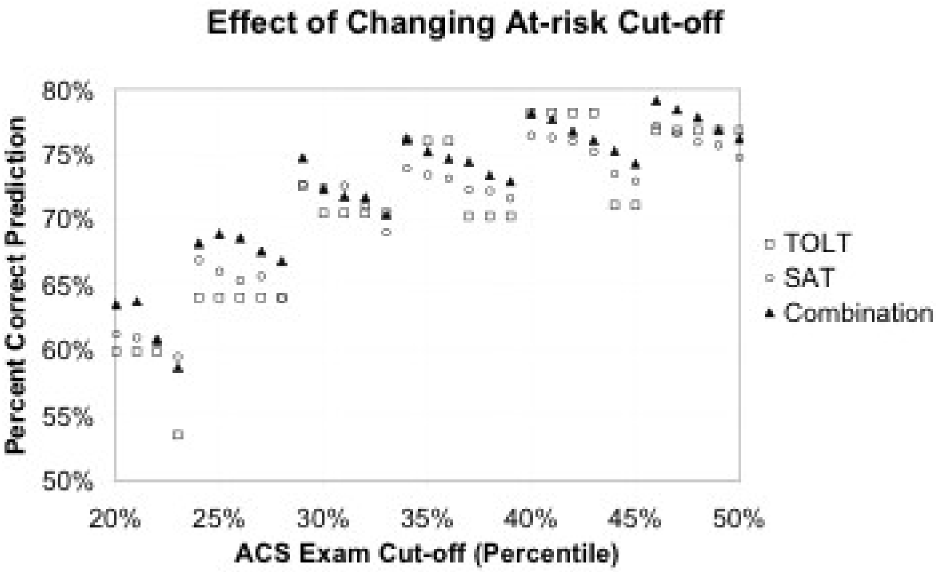
Of the 2275 students, 802 students finished the course below the ACS cut-off. Of these 802 students, 443 students (55.2%) were identifiable based on scores from either TOLT, SAT or a combination of the two. Thus a sizable portion of the students that performed poorly on the ACS exam was not identifiable by these models. We believe this finding may be representative of a need to include non-cognitive predictors, such as affective measures like motivation or confidence, if the goal is to predict all at-risk students. However, it is important to recognize that many non-cognitive factors may feature a strong correlation with general achievement (House, 1995). Rather than attempting to include non-cognitive factors in this study, it would be more appropriate to focus an additional study on the degree to which affective factors are distinct from general achievement, in the same manner as we have set out the comparison between formal thought and general achievement.

At-risk cut-off

It is recognized that the decision to employ the cut-off at the bottom 30% of the sample is somewhat arbitrary, as other values such as the bottom 25% or bottom 33% could reasonably suffice. To address these concerns and to understand the impact of this decision on the conclusions reached, a SAS program was developed to calculate the percent correct predictions for each model as the cut-off point is changed. The results have been plotted in Figure 3.

First, note from Figure 3 that the models switch places depending on the cut-off decision, but all of them remain relatively close together, so that no model offers a distinct advantage over the others in terms of accuracy in identifying at-risk students. Also note the general upward trend of percent correct predictions as the cut-off decision increases. This can be attributed to chance guessing. For example, if the cut-off is placed at 20%, randomly selecting students would get 20% correct prediction in identifying at-risk students. However, if the cut-off was 40%, there would be a 40% chance of identifying at-risk students by random selection. In general, each model stays approximately 35 - 45% above the random selection method of identifying students. Now that we can describe the role of the models in the identification of at-risk students for whom SAT scores were available, we will now turn our attention to another important aspect of this study, the consideration of students for whom SAT scores were not available.

Figure 3. Correct predictions for each model versus the cut-off point.



Those without SAT scores

As mentioned, cases for which SAT scores were unavailable were omitted so that the previous comparisons between the SAT model and the TOLT model could be undertaken for the same group of students. This omission presents some interesting implications for the study. A chief concern with missing data is the presence of a trend in those students who have missing data, because the presence of any such trend represents a limitation in the generalizability. By omitting students without available SAT scores, it is necessary to check if the group omitted differs from the group studied. If so, then the applicability of the analysis to those omitted may be questionable. Table 8 presents the results from this comparison

Table 8. Comparison of those with SAT scores to those without.

	Avg. score for those with SAT (n, st dev)	Avg. score for those without SAT (n, st dev)	t-test	p-value	d-value
TOLT	6.65 (2957, 2.656)	6.24 (841, 2.645)	3.934	0.000	0.155
ACS Exam	52.11 (2284, 16.724)	49.92 (587, 16.178)	2.839	0.005	0.133

The students without SAT scores scored significantly lower on both the TOLT measure and the ACS exam measure than students with SAT scores. The d-value is the effect size for comparing two means, with both values representing a small effect. Because of the differences between students with SAT scores and those without, the students without SAT scores likely represent a non-random population. For this reason we will examine these students separately in terms of the conclusions presented so far.

There were 841 students in the original sample without SAT scores, and 587 of those took the ACS exam. While no claim can be made regarding what the SAT model would have predicted for these students, the role of TOLT in identifying at-risk students can still be investigated. To do this, a new regression model (Table 9) was fitted for just these 587 students.

Table 9. TOLT model for students without SAT scores only.

Coefficient	Slope	Std. Error	t-value	p
Constant	31.910	1.599	19.962	<0.001
TOLT	2.799	0.230	12.149	<0.001

$R^2 = 0.201$ Model F = 147.587

As does the previous TOLT model, this model indicates a positive linear relationship between TOLT and ACS exam scores. An examination of the standard error associated with the TOLT coefficient (0.230) and the intercept (1.599) in this model indicates that it cannot be considered different from the original model. It appears the conclusions reached regarding the previous TOLT model also apply to the students without SAT scores. Of the 587 students, 150 scored at or below a 4 on the TOLT and would therefore be characterized as at-risk. Of these 150 students predicted to be at-risk, 101 scored below the ACS final exam cut-off, giving a 67.3% success rate in classification. Of the 49 incorrectly classified, 20 scored below the average score. In short, the inclusion of students for whom SAT scores were not available reveals nothing inconsistent with the previous findings regarding the utility of TOLT. The conclusion that TOLT as a formal thought measure identifies a barrier to the success of chemistry students holds true for those in our sample without SAT scores. The distinct advantage of the TOLT model over the SAT model here is that, even though SAT scores were unavailable, the ease of administration of the TOLT made it possible to identify correctly an additional 101 students as being at-risk.

Those who did not finish the course

As mentioned earlier, those who did not finish the course represent a significant portion of the at-risk student population. However, while leaving the course may be a function of academic performance during the course, there are also a variety of other reasons for such a departure, ranging from personal health to financial trouble. For this reason, we are hesitant to classify all students who did not finish the course as at-risk students. However, given the nature of the conclusions reached regarding the ability of TOLT to predict performance, it will only be necessary to examine those whose TOLT scores fell at or below 4, to determine if those students were in fact performing poorly when they left the course. This will be approximated by reviewing students' scores on four instructor-generated, multiple-choice in-course tests, in comparison to the class performance on the same test.

Of the 3798 students in this study, 927 students (24.4%) did not take the ACS exam. Of those 927 students, 263 students (28.4%) scored at or below a 4 on the TOLT, which was the criterion previously used for at-risk classification. Forty-two of the 263 students did not take any of the tests, so their decision to drop the course came relatively early, and unfortunately, little else can be said of them. However, of the remaining 221 students, 194 students scored in the bottom 30% within their class on every test they took. Of the remaining 27 students, 22 scored above this mark only once. While it is not possible to extrapolate an exact reason for leaving a course from the data available, and indeed the decision is likely attributable to a number of factors, the data do indicate that low academic performance probably played a role in the decision and that low performance on the TOLT would have served as a warning sign for this population.

Discussion

What we have identified with this study is threefold:

1) Formal thought (as measured by TOLT) and general achievement (as measured by SAT) represent separate and distinct factors, each of which can be used to identify at-risk students.
2) Neither the formal thought measure (TOLT) nor the general achievement measure (SAT) is clearly superior in terms of percent correct identification of at-risk students.
3) The ease of administration of the TOLT makes it possible to use this measure to identify additional at-risk students for whom SAT scores are not available

The fact that both general achievement and formal thought represent distinct factors (see discussion of Table 6) in this study is important. It seems there are at least three groups of at-risk students: those who do not have an appropriate knowledge of mathematics and language for success in chemistry, those who do not have the requisite reasoning skills for success in chemistry, and those who lack both. In other words, even if a student performs reasonably well on the SAT, low formal thought ability can still hinder his or her success in chemistry, and the reverse is also true. Since this result shows that mathematics achievement and reasoning ability represent different barriers to success, effective remediation aimed at these two different groups of students will incorporate a review of relevant mathematical and verbal skills as well as the opportunity to work on developing formal thought ability. A remedial course focused solely at reviewing fundamental mathematical rules in the abstract (e.g. how to isolate variables in an equation, how to manipulate logarithms) is definitely too narrow. Connecting mathematical manipulations with concrete observables in chemistry could provide some assistance to both groups; however, the chemistry content review in a standard remedial course has not typically led to success for a large proportion of students (Bentley and Gellene, 2005). We suggest that attention should be paid to the sequencing of concepts in the chemistry content review, presenting the most abstract only after a concrete foundation has been established (Tsaparlis, 1997; Sanger et al., 2001). Chemistry educators have also achieved some success with improving cognitive skills such as formal reasoning ability in the context of chemistry courses via the integration of carefully sequenced problem-solving activities, incorporating algorithms of increasing complexity, with conceptual instruction sufficient to explain the 'how' and 'why' of the calculations at each level (Tsaparlis, 2005).

From a pedagogical perspective, chemistry review lectures on concrete concepts but with few graphics, animations, or demonstrations require students to create their own mental models of these concepts, a skill associated with formal rather than concrete thinking. Taking advantage of the wide array of animations available for illustrating major concepts in chemistry is perhaps the simplest way to scaffold learners with low formal thought ability in the large lecture setting (Williamson and Abraham, 1995; Tasker et al., 1996; Sanger and Greenbowe, 2000; Wu et al., 2001; Stieff, 2005). Another alternative is the incorporation of computer-assisted learning activities to supplement lectures. Simulations with a focus on manipulating variables and repetition have shown promise with science learners of low formal reasoning ability (Huppert, 2002). In general, researchers have recommended the use of active learning practices to avoid over-dependence on lectures (Chandran et al., 1987; Shayer, 2003). Tien et al. (2002) and Lewis and Lewis (2005) provide examples of effective active learning reforms that de-emphasize lectures without moving completely from the lecture format.

Finally, in all cases, both quantitative and qualitative investigations of the effects of the reform on different groups of students are needed to provide insight into whether and how these reforms assist those who need to develop formal thinking skills as well as chemistry knowledge. How should these investigations be conducted? Indeed, considering formal thought as a factor distinct from general achievement in college chemistry has implications for research as well as for teaching. The relationship between formal thought as measured by TOLT and chemistry performance (displayed graphically in Figure 1) is important: TOLT is a better predictor of at-risk students than of successful students. Therefore, studies that investigate a relationship between TOLT and academic performance through the use of linear regression or correlation (BouJaoude et al., 2004) may be underestimating the importance of formal thought. A large amount of variation in academic performance for students with high TOLT scores, while congruent with cognitive development theory, would lead to a reduced proportion of variance explained by TOLT as compared with other predictors of performance. In other words, researchers may be misled into thinking formal thought is not relevant for a given situation, when, in fact, the association of low formal thought ability with poor performance is masked by the large variability in performance for those at the higher end of the TOLT. A suggestion for researchers who are considering such models is to dichotomize TOLT scores, creating a low TOLT score and high TOLT score classification, an approach that is better aligned with the theory of developmental stages. Another option would be to consider students with low TOLT scores as a unique subset of students. This latter option should be of particular use when evaluating whether pedagogical reforms are able to help different groups of at-risk students.

Even though this study focuses on college-level general chemistry, it is also worthwhile to consider broader teaching implications. Longitudinal work from Novak has shown that complex science instruction among elementary age students can show improved understanding at the high school level on similar concepts, far removed from the intervention (Novak, 2005). Therefore initiatives to improve formal thought ability could also be instituted earlier in the educational stream, with the strong possibility for improving the trends witnessed here at the college level. One such initiative, Shayer and Adey's Cognitive Acceleration program (Adey and Shayer, 1994), has shown promising results in promoting cognitive development among middle school students.

Conclusions

In this study, formal thought has been found to have a unique relationship to chemistry achievement apart from SAT sub-scores, even though the two constructs share a medium-sized correlation. Low formal thought ability impedes success in chemistry as much as low SAT sub-scores, and formal thought has been shown to represent a necessary factor for success in college-level general chemistry for a distinct group of students. Recommendations for remediation and for future research were discussed in light of these findings. It is important to note that, while both measures used in the study had reasonable success at identifying students at-risk of performing poorly in college-level general chemistry, there was an additional group of students whose poor performance was not predicted by either measure. Therefore, factors that are unaddressed in this paper are also likely to play a role in success in chemistry. Research into affective aspects of chemistry learning with specific emphasis on at-risk students would complement the cognitive approach taken in this paper. In particular, identifying those affective components that prevent students from achieving success despite high cognitive abilities may help identify other distinct groups of at-risk students and lead to the development of targeted remedies for these groups.

Notes

a. It was found that those students completing one year of high school chemistry or more scored significantly higher, with a consistent small effect size (d=0.2), than those who did not on all four measures (Table 10):

Table 10. Comparison of high school chemistry takers with non-takers.

	Average Score (standard deviation, n)		t-value	p-value	d-value
	< 1 full year of high school	≥ 1 full year of high school chemistry			
TOLT	6.19 (2.73, 936)	6.73 (2.61, 2678)	5.437	0.000	0.202
Math SAT	533.5 (84.8, 691)	558.2 (83.1, 2184)	6.770	0.000	0.294
Verbal SAT	523.8 (83.1, 691)	539.6 (81.7, 2184)	4.416	0.001	0.197
ACS Exam	48.0 (15.1, 660)	53.0 (16.9, 2107)	6.737	0.000	0.312

The high school chemistry taking population is therefore likely not representative of the college chemistry taking population. This also supports Chandran et al.'s (1987) postulate that students taking high school chemistry represent a population of students that are likely to succeed in chemistry.

References

Abraham M.R. and Williamson V.M., (1994), A cross-age study of the understanding of five chemistry concepts, *Journal of Research in Science Teaching,* **31**, 147-165.

Acevedo J.A. and Oliva J.M., (1995), Validacion y aplicaciones de un test de razonamiento logico, *Revista de Psicologia General y Aplicada,* **48,** 339-352.

Adey P.S. and Shayer M., (1990), Accelerating the development of formal thinking in middle and high school students, *Journal of Research and Development in Education,* **27**, 267-285.

Adey P.S. and Shayer M., (1994), *Really raising standards: cognitive intervention and academic achievement* (First ed.), New York: Routledge.

Bender D.S. and Milakofsky L., (1982), College chemistry and Piaget: the relationship of aptitude and achievement measures, *Journal of Research in Science Teaching,* **19**, 205-216.

Bentley A.B. and Gellene G.I., (2005), A six-year study of the effects of a remedial course in the chemistry curriculum, *Journal of Chemical Education,* **82**, 125-130.

BouJaoude S., Salloum S. and Abd-El-Khalick F., (2004), Relationships between selective cognitive variables and students' ability to solve chemistry problems, *International Journal of Science Education,* **26**, 63-84.

Bunce D.M. and Hutchinson K.D., (1993), The use of the GALT (Group Assessment of Logical Thinking) as a predictor of academic success in college chemistry, *Journal of Chemical Education,* **70**, 183-187.

Carmichael J.W.J., Bauer J.S., Sevenair J.P., Hunter J.T. and Gambrell R.L., (1986), Predictors of first-year chemistry grades for Black Americans, *Journal of Chemical Education,* **63**, 333-336.

Chandran S., Treagust D.F. and Tobin K.G., (1987), The role of cognitive factors in chemistry achievement, *Journal of Research in Science Teaching,* **24**, 145-160.

Cohen J., (1988), *Statistical power analysis for the behavioural sciences* (2nd ed.), Hillsdale, Lawrence Erlbaum Associates.

Cohen J., Cohen P., West S.G. and Aiken L.S., (2003), *Applied multiple regression / correlation analysis for the behavioral sciences* (3rd ed.), Mahwah, NJ, Lawrence Erlbaum Associates, Inc.

Cohen S.J. and Cronbach L.J., (1985), College Board Scholastic Aptitude Test and Test of Standard Written English, In Boros Institute of Mental Measurement (Ed.), *The Mental Measurements Yearbook* (9th ed.), New Brunswick, Rutgers University Press.

Craney C.L. and Armstrong R.W., (1985), Predictors of grades in general chemistry for allied health students, *Journal of Chemical Education,* **62**, 127-129.

Demerouti M., Kousathana M. and Tsaparlis G., (2004), Acid-base equilibria, Part II. Effect of developmental level and disembedding ability on students' conceptual understanding and problem-solving ability, *The Chemical Educator,* **9**, 132-137.

Educational Testing Service, (2006), About the SAT, Accessed December 2006; http://www.collegeboard.com/student/testing/sat/about.html

Examinations Institute of the American Chemical Society, Division of Chemical Education (1997), *First term general chemistry (special examination)*, Clemson, SC, Clemson University.

Haidar A.H. and Abraham M.R., (1991), A comparison of applied and theoretical knowledge of concepts based on the particulate nature of matter, *Journal of Research in Science Teaching,* **28**, 919-938.

Herron J.D., (1975), Piaget for chemists, *Journal of Chemical Education,* **52**, 146-150.

House J.D., (1995), Noncognitive predictors of achievement in introductory college chemistry, *Research in Higher Education,* **36**, 473-490.

Huppert J., (2002), Computer simulations in the high school: students' cognitive stages, science process skills, and academic achievement in microbiology, *International Journal of Science Education,* **24**, 803-821.

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

Jones K.B. and Gellene G.I., (2005), Understanding attrition in an introductory chemistry sequence following successful completion of a remedial course, *Journal of Chemical Education,* **82**, 1241-1245.

Lawson A.E., (1978), The development and validation of a classroom test of formal reasoning, *Journal of Research in Science Teaching,* **15**, 11-24.

Lawson A.E., (1979), The developmental learning paradigm, *Journal of Research in Science Teaching,* **16**, 501-515.

Lawson A.E., (1982), Formal reasoning, achievement, and intelligence: an issue of importance, *Science Education,* **66**, 77-83.

Lawson A.E., (1983), Predicting science achievement: the role of developmental level, disembedding ability, mental capacity, prior knowledge, and beliefs, *Journal of Research in Science Teaching,* **20**, 117-129.

Lawson A.E., (1985), A review of research on formal reasoning and science teaching, *Journal of Research in Science Teaching,* **22**, 569-617.

Lawson A.E. and Nordland F.H., (1976), The factor structure of some Piagetian tasks, *Journal of Research in Science Teaching,* **13**, 461-466.

Lawson A.E. and Renner J.W., (1975), Relationships of science subject matter and developmental levels of learners, *Journal of Research in Science Teaching,* **12**, 347-358.

Lawson A.E., Adi H. and Karplus R., (1979), Development of correlational reasoning in secondary schools: do biology courses make a difference?, *The American Biology Teacher,* **41**, 420-425.

Lawson A.E., Karplus R. and Adi H., (1978), The acquisition of propositional logic and formal operational schemata during the secondary school years, *Journal of Research in Science Teaching,* **15**, 465-478.

Legg M.J., Legg J.C. and Greenbowe T.J., (2001), Analysis of success in general chemistry based on diagnostic testing using logistic regression, *Journal of Chemical Education,* **78**, 1117-1121.

Lewis S.E. and Lewis J.E., (2005), Departing from lectures: an evaluation of a peer-led guided inquiry alternative, *Journal of Chemical Education,* **82**, 135-139.

McFate C. and Olmsted III J.; (1999), Assessing student preparation through placement tests, *Journal of Chemical Education,* **76**, 562-565.

Nakhleh M.B., (1993), Are our students conceptual thinkers or algorithmic problem solvers?, *Journal of Chemical Education,* **70**, 52-55.

Nakhleh M.B., Lowrey K.A. and Mitchell R.C., (1996), Narrowing the gap between concepts and algorithms in freshman chemistry, *Journal of Chemical Education,* **73**, 758-762.

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., (1996), Reasoning strategies of students in solving chemistry problems as a function of developmental level, functional M-capacity, and disembedding ability, *Journal of Chemical Education,* **64**, 502-505.

Nordstrom B.H., (1990), *Predicting performance in freshman chemistry,* Paper presented at the American Chemical Society National Meeting, Boston, Massachusetts.

Novak J.D., (2005), Results and implications of a 12-year longitudinal study of science concept learning, *Research in Science Education,* **35**, 23-40.

Oliva J.M., (2003), The structural coherence of students' conceptions in mechanics and conceptual change, *International Journal of Science Education, 25*(5), 539-561.

Ozsogomonyan A. and Loftus D., (1979), Predictors of general chemistry grades, *Journal of Chemical Education,* **56,** 173-175.

Pederson L.G., (1975), The correlation of partial and total scores of the Scholastic Aptitude Test of the College Entrance Examination Board with grades in freshman chemistry, *Educational and Psychological Measurement,* **35**, 509-511.

Pickering M., (1975), Helping the high risk freshman chemist, *Journal of Chemical Education,* **52**, 512-514.

Pickering M., (1990), Further studies on concept learning versus problem solving, *Journal of Chemical Education,* **67**, 254-255.

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 of Research in Science Teaching, Dallas.

Russell A.A., (1994), A rationally designed general chemistry diagnostic test, *Journal of Chemical Education,* **71**, 314-317.

Sanger M.J., Brinks E.L., Phelps A.J., Pak M.S. and Lyovkin A.N., (2001), A comparison of secondary chemistry courses and chemistry teacher preparation programs in Iowa and Saint Petersburg, Russia, *Journal of Chemical Education,* **78**, 1275-1280.

Sanger M.J. and Greenbowe T.J., (2000), Addressing student misconceptions concerning electron flow in aqueous solutions with instruction including computer animations and conceptual change strategies, *International Journal of Science Education,* **22**, 521-537.

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

Shayer M., (2003), Not just Piaget; not just Vygotsky, and certainly not Vygotsky as alternative to Piaget, *Learning and Instruction,* **13**, 465-485.

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

Shayer M. and Adey P.S., (1992a), Accelerating the development of formal thinking in middle and high school students II: post-project effects on science achievement, *Journal of Research in Science Teaching,* **29**, 81-92.

Shayer M. and Adey P.S., (1992b), Accelerating the development of formal thinking in middle and high school students III: testing the permanency of effects, *Journal of Research in Science Teaching,* **29**, 1101-1115.

Shayer M. and Adey P.S., (1993), Accelerating the development of formal thinking in middle and high school students IV: three years after a two-year intervention, *Journal of Research in Science Teaching,* **30**, 351-366.

Spencer H.E., (1996), Mathematical SAT test scores and college chemistry grades, *Journal of Chemical Education,* **73**, 1150-1153.

Staver J.R. and Gabel D.L., (1979), The development and construct validation of a group administered test of formal thought, *Journal of Research in Science Teaching,* **16**, 535-544.

Stevens J.P., (1999), *Intermediate statistics: a modern approach* (2nd ed.), Mahwah, New Jersey, Lawrence Erlbaum Associates.

Stieff M., (2005), Connected chemistry: a novel modeling environment for the chemistry classroom, *Journal of Chemical Education,* **82**, 489-493.

Tai R.H., Sadler P.M. and Loehr J.F., (2005), Factors influencing success in introductory college chemistry, *Journal of Research in Science Teaching,* **42**, 987-1012.

Tasker R.F., Chia W., Bucat R.B. and Sleet R., (1996), The VisChem Project: visualising chemistry with multimedia, *Chemistry in Australia,* **63**, 395-397.

Tien L.T., Roth V. and Kampmeier J.A., (2002), Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course, *Journal of Research in Science Teaching,* **39**, 606-632.

Tobin K.G. and Capie W., (1981), The development and validation of a group test of logical thinking, *Educational and Psychological Measurement,* **41**, 413-423.

Treagust D.F., (1988), Development and use of diagnostic-tests to evaluate student misconceptions in science, *International Journal of Science Education,* **10**, 159-169.

Tsaparlis G., (1997), Atomic and molecular structure in chemical education: a critical analysis from various perspectives of science education*, Journal of Chemical Education,* **74**, 922-925.

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.

Tsaparlis G. and Angelopoulos V., (2000), A model of problem-solving: its operation, validity, and usefulness in the case of organic-synthesis problems*, Science Education,* **84**, 131-153.

Tsaparlis G., Kousathana M. and Niaz M., (1998), Molecular-equilibrium problems: manipulation of logical structure and of M-demand and their effect on student performance*, Science Education,* **82**, 437-454.

Wagner E.P., Sasser H. and DiBiase W.J., (2002), Predicting students at risk in general chemistry using pre-semester assessments and demographic information, *Journal of Chemical Education,* **79**, 749-755.

Williamson V.M. and Abraham M.R., (1995), The effects of computer animation on the particulate mental models of college chemistry students, *Journal of Research in Science Teaching,* **32**, 521-534.

Wu H.K., Krajcik J.S. and SolowayE., (2001), Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom, *Journal of Research in Science Teaching,* **38**, 821-842.

Yager R.E., Snider B. and Krajcik J.S., (1988), Relative success in college chemistry for students who experienced a high-school course in chemistry and those who did not, *Journal of Research in Science Teaching,* **25**, 387-396.

Yarroch W.L., (1991), The implication of content versus item validity on science tests, *Journal of Research in Science Teaching,* **28**, 619-629.

Primary teachers' views and descriptions regarding some science activities

George Papageorgiou*, Efthalia Kogianni and Nicolaos Makris

Democritus University of Thrace, Department of Primary Education, Alexandroupolis 68100, Greece,
e-mail: gpapageo@eled.duth.gr

Received 5 September 2006, accepted 8 December 2006

Abstract: The views and the descriptions of 228 primary teachers on the use and the carrying out of some chemistry-related practical activities were studied. The research took place in the context of an in-service training course. As research tool, a questionnaire was used, where tasks dealing with three hands-on activities on chemistry topics, were included. According to the results, teachers had generally both the necessary knowledge to carry out such activities and the ability to describe procedures with the use of appropriate equipment and materials, to a satisfactory degree. However, their responses ranged from a more correct and complete level to a more generally descriptive and incomplete level. Results also suggest that, the gender and the use of science activities in school had an impact on their descriptions. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 52-60]

Key words: Primary teachers, descriptions, primary science, (hands-on) science activities.

Introduction

Although teaching science is one of the main objectives for a primary teacher, many avoid it – especially chemistry topics. They give higher priority to other subjects (e.g., reading), they often do not have the requisite subject matter knowledge, and they don't feel very comfortable with science (Smith and Neale, 1989; Appleton and Kindt, 1999; Schibeci and Hickey, 2000; Appleton, 2003). This could be also related to the female teachers' predominance in primary education (Johnston et al., 1999), since the female identity is more connected to theoretical sciences (such as humanities, social sciences or art), than to practical sciences (Kelly, 1987; Arnot et al., 1999).

Many of the above preconditions for a successful implementation of hands-on science activities in primary education are lacking, and such activities are not very common in primary schools. Since science activities are very important for primary education, they should be included in the primary teachers' practice, as they can help children to overcome their difficulties in understanding science concepts to a certain degree (Gabel, 1999; Lavonen et al., 2004; Papageorgiou and Tsiropoulou, 2004). So, the question addressed here is this: What are the views and descriptions of the teachers themselves regarding such hands-on science activities in school?

As Appleton (2002) suggests, among the requirements for a hands-on science activity is that the background science content for the activity should be already known to the teacher. In particular, such science activities on chemistry and physics topics need further training support and thorough knowledge and understanding of the science subject matter; accordingly, the teachers requested more practice in activity planning (Asunta, 1997). Additionally, teachers who intend to carry out hands-on activities in school should both be knowledgeable about the relevant topics, and know how to use particular materials and equipment (Kepler, 1998).

In this context, we decided to study the views and the descriptions of primary teachers regarding both their ability (in the sense of the requisite knowledge) to carry out some hands-on science activities on chemistry topics and their use of the appropriate equipment and materials.

Research methods

In this study we attempt to answer the following questions:

- What are the teachers' descriptions regarding both their ability to carry out some hands-on science activities on chemistry topics, and the practical handling of appropriate equipment and materials?
- What are the teachers' views on the use of such science activities in primary schools?
- To what extent teachers' views and descriptions are related to their gender?

In order to answer these questions, 228 teachers of primary education (97 male and 131 female) who have been working in Greek primary schools in the region of East Macedonia and Thrace participated in this study. The time teachers had been working in schools, ranged from 5 to 20 years. The study took place during an in-service teachers' training program. The teachers of the sample who participated in this program were divided into nine classes.

Data were collected using an anonymous questionnaire, especially constructed for the purposes of the present study, which asked teachers to comment on their use of hands-on science activities in school. The three questions of the questionnaire are described in Figure 1., In addition, the teachers were asked to work on three tasks concerning hands-on science activities; the tasks are given in Figure 2. The time for the completion of the questionnaire and the tasks was one hour. The data collected were qualitatively analysed by two of the authors independently, who classified the data according to a category scheme especially devised for the present study. Additionally, in order to compare the teachers' ability to carry out the activities in the three tasks and to determine the effect of the 'gender' on the teachers' descriptions, we coded the data quantitatively. The coding was done by the two authors, who had also analysed the teachers' responses qualitatively.

Figure 1. Description of the questionnaire.

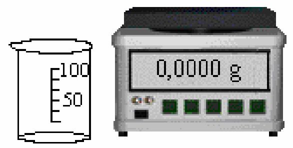
Question 1: In this multiple choice question, teachers were asked to check how often they used hands-on science activities in school (often/sometimes/never)

Question 2: Teachers were asked to give some possible reasons for not using hands-on science activities in school as often as they would have liked, if this was the case. There were some possible reasons suggested (Hands-on science activities are too difficult / There are hazards using such activities / I haven't been trained for that), but they could also give their own reasons.

Question 3: Teachers were asked to give some possible reasons, in case they did not want to use hands-on science activities in school at all. There were also some possible reasons suggested (I don't think that such activities help the learning process / There are hazards using such activities), but they could also provide their own reasons.

Figure 2. The three tasks.

Task 1: It is assumed that a balance and a beaker with volume graduations are available (the relevant equipment is shown below). Teachers are asked to describe the procedure that they would follow in order to determine the density of an unknown liquid. The following reminder is also given: the density of a liquid can be calculated by dividing the mass by the volume.



Task 2: It is assumed that there is a mixture of water, acetone and table salt in a vessel and that any necessary equipment for the separation of the mixture into its components is available. Teachers are asked to describe the procedure they would follow in order to separate the components of the mixture, describing the necessary equipment. Some relevant data are also given:

	Water	Acetone	Table salt
Melting point (°C)	0	-95	801
Boiling point (°C)	100	56	1413

Task 3: Teachers are asked to describe the procedure they would follow in order to determine the solubility of table salt in water, describing any equipment necessary for this purpose. The following reminder is also given: the solubility of a substance in water is the maximum amount of this substance that can be dissolved in a given quantity of water (usually 100 g) at a specific temperature.

Results

How often do teachers use hands-on science activities in school?

Table 1 shows primary teachers' views on how often they used hands-on science activities in school. The teachers seemed to participate in such activities to a significant degree, although male and female teachers did it to a different extent. Male participants estimated that they used hands-on science activities more often than their female colleagues, supporting the view that male teachers are more practical than the female ones (Arnot et al., 1999). The difference between the two genders is found to be significant [$\chi^2(3)=15.766$, $p=.001$].

Table 1. Teachers' answers about their use of hands-on science activities in school.

Category	Teachers' answers	Number of teachers, N, and Percentage (%)		
		Total	Male*	Female*
1	Often	93 (40.8)	53(54.6)	40(30.5)
2	Sometimes	125 (54.8)	43(44.3)	82(62.6)
3	Never	8 (3.5)	1(1.1)	7(5.4)
4	No answer	2 (0.9)	0(0.0)	2(1.5)

*Percentages refer to the total number of male teachers or female teachers, respectively.

Teachers' descriptions regarding hands-on science activities

Task 1

For task 1, teachers were asked to determine the density of an unknown liquid. Their answers were categorized according to their correctness and completeness (see Table 2). In category 1, 44.3% of teachers described in detail a correct and complete procedure. Another 27.2% of the teachers (category 2) gave a correct, but not totally complete answer, as they responded similarly to those in category 1, with the exception of the determination of volume: They did not specify how the volume is measured (i.e., by the volume graduations on the beaker). Teachers who fell into category 5 (sub-categories 5a and 5b) of Table 2 could not determine the density of the liquid because either they used the total weight of beaker and liquid in their calculations without subtracting the tare (sub-category 5a) or they did not proceed to the division of *mass* by *volume* (sub-category 5b).

Table 2. Teachers' categories according to their answers in tasks 1, 2, 3.

Cat.	Characteristics of the teachers' descriptions		Task 1	Task 2	Task 3	Score
	Correctness	Completeness	N (%)	N (%)	N (%)	
1	Correct procedure	Complete description of the procedure	101 (44.3)	**1a.** 17 (7.6) **1b.** 40 (17.5)	49 (21.5)	4
2	Correct procedure	Almost complete description of the procedure (*Clarifications in some points are missing or they are implied*)	62 (27.2)	68 (29.8)	74 (32.5)	3
3	(Mainly) correct part of the procedure	Partial description of the procedure (*A part of the procedure is missing*)	-	52 (22.8)	-	2
4	Correct outline of the procedure	Description of an outline of the procedure	-	25 (11.0)	66 (28.9)	1
5	Incorrect procedure/ No answer	Complete or incomplete description/ No answer	**5a.** 39 (17.1) **5b.** 17 (7.5)	4 (1.7)	7 (3.1)	0
6	Miscellaneous answers		13 (5.7)	22 (9.6)	32 (14)	0

Note: Sub-categories 1a and 1b, and 5a and 5b are explained in the text.

Task 2

Besides the teachers' ability to carry out the activity, the use of the appropriate equipment and materials was also an issue in task 2. The categorization (Table 2 for task 2) was not simple as there was too great a variety in teachers' working and their descriptions. Generally, in all tasks, teachers had problems in expressing their thoughts and in using appropriate words or terms as well, something that had also been pointed out in other studies in Greece (Papageorgiou and Sakka, 2000). For instance, *boiling*, *evaporation* and *heating* were used as synonyms or in an incorrect way by the teachers. Furthermore, these problems in teachers' descriptions sometimes resulted in incoherent answers, which is one of the main reasons for the formation of the category *miscellaneous* in each one of the three tasks (see Table 2).

Among the categories of Table 2 for task 2, teachers' descriptions in the 1a and the 1b sub-categories were correct and complete in carrying out the activity; the main difference between them was the use of the necessary equipment and materials. Teachers in the 1b sub-category did not explain what equipment and materials are needed for this activity (in all or in parts of the procedure) and how they are used.

Teachers' descriptions in the rest of the categories for task 2 did not report any use of the necessary equipment and materials. In category 2, the descriptions did not explain how the collection of the vapors of a substance could be achieved. Teachers of category 3 gave descriptions, which were mainly correct but incomplete, as they did not explain how the isolation of the components of the mixture is achieved. It should be mentioned at this point that, although teachers seemed to get along task 2 in general terms, the part of the procedure concerning the isolation of the components of the mixture (all or part of this procedure) was a problematic one for a remarkable number of them, independently of the categories of Table 2 (for task 2). In cases where teachers referred to this part, the main difficulties concerned the condensation of vapors (water and acetone) and their collection. However, although teachers' descriptions at this point were 'alternatives', 'unclear' or 'incomplete', they could be considered as acceptable to a certain degree, because primary teachers are not expected to have any specific knowledge about such procedures. Some of these teachers' descriptions are: "*In order to condense vapors, we use a cold plate…*", *"…we collect vapors in a vessel, where, they are left in order to cool," "…vapors are collected through a long tube, in order to be condensed*", "*…we use an apparatus similar to that we use in the production of tsipouro"* (*tsipouro* is a traditional Greek drink produced by distillation), etc.

It should also be noted that, although in categories 1-3 boiling and melting points were used by the teachers, there were some cases in category 3 (and in category 6, as well), where the use of these properties was incorrect. For instance, a teacher used the melting point of acetone (-95 °C), instead of the boiling point 56 °C, for the isolation of acetone. As Johnson (1996) has underlined, these are indications of an incomplete understanding of the concept of *substance*. As del Pozo (2001) also has suggested, prospective teachers have difficulties in concepts like *substance* and *mixture*, as well as in relationships between concepts concerning the composition of matter, in general.

Finally, there were a number of teachers, who simply described a general scheme of the procedure (category 4).

Task 3

In task 3, as also in task 2, the use of the appropriate equipment and materials was also an issue (besides the teachers' ability to carry out the activity). Teachers' answers, which are presented in Table 2, were also categorized according to their correctness and completeness.

The 123 teachers, who fell into the categories 1 and 2, described satisfactorily a procedure to find the solubility of the table salt in water. In fact, the differences between these two categories are that: a) in the 2nd category there wasn't a clarification at the end of the procedure of how the solubility is obtained (or it was implied) and b) the use of the necessary equipment and materials was not always complete. However, in both categories, teachers proposed some ways, 'scientific' or 'alternative', in order to measure the amount of table salt dissolved in water, which is a crucial point indicating their ability to carry out this activity. Some of these 'alternative' ways included: addition of small pre-weighed parts of salt in water; use of a volumetric tube and calculation of the amount of salt on the basis of volume; addition at once of a pre-weighed quantity of salt and weighing of the precipitate after filtration, etc. Also, 28 out of these 123 teachers described a more 'scientific' way, based on the weighing of an amount of salt, pre- and post- the dissolution of the necessary part of it in water.

The teachers placed into category 4 did not describe any particular procedure for the determination of the solubility of the table salt in water, and they only referred to its dissolution (with no reference to the quantification of the dissolved salt or to the equipment).

Comparing tasks - How 'gender' is related to teachers' descriptions?

In order to compare teachers' descriptions among the three tasks, a scoring scale was constructed, common for all tasks, on the basis of the particular characteristics of the categories throughout the tasks. For this analysis, only the ability of the teachers to carry out the activity was considered, as the use of the necessary equipment and materials was an issue only in two of the three tasks (tasks 2 and 3). As Table 2 shows, this is a 5-points scoring scale (0-4), each step of which has a distinct difference from the next one regarding the correctness and the completeness of the teachers' descriptions. Using this scale, a corresponding score (0-4) was awarded for each one of the categories of Tables 2 for tasks 1, 2 and 3, respectively. The percentage agreement between the two coders was 94.7%, 96.0% and 95.2% for the tasks 1, 2 and 3, respectively. After discussion, the percentage became 100% for each of the three tasks. Assuming that there is also an impact of the 'gender' on participants' descriptions in each one of the three tasks, a set of Multiple Analysis of Variance (MANOVA) was applied on the relevant scores.

Data analysis showed that the main effect of teachers' descriptions regarding the activities is significant, $F(2,223)= 7.642$, $p=.006$. As Figure 3 illustrates, teachers had more difficulties in carrying out task 3 than the other two tasks, and task 2 was found to be more difficult than task 1. However, it is worth noting that these results do not take into account teachers' descriptions concerning the use of the necessary equipment and materials. Thus, results could be different if teachers' difficulties regarding the use of appropriate equipment and materials for the isolation of the components of a mixture (especially for the condensation of vapors), which was revealed in task 2, were included in this comparison.

Figure 3. The main effect of teachers' descriptions regarding the activities.

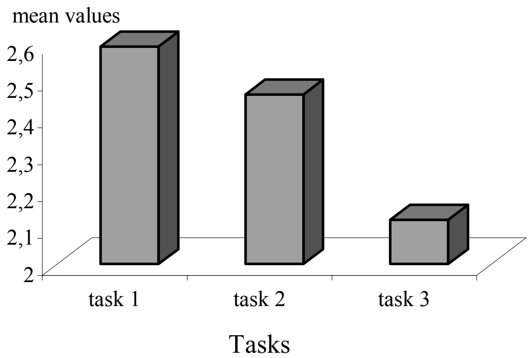
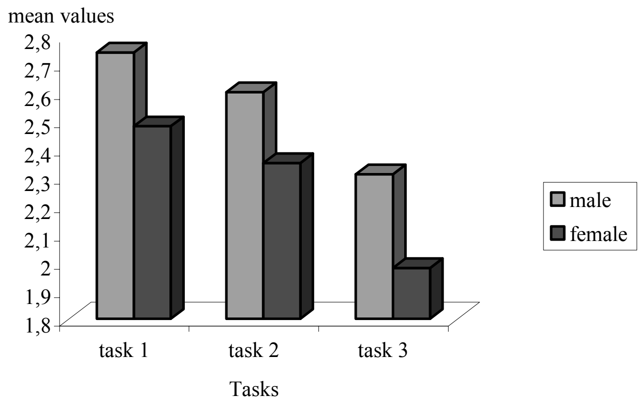


Figure 4. The teachers' descriptions of the three tasks in relation to their gender.



As far as the relation between 'gender' and teachers' descriptions is concerned, it was found to be significant $F(1,226)=3.942$, $p=.048$. Figure 4 shows that the male teachers were found to give more efficient descriptions regarding the procedures of the three science activities than their female colleagues. However, as it has been already reported, male participants estimated that they had used hands-on science activities in school more often than female ones, a fact which had also an impact on their descriptions. As a result, we could not draw a definite conclusion about the main reason for this particular effect on their descriptions of carrying out the three tasks.

Nevertheless, taking into account the teachers' descriptions concerning the use of the necessary equipment and materials (in tasks 2 and 3) the general impression (according to the qualitative analysis of the data) is that male teachers are working in a more practical way, whereas female teachers are more theoretical (and more general) in their descriptions. In particular, in task 2 (Table 2) 82.4% of teachers of sub-category 1a (where there is a use of the necessary equipment and materials) were male, $\chi^2(1)=11.91$, $p=.001$. In sub-category 1b and the two other categories that follow (where the use of the equipment and materials is progressively declined), the proportion of female teachers increased, making no significant the difference between the two genders. In these categories, the corresponding percentages of female teachers were 60.3%, 61.5% and 60.0% per category, respectively. Similarly, in task 3 the majority of teachers in category 1 of Table 2 were male (33 out of the 49 teachers, χ^2 (1)= 8.91, $p=.002$); the opposite holds true for category 2 (49 out of the 74 teachers were female, χ^2 (1)=3.44, $p=.043$.), whereas there is not a statistically significant difference between the two genders in the rest of the categories of Table 2 for task 3 [category 4 (37 out of the 66 teachers were female), category 5 (5 out of the 7 teachers were female), category 6^h (21 out of the 32 teachers were female)].

However, it should be noted that the above analysis relates to teachers' descriptions of carrying out practical work; they are not statements about any teachers' science skills. As a result, any further conclusion drawn for the gender differences in the ability to carry out hands-on sciences activities would be unjustified.

Conclusions and implications

According to the results of this study, although the Greek primary teachers' descriptions regarding some hands-on activities on chemistry topics ranged from a more correct and complete level to a more generally descriptive and incomplete level, it seems that the majority of teachers could carry out such activities to a satisfactory degree.

As far as the teachers' ability to carry out these activities is concerned, task 3 appeared to be the most difficult. This could be seen as an indication that teachers were not used in working on such activities where a number of issues are involved at the same time, such as the dissolving of a substance in water, the measurement of the dissolved amount of this substance, and the effect of the temperature on its solubility. Indeed, hands-on activities which are related to chemistry topics, and especially to solubility, are not very common in primary education, in Greece at least (Greek Pedagogical Institute, 1999).

However, if the teachers' descriptions regarding the use of the necessary equipment and materials are the issue, task 2 appeared to be the most problematic, especially its part concerning the isolation of some of the components. A possible reason might be that this activity has the greatest number of steps related to changes of state. In these steps, the use of 'specific' equipment is required, which is not very familiar to primary teachers. Among these changes of state, condensation seems to be the most difficult for the participants. This may have contributed to the result that task 2 was also the second in difficulty among the three tasks concerning the teachers' ability to carry out the activities. Indeed, although there are no data available about primary teachers' conceptions concerning condensation in literature, this change of state has been reported as one of the most problematic for children (Johnson, 1998; Paik et al., 2004; Papageorgiou and Johnson, 2005). Given that teachers' misconceptions have common characteristics with those of children (Kruger and Summers, 1988; Kokkotas and

Hatzinikita, 1994; Papageorgiou and Sakka, 2000), primary teachers could also have difficulties in understanding condensation, as was seen in the problems in carrying out the corresponding part of task 2.

The gender of the teachers seems to have played a significant role in their descriptions. Additionally, male teachers stated that they had used such activities in school more often than female ones had. As a result it is not clear whether it was the male teachers' more frequent use of hands-on science activities in school that made them better able to describe a procedure of a science activity, or whether it was due to the male identity being connected more to practical sciences, while the female identity is connected to more theoretical sciences (Arnot et al., 1999). However, as has already been discussed, the above results indicate 'gender differences' regarding only the teachers' descriptions of science activities, they offer no insight into any connection to science skills.

Although we acknowledge the simplicity of our research instrument and the limitations of this study as they have been already reported, the findings can suggest some answers to the main questions of this work. These seem to be promising for a successful incorporation of such science activities in the National Curriculum for Primary Education. However, the implementation of science activities in primary schools is also a matter of primary teachers' training on scientific ideas. As it is reported by a number of researchers (Jarvis et al., 2001, 2003; Lavonen et al., 2004) an appropriate in-service training program can significantly improve teacher practice in school. More effective science training programs mean more teachers able to use science activities frequently and effectively towards a better teaching and learning science.

References

Appleton K. and Kindt I., (1999), Why teach primary science? Influences on beginning teachers' practices, *International Journal of Science Education*, **21**, 155-168.

Appleton K., (2002), Science activities that work: perceptions of primary school teachers, *Research in Science Education*, **32**, 393-410.

Appleton K., (2003), How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice, *Research in Science Education*, **33**, 1-25.

Arnot M., David M. and Weiner G., (1999), *Closing the gender gap,* Cambridge: Polity Press.

Asunta T., (1997), In-service science courses for primary teachers: implementation of different types of in-service training courses in Finland, *Science Education International*, **8**, 18-23.

Del Pozo M.R., (2001), Prospective teachers' ideas about the relationships between concepts describing the composition of matter, *International Journal of Science Education*, **23**, 353-371.

Gabel D., (1999), Improving teaching and learning through chemistry education research: a look to the future, *Journal of Chemical Education*, **76**, 548-554.

Greek Pedagogical Institute, (1999), *National program of study for primary and secondary education: science*, Athens (Greece): Greek Pedagogical Institute Publications.

Jarvis T., Mckeon F., Coates D. and Vause J., (2001), Beyond genetic mentoring: helping trainee teachers to teach primary science, *Research in Science and Technological Education*, **19**, 5-23.

Jarvis T., Pell A. and McKeon F., (2003), Changes in primary teachers' science knowledge and understanding during a two year in-service programme, *Research in Science and Technological Education*, **21**, 17-42.

Johnson P.M., (1996), What is a substance? *Education in Chemistry*, **33**, 41-45.

Johnson P.M., (1998), Children's understanding of state involving the gas state, Part 2. Evaporation and condensation below boiling point, *International Journal of Science Education*, **20**, 695-709.

Johnston J., McKeon E. and McEwen A., (1999), Choosing primary teaching as a career: the perspectives of males and females in training, *Journal of Education for Teaching,* **25**, 55-64.

Kelly A., (1987), *Science for girls,* Philadelphia: Open University Press.

Kepler L., (1998), Hands-on science. Getting set up for science, *Instructor*, **108**, 56-57.

Kokkotas P. and Hatzinikita V., (1994), The concept of the molecule in fourth year primary education students of the University of Athens. *In Proceedings of ATTI and European Conference on Research in Chemical Education (2nd ECRICE)*, University of Pisa, Italy.

Kruger C. and Summers M., (1988), Primary school teachers' understanding of science concepts, *Journal of Education for Teaching*, **14**, 13-17.

Lavonen J., Jauhiainen J., Koponen I.T. and Kurki-Suonio K., (2004), Effect of a long-term in-service training program on teachers' beliefs about the role of experiments in physics education, *International Journal of Science Education*, **26**, 309-328.

Paik S.-H., Kim H.-N., Cho B.-K. and Park J.-W., (2004), K-8th grade Korean students' conceptions of 'changes of state' and conditions for changes of state', *International Journal of Science Education*, **26**, 207-224.

Papageorgiou G. and Johnson P. M., (2005), Do particle ideas help or hinder pupils' understanding of phenomena?, *International Journal of Science Education*, **27**, 1299-1317.

Papageorgiou G. and Sakka D., (2000), Primary school teachers' views on fundamental chemical concepts, *Chemistry Education Research and Practice*, **2**, 237-247.

Papageorgiou G. and Tsiropoulou S., (2004), The impact of experiments on students' knowledge and explanations of significant aspects of the greenhouse effect, *Journal of Science Education*, **5**, 28-33.

Schibeci R.A. and Hickey R., (2000), Is it natural or processed? Elementary school teachers' and conceptions about materials, *Journal of Research in Science Teaching*, **37**, 1154-1170.

Smith D.C. and Neal D.C., (1989), The construction of subject matter knowledge in primary science teaching, *Teaching and Teacher Education*, **5**, 1-20.

New visualization tools for learning molecular symmetry: a preliminary evaluation

Inbal Tuvi-Arad* and Paul Gorsky

Department of Natural Sciences, The Open University of Israel, Israel.
e-mail: inbaltu@openu.ac.il

Received 2 August 2006, accepted 20 December 2006

Abstract: A Website that helps students visualize and locate symmetry elements on three-dimensional molecular structures was developed. It includes textual explanations, an interactive example window and a Microsoft-Excel based symmetry toolkit that enables students to draw symmetry elements in three dimensions. Preliminary qualitative research aimed at exploring how students learned with this tool was performed. It was found that the three-dimensional graphical capabilities of the toolkit (1) helped students overcome difficulties in three-dimensional visualization, (2) enabled students to find symmetry elements of complex molecules generally not accessible from drawings and (3) contributed to a deeper understanding of molecular structure and chemical symmetry. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 61-72]

Keywords: Molecular visualization, spatial ability, molecular symmetry, point-group symmetry operations, Web based learning.

Introduction

Molecular visualization is playing a central role in science education in general and in chemistry education in particular (Mathewson, 1999; Jones, Jordan and Stillings, 2001, 2005; José and Wiliamson, 2005). In a recent review, Wu and Shah (2004) summarized the literature regarding visual-spatial thinking in chemistry. They found a positive correlation between students' visual-spatial abilities and their achievements in learning chemistry, in that students with higher visual-spatial abilities performed better than their peers on problem solving in chemistry (both spatial and non-spatial problems). They also pointed out that good three-dimensional (3D) visualization tools (both plastic models and computerized models) enhance students' understanding of molecular structure.

The study of Ferk and Vrtacnik (2003) supports these claims. They investigated how students understand different kinds of molecular representations, and their ability to apply mental operations to molecules. Their findings emphasized the importance of the correct perception of the 3D molecular structure as the basic step preceding any further mental operations. They found that undergraduate university students preferred 3D computerized models or 3D photographs of molecules over other representations (e.g., two-dimensional stereochemical formulas and plastic models). On the other hand, plastic models were more effective for secondary school students.

Molecular symmetry is a topic in chemistry grounded in visualization. It is taught in several undergraduate chemistry courses (e.g., inorganic chemistry, spectroscopy and quantum chemistry) each with its own context and at varying conceptual levels. Understanding the basic concepts of molecular symmetry involves identifying particular symmetry elements of molecules such as rotation axes, reflection planes and inversion points.

To do so requires three skills: (1) the ability to visualize and to understand molecular structures in three-dimensions (spatial visualization), (2) the ability to see and to describe the molecular structure after a rotation, reflection or inversion (spatial orientation skills), and (3) the ability to define precisely the location of symmetry elements in relation to particular atoms. Achieving such conceptual understanding is made harder by the need for students to recognize 3D structures from the static two-dimensional representations found in textbooks. The task becomes even more difficult as the number of atoms per molecule increases and the 3D structure becomes more branched.

A Website entitled "*Molecular Symmetry Online"* that helps students learn molecular symmetry was recently developed at the Open University of Israel by the first author. It includes textual explanations, an interactive example window and a Microsoft-Excel based symmetry toolkit that can be used for drill and practice. The toolkit is an open tool – it enables users to create and draw symmetry elements for *any* molecule in 3D. In this sense it is very different from other recently developed Websites that present symmetry elements within a closed set of molecules (e.g., Korkmaz and Harwood, 2004; Cass and Rzepa, 2005a, 2005b; Charistos et al., 2005).

This study focuses primarily on how students learned while using the symmetry toolkit. A secondary focus was on the capabilities and limits of the tool itself. In what follows, we briefly present the basic concepts of molecular symmetry. We then describe the symmetry toolkit and elaborate on its usage. In the second part of the paper, we describe the findings of preliminary qualitative research performed at the Open University of Israel aimed at exploring how students learned with the toolkit.

Symmetry elements and operations

The symmetry of molecules is described in terms of elements and operations. Elements are geometric entities such as axes, planes and points in space used to define symmetry operations. Operations involve the spatial re-arrangement of atoms in a molecule by rotation about an axis, by reflection through a plane or by inversion through a point. A rotation, reflection or inversion operation will be called a symmetry operation if, and only if, the new spatial arrangement of the atoms in the molecule is *indistinguishable* from the original arrangement. The definitions and applications of symmetry elements and operations are discussed in many textbooks (e.g., Atkins and de Paula, 2006) and will not be repeated here. Table 1 summarizes the symbols of all the possible symmetry elements and operations in molecules.

Table 1. Symmetry elements and operations in molecules.

Symmetry Operation	Symmetry Element	Symbol
Identity	None	E
Rotation about an axis by 360°/*n*	Axis of order *n* (proper)	C_n
Reflection through a mirror plane	Plane	σ_h, σ_v, σ_d
Inversion through a point	Point	i
Rotation about an axis by 360°/n followed by reflection through a plane perpendicular to the rotation axis	Axis of order n (improper)	S_n

The *Molecular Symmetry Online* Website

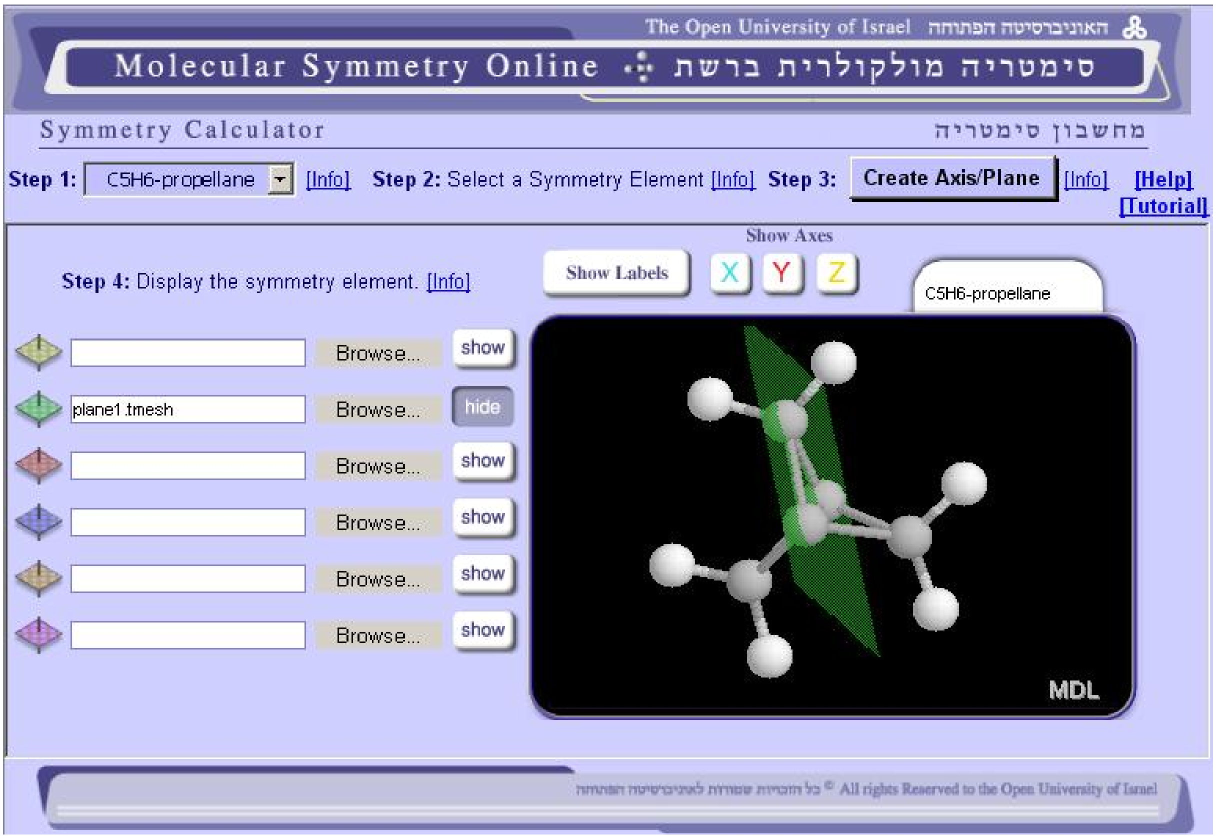
In order to improve visualization of symmetry elements of molecules, and to improve the way students learn symmetry at the Open University of Israel, we recently developed the

Molecular Symmetry Online Website (http://telem.openu.ac.il/symmetry); it comprises two main sections:

- An interactive *example window* presenting six molecules in three-dimensions that users can rotate and translate in space, with or without their symmetry elements. The example window also includes a chemical datasheet for each molecule, a basic tutorial on molecular symmetry and a technical help file.
- The *symmetry toolkit* including a *symmetry calculator* that can be used to create symmetry elements (rotation axes and reflection planes) for any molecule and display them on screen in an interactive, three-dimensional format. The toolkit also includes a database of about 60 molecules, as well as comprehensive didactical and technical help files.

A snapshot of the symmetry calculator is presented in Figure 1.

Figure 1. A snapshot of the symmetry calculator. Working with the calculator is done in four steps. Step 1: loading a molecule from the database, or from the user's computer. Step 2: deciding which element to draw. This is a virtual step. Users can use the 'show label' button to draw the molecule in wireframe with serial numbers on each atom. Step 3: creating the symmetry element using the element and planes wizard. Step 4: loading the symmetry elements created in step 3. Clicking on 'info' at each step opens the relevant help file.



Technology

The Website was built using the Open University of Israel's OPUS system for online courses. The *example window* is a dynamic HTML page in which an MDL-CHIME window is embedded to view molecules in three-dimensions. The *symmetry calculator* is constructed of such an HTML page embedded in Microsoft-Excel and enhanced with a Visual Basic user-friendly interface.

Working with the symmetry calculator is done in 4 steps:

1. **Load a molecule**. The current database contains about 60 molecules, but additional molecules in MDL mol format can be loaded through a browse option.
2. **Decide where to draw a symmetry element**. A virtual step. To define a symmetry element mathematically, the points through which it passes must be specified. Two points in space are used to define a rotation axis. Three points are needed to define a mirror plane. At this stage, inversion points can not be drawn. A point may be defined in one of three ways:
   - a specific atom in a molecule
   - a mid-point between two atoms
   - any point in space

   Atom labels (serial numbers) and Cartesian axes can be displayed in order to help users visualize and specify these points. The molecules in the database were set such that their center of mass is located at the origin. This makes it easier to use the Cartesian axes to define rotation axes that do not go through atoms of the molecule (e.g., C_6 axis in benzene).
3. **Create a symmetry element**. A specially designed 'axes and planes wizard' is used to enter the coordinates of the points in space through which the symmetry element passes. Elements are saved on the user's hard disk. A snapshot of the wizard is shown in Figure 2.
4. **Display the symmetry elements**. Up to six different elements per molecule can be displayed simultaneously. The toolkit displays the elements as defined by the user; it does not evaluate whether the elements are correct or not. Therefore, users can create any element, regardless of it being a symmetry element or not.

Figure 2. A snapshot of the 'axes and planes wizard'. Top: in this screen the user chooses the type of point. Bottom: in this screen the user enters atom types and serial numbers to define a mid point between atoms. The table on the top shows the coordinates of the chosen points.

Usage

The toolkit can be used for drill and practice, testing and classroom presentations. It has been designed as an open tool; therefore it can be used at various content levels and for diverse audiences. Teachers can add molecules to the database thereby allowing students to work on any molecule of interest. Furthermore, teachers may create a solution database so that students can compare their answers with the correct ones. Students may work alone or in groups. The toolkit can also be used for testing purposes. To do so, students submit their calculated symmetry elements for teacher or peer evaluation.

Here are examples for three assignments, based on the symmetry toolkit that can be given to students.

1. Creating symmetry elements of molecules

In this assignment, students are given a list of molecular files for which they need to find part or all of the symmetry elements and draw them on the screen. The complexity of the assignment is a function of several factors:

a. The way symmetry elements are defined
b. The number of symmetry elements in the molecule
c. The spatial complexity of the three-dimensional structure.

The importance of each of the above factors depends on the molecule in question and the student's experience with the symmetry calculator. For example, the water molecule, H_2O, has a very simple planar structure, but defining the reflection plane perpendicular to the molecular plane is a bit tricky since, in the language of the symmetry calculator, it cannot be defined solely by the atoms of the molecule. In contrast, the phosphorous pentachloride molecule, PCl_5, has more atoms and more symmetry elements than water, but these are easily defined in terms of the molecule's atoms. A third example, finding the symmetry elements of urazole, $C_2H_3N_3O_2$ (Figure 3), is more difficult than water, even though they both belong to the same point group (C_{2V}), because the urazole molecule has more atoms than the water molecule and the structure is branched.

For tutoring purposes, we recommend starting with the molecules presented in the example window: CCl_2O, NH_3, N_2F_4, C_3H_4 (allene), PCl_5 and CH_4 (see Figure 4). Although students can see the solutions, locating the symmetry elements and drawing them on the screen is a different task that leads to a deeper understanding of molecular structure. Advanced students can find molecules like twistane (D_2) and Paddle (D_{3h}) quite challenging (see Figure 5, later).

Figure 3. Urazole.

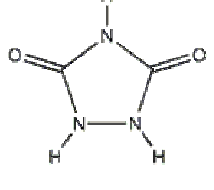
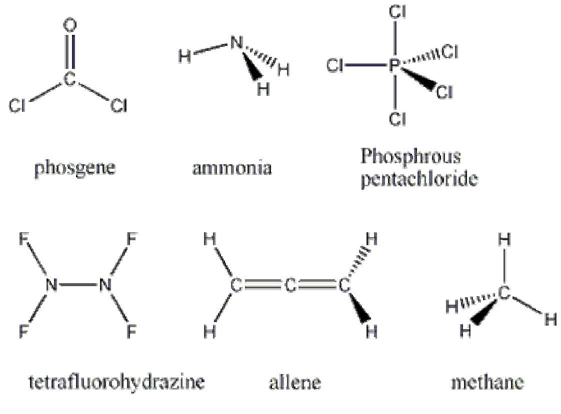


Figure 4. Molecules presented in the example window.



Within the frame of the toolkit, there are several options to provide feedback for the students:

1. Providing files containing the correct symmetry elements. Students can overlay these files onto their own files within the symmetry calculator and see if they match.
2. Providing the answer in the form of the points in space (atoms, midpoints or others) through which each element passes.
3. On-site feedback from a teacher/tutor who watches the student's work.
4. Asynchronous textual feedback from a teacher/tutor.

2. Symmetry Reduction

Another type of assignment involves the analysis of symmetry after substitutions. For this purpose, the teacher should prepare a series of molecules where one or more of the atoms are substituted by other groups. Planar structures like aromatic cyclic molecules or crown ethers are good examples, but more complex structures can be used for advanced students. One can start with benzene, for example, and ask students to find all of its symmetry elements. In the next step, the students are given chlorine derivatives of benzene. Students should then decide which of the symmetry elements of benzene remain symmetry elements for each derivative. For this purpose, they do not need to redraw the elements, just overlay them on each derivative.

3. Defining the location of symmetry elements

In this assignment, the teacher prepares in advance files of elements for a number of molecules. The students then need to decide whether these elements are symmetry elements for the listed molecules, and describe the location of these elements, either in words, or by creating their own elements to match the teacher's elements.

The current study

Background

The Open University of Israel is a distance education university designed to offer academic studies to students throughout Israel. Established in 1974, the university offers a home study system based on textbooks, instructors, face-to-face tutorials and course Websites. The university offers a variety of programs for both bachelor and master degrees. Enrollment for the academic year 2004-2005 was about 40,000 students. As course study loads are relatively high, most students register for only two or three courses per semester; most do not complete any degree program. On average, it takes about seven years to finish a bachelor's degree.

The course, inorganic chemistry, from which students volunteered to participate in the study, is an intermediate level course required for all chemistry majors. It accounts for 6 credits out of a total 108 needed for graduation. The course material includes 12 self-study text units, based on Shriver and Atkins' (1999) *Inorganic Chemistry* textbook. The second unit deals with molecular symmetry. Molecules and their symmetry elements are presented in this unit by two-dimensional drawings. As part of the learning materials, students receive kits (plastic balls and sticks) that are used to build three-dimensional models of molecules. There are six face-to-face tutorial sessions during the semester in which students may participate. Participation is not obligatory. The first session is devoted to molecular symmetry. In addition, students submit written assignments during the course. The first assignment includes problems in molecular symmetry.

Research Rationale and Objective

After an extensive literature review, no previous studies were found that tested the way students learn molecular symmetry using interactive three-dimensional representations of molecules and symmetry elements. Initial investigations into a phenomenon as subtle and complex as student learning require an in-depth examination of individual learners in order to produce preliminary findings. A grounded theory approach (Glaser and Strauss, 1967; Glaser, 1978; Cohen and Manion, 1989; Charmaz, 1995; Strauss and Corbin, 1998) generally begins with small naturalistic studies carried out in order to generate models and hypotheses that can then be tested on larger populations using traditional statistical techniques.

Based on these assumptions, a small naturalistic study was planned and carried out. Objectives were twofold. First, to describe how students learn molecular symmetry by means of the Website's features (the interactive example window and the symmetry toolkit), and second, to judge their perceived effectiveness in relation to the two-dimensional methods used during the course.

Participants

18 students enrolled in the course 'Inorganic Chemistry' during the spring semester of 2004. All students were offered the opportunity to participate in the study. Six chose to do so (five men and one woman). All students who took part in the study had previously completed the course 'General Chemistry' which is a prerequisite for registration in 'Inorganic Chemistry'. They were, therefore, experienced in distance learning at the Open University. All students submitted the assignments on molecular symmetry and received passing grades (above 75 on a scale from 0 to 100). The average grade was 87. The average grade of the whole class was similar (10 out of 18 students submitted this assignment).

Interviews with students

Students participating in the study were invited to a personal interview/tutorial session that lasted between 90-120 minutes. Interviews/tutorials were carried out by both authors who were in no way directly related to the course. This was emphasized so that students would not be reticent or apprehensive about offering information freely. Complete confidentiality was assured. The interviewers did not know any of the participating students. Sessions took place towards the end of the semester, prior to the final exam and about two months after the students had completed the second unit of the course which dealt with symmetry. From the students' point of view, these sessions were an opportunity to have a private lesson prior to the final exam.

All sessions were tape-recorded. During the session, one interviewer made brief notes; immediately upon its completion, both interviewers made extensive notes. The interview/tutorial had four parts: (1) preliminary discussion, (2) pretest, (3) instructional session and (4) summation.

The preliminary discussion was intended to create a friendly and open environment. Students were told about the study and its goals. They spoke about themselves, their studies, goals, future plans, etc.

The pretest included the following questions about basic symmetry concepts:

- What is a rotation axis? Draw an example.
- What is a reflection plane? Draw an example.

Students were also asked to find the symmetry elements of water and methane. These molecules were chosen on the basis that students were familiar with their symmetry, since they were previously discussed in the course. For each molecule they were asked to:

- Draw the molecule and all of its symmetry elements.
- Describe verbally the spatial location of the symmetry elements.

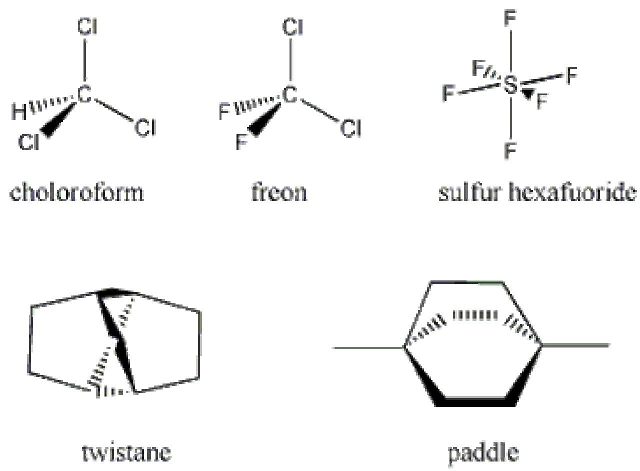
- Draw the molecule before and after specific symmetry operations by assigning serial numbers to the atoms.

Students were expected to answer these questions on paper, and/or orally. Plastic models were not provided.

The instructional session included:

- An introduction to the symmetry toolkit by using the "example window" where they could see the symmetry elements of methane in interactive 3D format.
- A short demonstration of how to use the symmetry calculator, during which the symmetry elements of the water molecule were drawn on screen.
- A work session in which students were asked to find symmetry elements of molecules by using the symmetry calculator. The set of molecules included: chloroform, Freon, paddle, twistane, and sulfur hexafluoride. These molecules are shown in Figure 5.

Figure 5. The set of molecules included in the instructional session.



The summation included questions about students' use of the toolkit, especially how they learned and their evaluation of its effectiveness.

Results

Preliminary discussion

Preliminary discussion created a friendly and open environment for all participants. Students spoke freely, and any initial unease or tension about being questioned and judged quickly dissipated. Students viewed the session as a tutorial aimed to help them understand symmetry concepts which, at the same time, aimed to help us understand the learning process.

Pretest results

All students were able to identify the symmetry elements of water (however, 2 out of 6 had problems with one of the reflection planes). Most of them failed to see all the symmetry elements of methane. Since the C_3 axes of methane overlay with the C-H chemical bonds, these were easily identified by all students. However, only one student identified the C_2 axes that are located between the bonds. The reflection planes were identified partially by four students, fully by one student and not identified at all by a sixth student. Table 2 summarizes the pretest findings. For the sake of privacy, students' names were changed.

Table 2. Students' spatial abilities prior to working with toolkit.

Molecule →	H_2O			CH_4		
Element → / **Name ↓**	C_2	σ_v	σ_v'	$3C_2(=S_4)$	$4C_3$	$6\sigma_d$
Oren	Yes	Yes	No	0/3	4/4	3/6
Moshe	Yes	Yes	Yes	0/3	4/4	3/6
Danny	Yes	Yes	Yes	0/3*	4/4	3/6
Jacob	Yes	Yes	Yes	0/3	4/4	1/6**
Hanna	Yes	Yes	Yes***	0/3	1/3**	No
David	Yes	Yes	Yes	3/3	4/4	6/6

Instructional sessions and summations

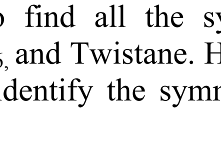
All students were introduced to the Website by using the 'example window' where they could see the symmetry elements of methane in interactive 3D format. This was followed by a short demonstration of how to use the 'toolkit', during which the symmetry elements of the water molecule were drawn on screen. All students readily grasped the toolkit's methodology for drawing symmetry elements; that is, using atoms, midpoints between atoms and any point in space to define axes and planes. In what follows we describe in detail the work session of four students.

Moshe dealt with the first molecule presented, chloroform, by creating his own mental image of the molecule from the diagram in order to identify the symmetry elements. He began with reflection planes and found all of them, without using the symmetry toolkit. Moshe reported that he was able to visualize the model in his own mind and see the symmetry elements clearly. Immediately afterwards, he replicated the same task, this time using the toolkit. He easily identified and drew the symmetry elements. The 3D representations enabled him to confirm that his mental images were indeed correct. He said: "*I can see the symmetry element in my mind, but I'd rather use the toolkit to draw it and check if I'm right*."

Next, Moshe chose to work on acetone (CH_3COCH_3, see Figure 6). Although this molecule was not on the list of molecules we intended to examine, we allowed him to continue. He created his own mental image of the molecule from the diagram but failed to identify the symmetry elements. Using the toolkit, however, he was able to identify all the rotation axes and reflection planes. For the following two molecules, paddle and twistane, Moshe made no attempt to create his own mental images (the molecules were too complex) and began directly with the toolkit. For both molecules, he identified the rotation axes. He also identified some of the reflection planes of paddle.

To conclude, Moshe pointed out that the toolkit enabled him to skip the critical step of internal visualization. Since the toolkit enabled him to view the three dimensional structure directly, he could concentrate on finding symmetry elements. Using the toolkit, he was able to find elements that otherwise would not have been found.

Figure 6. Acetone.



Danny used only the toolkit to find all the symmetry elements for the following molecules: $CHCl_3$, CCl_2F_2, Paddle, SF_6, and Twistane. He did not draw the symmetry elements on the screen since he was able to identify the symmetry elements with a high degree of certainty just by viewing the 3D molecular structure. He found the 'axes and planes wizard' of the toolkit "*tedious and time consuming*". Danny reported that, although he was capable of imagining simple 3D structures in his mind, he preferred to work directly from the screen without any internal visualization. He said: "*I no longer need to imagine the structure in my mind*."

Hanna was incapable of visualizing any three dimensional molecular structures from their drawings on paper. This limitation severely handicapped her understanding of the basic concepts of symmetry. She was unable to define rotation axes and reflection planes since she had no language through which she could do so. She was unable to draw any molecule except HCl.

In order to proceed, Hanna was reintroduced to the concepts of symmetry by using the 'example window'. She was then introduced to the toolkit. Together with the researcher, she worked on two molecules, water and methane, for about twenty minutes. Upon completion, she was able to define and to identify rotation axes and mirror planes. She said: "*Before I couldn't see anything. Now I do*."

Next, she worked alone on two molecules, chloroform and Freon. She was able to identify rotation axes and mirror planes in both. She was extremely excited, even thrilled, at her new found visualization skills. Despite the two hour length of the tutorial session, Hanna wanted to continue and work on additional molecules. The authors politely declined.

David used the toolkit to find *all* the symmetry elements for the following molecules: Freon, Paddle and Twistane. For the first two molecules, he did so quickly and easily, claiming that the 3D visualization is not really essential since he can visualize them in his mind. However, he noted that the 3D representations of the supposed symmetry elements helped him to decide if, in fact, they were correct. For the twistane molecule, he worked confidently, but required more time to complete the tasks. He said: "*The toolkit really helps here because I can't visualize this molecule at all without it*".

Concluding Remarks

The Website's primary features, the 'example window' and toolkit, helped students to overcome difficulties in 3D visualization. Even students with good visualization and orientation skills preferred to work with the computer graphics, especially for complex molecules, thereby eliminating the need for the now redundant step of internal visualization. Interviews with the other two students, Oren and Jacob, not described above, support these findings.

Usually, tasks that require finding symmetry elements of molecules rely on student's intuition and drawing capabilities. Questions of this type are usually phrased: "*Find the symmetry element of ...*" Students' answers to such questions are usually given in the form of a drawing, without explanation. The symmetry toolkit provides an alternative didactical way to solve these problems. In order to draw a symmetry element on screen, it must be defined as a geometrical entity, that is, the points in space through which it passes must be specified. In our opinion, solving the task in this way contributes to a deeper understanding of the basic concepts of symmetry, as was seen in the case of Hanna described above. She represents an extreme example of a student with difficulties in spatial ability. The toolkit not only helped her understand the three-dimensional structure of molecules – it also provided her with a language to solve problems. Nevertheless, further research is needed to generalize this finding.

From an instructional point of view, the Website can be used for drill and practice, testing and classroom presentations. It is an 'open tool' that provides teachers with an interactive framework to teach the molecular symmetry of any molecule. It can therefore be used for various courses at different levels and for diverse audiences. Further research with this tool is required in order to optimize its use. Some tentative research ideas are: exploring the sequence of molecules that will allow students to learn basic concepts of symmetry with maximal success and minimal frustration; understanding the origin of students' difficulties in finding symmetry elements – spatial visualization, spatial orientation or the ability to define the location of symmetry elements; comparing the way students solve problems on paper before and after working with the toolkit.

From a technological point of view, the major advantage of the Website is its ability to define, save and upload symmetry elements of any molecular structure so desired by teacher, student or researcher. Further avenues of development might include creating an online version of the toolkit; replacing Chime with Jmol (http://jmol.sourceforge.net); adding other features such as solutions for the set of molecules in the database and assigning point groups for these molecules. More sophisticated developments may include applying the method of the continuous symmetry measure (Zabrodsky et al., 1992; 1993) in order to provide an immediate feedback to symmetry elements drawn for any molecule.

The authors will be happy to collaborate in further research and development of the Molecular Symmetry Online Website. Please contact the first author for details.

Acknowledgements

We would like to thank Dov Garmise and Vered Tubi from the Center for Information Technology in Distance Education at the Open University of Israel, for their help in developing the *Molecular Symmetry Online* Website. The technology that enabled the development of the Website is based on the Symmetry Tutorial Website developed by Prof. Dean H. Johnston of Otterbein College in Westerville, Ohio (http://www.otterbein.edu/home/fac/dnhjhns/symmetry/symmetry.html). We greatly appreciate Prof. Johnston's help. We would also like to thank Prof. Eric Martz of the University of Massachusetts, Amherst for fruitful discussions regarding the development of the Website. Finally, we would like to thank Dr. Dina Yogev-Einot, lecturer of the course Inorganic Chemistry at the Open University, for her cooperation in conducting this research.

References

Atkins P.W. and de Paula J., (2006), *Physical chemistry*, 8th ed., Oxford University Press, UK.

Cass M.E. and Rzepa H.S., (2005a), The use of the free, open-source program jmol to generate an interactive web site to teach molecular symmetry, *Journal of Chemical Education*, **82**, 1736-1740.

Cass M.E. and Rzepa H.S., (2005b), An animated interactive overview of molecular symmetry, *Journal of Chemical Education*, **82**, 1742-1743.

Charistos N.D., Tsipis C.A. and Sigalas M.P., (2005), 3D molecular symmetry shockwave: a web application for interactive visualization and three-dimensional perception of molecular symmetry, *Journal of Chemical Education*, **82**, 1741-1742.

Charmaz K.. (1995), Grounded theory, in: J. Smith, R. Harre and L. Van Langenhove (Eds.). *Rethinking methods in psychology.* Sage Publications, London, UK.

Cohen L. and Manion L., (1989), *Research methods in education*. Routledge, London, UK.

Ferk V. and Vrtacnik M., (2003), Students' understanding of molecular structure representations, *International Journal of Science Education*, **25**, 1227-1245.

Glaser B., (1978), *Theoretical sensitivity: advances in the methodology of grounded theory,* The Sociology Press, California, USA.

Glaser B. and Strauss A., (1967), *The discovery of grounded theory: strategies for qualitative research.* Aldine de Gruyter, New York, USA.

José T.J. and Williamson V.M., (2005), Molecular visualization in science education: an evaluation of an NSF-sponsored workshop, *Journal of Chemical Education*, **82**, 937-943.

Jones L., Jordan K., Stillings N., (2001), Report from the molecular visualization in science education workshop. http://pro3.chem.pitt.edu/workshop/workshop_report_180701.pdf .

Jones L.L., Jordan K.D. and Stillings N.A., (2005), Molecular visualization in chemistry education: the role of multidisciplinary collaboration, *Chemistry Education Research and Practice*, **6**, 136-149.

Korkmaz A. and Harwood S.W., (2004), Web-Supported chemistry education: design of an online tutorial for learning molecular symmetry, *Journal of Science Education and Technology,* **13**, 243-253.

Mathewson J.H., (1999), Visual-spatial thinking: an aspect of science overlooked by educators, *Science Education,* **83**, 33-54.

Shriver D.F. and Atkins P.W., (1999), *Inorganic chemistry*, 3rd ed., Oxford University Press, UK.

Strauss A. and Corbin J., (1998), *Basics of qualitative research: techniques and procedures for developing grounded theory*, Sage Publications, Thousand Oaks, CA.

Wu H-K. and Shah P., (2004), Exploring visuospatial thinking in chemistry learning, *Science Education*, **88**, 465-492.

Zabrodsky H., Peleg S. and Avnir D., (1992), Continuous symmetry measures, *Journal of the American Chemical Society,* **114***,* 7843-7851.

Zabrodsky H., Peleg S. and Avnir D., (1993), Continuous symmetry measures II: symmetry groups and the tetrahedron, *Journal of the American Chemical Society,* **115**, 8278-8289.

Lectures: electronic presentations versus chalk and talk – a chemist's view.

Dudley E. Shallcross* **and Timothy G. Harrison**†

Bristol ChemLabS, School of Chemistry, Bristol University, BS8 1TS.
e-mail: d.e.shallcross@bris.ac.uk

Received 28 January 2006, accepted 8 December 2006

Abstract: An extensive survey of undergraduate Chemistry lectures from years 1-4 during 2004-2005 has been undertaken. They were categorised according to the method used for delivery, where category 1 used only electronic media to deliver courses, category 2 used a mixture of electronic and non-electronic and category 3 used non-electronic only. Analysis of student questionnaires, coupled with interviews with a selection of students and lecturers from each category, revealed that the impact of the method of lecture delivery is very slight indeed. Non-electronic methods were preferred, but the differences were not significant. The main problems identified with electronic presentations were: that too much material was covered, hard copies of the notes were not provided, the presentation contained particularly complicated diagrams or seemingly irrelevant images, and lectures were presented too quickly. In addition, it was observed that there was a tendency for lectures given using electronic media to have fewer (or no) breaks, natural or otherwise. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 73-79]

Keywords: Electronic presentations, lectures,

Introduction

Lectures can be traced as far back as the Greeks of the fifth century BC, and in medieval times lectures were the most common form of teaching (Brown and Atkins, 1988). Therefore, the lecture has its merits, otherwise this form of teaching would have ceased. Cannon (1988) noted that research comparing lecturing with other teaching methods provides insufficient evidence to favour one method over another. However, he also notes that discussion methods in small groups appear to be a superior method of attaining higher-level intellectual learning. Once students enter University, it is almost inevitable that they will experience lectures, irrespective of the chosen subject. It is often impossible, given the limited number of academics in any University department, to provide small group classes to cover a particular course or module when the number of students attending is so large. In some countries limited resources, e.g. where perhaps only the lecturer has a textbook, force the wholesale use of lectures as the medium for education. Hence, as Walton (1972) notes, the lecture is here to stay, so it is imperative that it should be as an effective teaching method as possible.

In recent times the use of electronic media has become commonplace in Universities, as well as secondary and primary schools. Recent studies have sought to determine whether using PowerPoint or other such media are superior forms of delivery for lecturing over the traditional 'chalk and talk' or the use of transparencies and an overhead projector (TOHP). The study of Bartsch and Cobern (2003) noted that students preferred PowerPoint over the

† Presently School Teacher Fellow at Bristol ChemLabS

use of TOHP, but that in some instances the content of the PowerPoint presentation distracted students and they performed less well on tests compared with a control group. Szabo and Hastings (2000) carried out an extensive study comparing PowerPoint and TOHP and observed no difference in student performance in tests; the most important factor was lecture subject difficulty in determining the students' performance in these tests. They concluded that the efficacy of using PowerPoint was case specific rather than universal. The study of Lowry (1999) saw a marked improvement in examination results when PowerPoint replaced the use of TOHP. Therefore, there is a mixture of views based on recent studies. In this study we investigate students' opinions of the impact of electronic presentations in lectures in undergraduate chemistry compared with TOHP and other traditional non-electronic approaches, and compare the impact on examination results.

Methodology

The School of Chemistry at Bristol is one of the largest undergraduate teaching institutes in Chemistry in the UK with 130-200 students in any one year. Lectures are arranged, for administrative purposes, under the traditional headings of Physical and Theoretical, Inorganic and Materials, and Organic and Biological. Students either take a 3 year course leading to a B.Sc. or a 4 year course leading to an M.Sc. This work is based on data taken from the academic year 2004-2005. We first invited lecturers to declare which of three categories they fall into for each course that they lecture. The categories are; 1, always uses electronic based media (mainly PowerPoint), 2, uses a mixture of electronic media and traditional methods such as TOHP, or chalk and blackboard, and finally those who use only non-electronic methods are in category 3. The responses to student questionnaires on all courses have then been analysed; first, from the perspective of each category of delivery as a whole, and second, the differences between undergraduate years 1 to 4, and third, differences between lecture groupings of Physical and Theoretical, Inorganic and Materials, and Organic and Biological. Some students and lecturers were interviewed further in the light of the analysis of questionnaires, and the salient points will be discussed later. The student questionnaire asked the students to rank the following (among other questions), where the scale of 1 (Agree strongly) to 5 (Disagree strongly) could be augmented by written comments on the particular lecture course.

1. The lectures were well organised	1 2 3 4 5
2. The lecturer was audible	1 2 3 4 5
3. The board work and/or visual aids were clear	1 2 3 4 5
4. The associated handouts and/or web pages were useful	1 2 3 4 5
5. The lectures stimulated my interest	1 2 3 4 5
6. This course of lectures advanced my understanding	1 2 3 4 5

Between 1999 and 2003, a variety of audio-visual aids had been added to lecture theatres and lecturers began to use electronic media for lectures. In 2004 the lecture theatres were refitted and in each one there is the option to use the web, PowerPoint and other packages, TOHP, blackboard and chalk, a volumiser (where writing can be projected onto a screen from a normal piece of paper), as well as videos.

Results of the questionnaires and their discussion

First, of the 45 academics interviewed, 9 (20%) of lecturers fell into category 1, *i.e.* they used electronic media only, 22 (49%) of lecturers fell into category 2, *i.e.* they used a mixture of electronic and non-electronic and 14 (31%) of lecturers fell into category 3, *i.e.* they used no electronic media. Second, the overall score from the questionnaire was used in these results. The results from the analysis are presented in Tables 1 and 2.

Table 1. Questionnaire scores expressed as a percentage for each of the 5 options for lecturers in the 3 categories of preferred presentation medium, first for all four years and then in individual years. See text for comments on year 4 data in categories 2 and 3.

All years

Questionnaire Scores %	1	2	3	4	5
Category 1	6	60	33	1	0
Category 2	10	63	27	0	0
Category 3	12	70	16	0	0

Year 1

Questionnaire Scores %	1	2	3	4	5
Category 1	11	68	19	2	0
Category 2	12	75	12	1	0
Category 3	14	78	8	0	0

Year 2

Questionnaire Scores %	1	2	3	4	5
Category 1	6	59	34	1	0
Category 2	8	65	27	0	0
Category 3	11	67	22	0	0

Year 3

Questionnaire Scores %	1	2	3	4	5
Category 1	8	58	31	3	0
Category 2	13	75	12	0	0
Category 3	11	75	14	0	0

Year 4

Questionnaire Scores %	1	2	3	4	5
Category 1	7	60	33	0	0
Category 2	6	49	39	6	0
Category 3	35	64	1	0	0

Comparisons across all categories

Table 1 shows the questionnaire scores for each category of lecturer for all years and also split into individual years. In year 4 in category 2 there were two courses given by the same lecturer that skews the result and in category 3 there was just one course given. Although the scores for years 1-3 and also the overall scores are very similar across categories, it could be argued that there is a small but consistent pattern in years 1-3 where category 1 fares less well than categories 2 and 3. The observation that there is neither a negative or positive effect of using electronic media is consistent with the work of Szabo and Hastings (2000). Many examination questions contain elements from several courses, and in each examination paper there is an element of internal choice. Therefore, an analysis of examination score from every course is difficult to obtain. However, where it is possible to compare examination result with a course there is a clear trend; lecture courses that yielded a higher score on questionnaires (less well received) had the lowest examination score. Where courses are perceived to be hard to understand or contain difficult material, examination marks are lower (Szabo and Hastings, 2000). Accompanying comments on questionnaires suggest that the courses which suffer poorest examination results are indeed the ones difficult to understand. We find no statistical difference between method of delivery and poor examination score. Our evidence does not support the study of Lowry (1999) and Bartsch and Cobern (2003) and this will be discussed later. There is also a remarkably consistent pattern across the years.

On interviewing students from various years and lecturers in each category, some useful practical comments were recorded.

- The most effective lectures, regardless of the method of delivery, were ones where lecture notes were supplied that could be annotated during the lecture. Students wanted to be able to listen to the lecturer and make their own notes, in addition to the material presented. They did not want to be simply writing material down all the time. They particularly stressed that copies of diagrams, etc. were essential.
- Of those electronic presentations that were disliked, the main reasons were: hard copies of the notes were not provided, the presentation contained too much material, the presentation contained particularly complicated diagrams or seemingly irrelevant images, and the lectures were delivered too quickly.
- Electronic presentations avoided the issues of poor handwriting, seen with TOHP or chalk and talk methods, and the perennial issue of dirty blackboards impairing legibility of material, especially if several lectures have used the blackboards that day.
- Lectures with a heavy mathematical bias were universally not popular when presented using only electronic media. However, students did find that good visual displays that augmented mathematical material were very beneficial, especially where it was interactive and allowed the students to explore concepts off-line for revision purposes.
- Several lecturers felt that using PowerPoint or html forced them to think very carefully about the material they presented and the way they presented it. Lecturers who never used electronic media found this medium inflexible. Many said that they wanted to have the ability to respond to the audience, and when they felt that a particular diagram or point required amplification it was very hard to do this with a ‘rigid’ electronic presentation format. Equally, those lecturers who only used electronic media said that the major advantage was the way that complex diagrams could be built up piece by piece. In the past they had used overlaid transparencies, which had been unsatisfactory. In particular, being able to visualise and manipulate 3D structures was seen as a real advantage in chemistry.

Of course it is now possible to annotate electronic presentations through interactive whiteboards and related devices, and some lecturers have been piloting these in tutorials. Such an extension would go some way to providing the flexibility sought by lecturers in category 3.

Having observed some lectures from all categories, we noted greater confidence shown by lecturers in category 3 over category 1. This derives in the main from the fact that no matter how much preparation one does, under the current set-up at Bristol, there is usually only a maximum of 10 minutes between lectures to ‘fire up’ a presentation and to make sure all the links work as intended on the day on that machine. There was also a tendency for the pace of a lecture in category 1 to be ‘flat’ and to have no breaks, intended or otherwise. Lectures in the other categories were varied in pace and had natural breaks (cleaning the blackboard, moving around the lecture theatre etc.). Chemistry lectures will, by their nature, contain a lot of symbols, and in some cases mathematical material; this will be discussed further in the next section. Taking questionnaire and interview material together it would seem that there are indeed some differences between lecture categories, but that their impact is slight.

Table 2. Average questionnaire scores for lecturers in the three categories of preferred presentation medium, split into the three overarching teaching strands.

Physical

Questionnaire Scores %	1	2	3	4	5
Category 1	6	59	35	0	0
Category 2	11	61	28	0	0
Category 3	14	76	10	0	0

Inorganic

Questionnaire Scores %	1	2	3	4	5
Category 1	7	61	32	0	0
Category 2	15	75	10	0	0
Category 3	12	75	13	0	0

Organic

Questionnaire Scores %	1	2	3	4	5
Category 1	6	48	43	3	0
Category 2	6	58	35	1	0
Category 3	10	71	19	0	0

Comparisons across lecture groupings

In Table 2 the effect of the subject area of lecture material is presented, and again any differences are small, with a slight bias towards Physical and Inorganic over Organic in all categories. The feedback from students emphasised three important points regarding differences in subject area.

- Lectures covering organic and biological chemistry tend to involve writing structures of compounds, with an emphasis on synthetic pathways. In general, students found this material harder to write down or annotate, regardless of lecture method used, especially when stereochemistry is important.
- When synthetic pathways were explained in organic and biological chemistry lectures, students preferred those that used TOHP or chalk and talk, because the material was presented at a slower pace, allowing students to follow the material. Those organic and biological chemistry lectures using PowerPoint or other electronic media that were perceived as being ‘weaker’ were reported to contain too much material and were presented too quickly. In year 1 in particular, students reported that they were unfamiliar with writing structures, and found it hard to follow pathways.
- A major concern of students in Physical and Theoretical lectures was the need for consistency of symbols used across courses, and it was observed that some electronic presentations suffered from not being able to write symbols in a font that was consistent with other lectures.

Cannon (1988) has set out characteristics that appear to be related to effective lecturing, notably preparation (Falk and Dow, 1971; Brown, 1978), presentation (Brown, 1978; Brown and Atkins, 1988) and evaluation. Under presentation, Brown (1978) and Brown and Atkins (1988) describe in detail the perceived process by which students learn from lectures. The lecturer sends a message, which may be verbal, non-verbal, a gesture, or some other medium. Messages that are received by the students are filtered and stored temporarily in the short-term memory. If these messages cannot be rehearsed or cannot be transferred to the long-term memory after about thirty seconds, they are forgotten. In addition, attention will inevitably fluctuate throughout a fifty-minute lecture. After twenty minutes there is a marked decline in attention, followed by a peak just before the lecture ends (e.g. Johnstone and Percival, 1976). Of the lectures observed in each category, there were fewer deliberate or forced breaks in the presentation in category 1, and in the most extreme case there was a solid 50 minutes of information presented.

The work of Lowry (1999) focused on a first year environmental science course, and gave several possible reasons why there may be a marked improvement in student performance after the introduction of PowerPoint, but the most striking was the fact that the style of workshops were changed, and that these may have led to the improvement in cognition. Lowry made some excellent points concerning the use of electronic media and these have already been commented upon in this paper, yet there is no strong evidence from this study based on questionnaires, interviews and examination results, to support the assertion that electronic presentations produce deeper level learning that non-electronic ones. Indeed, if the very small differences between scores from the three categories are taken at face value they suggest the opposite is more likely to be true.

A secondary school perspective

It is useful to have a secondary school perspective on styles of teaching and to know what students of the future will have experienced. In the early days of IT in schools it was rare for teachers to make regular use of the technology in class teaching because of a lack of training and equipment. New Opportunities Fund (NOF) Training (1999-2003) with the aim of increasing the IT expertise of serving teachers (and librarians) to that of newly qualified teachers in 1999, has been successful in up-skilling most teachers (Preston, 2004). Electronic media in laboratories, classrooms and computer suites is increasingly available in schools and is being used in science lessons by "*teachers who are confident in the classroom use of modern technology*" (Ofsted, 2002).

It is now common for teachers at school to use PowerPoint and other packages to deliver lessons, particularly during practical science classes. Students are therefore very familiar with this format, however, the dominant form of delivery is still a chalk and talk approach.

Concluding remarks

There are clearly some essential characteristics for effective lecturing. Good preparation, from planning the course structure to being able to work the lights in the lecture theatre, is vital. Good presentation skills are important; students must be able to see and hear the relevant information. Ideally, the course must motivate and enthuse the students without overloading them. Ausubel (1968) has claimed that "*The most important single factor influencing learning is what the learner already knows. Ascertain this and teach them accordingly*", although the student's own motivation will also play a role.

Having analysed questionnaires, carried out interviews and also inspected examination results, we conclude that the method of delivery has no significant impact on learning outcomes. Questionnaires and interviews with students establish that the most important aspect of the lecture is clarity; there are good examples of clear lectures using each category of delivery inspected. In addition, there is a strong positive correlation between examination result and questionnaire score, where interviews suggest that the reason for poor examination results is that material has been hard to understand.

Brown (1978) argues that it does not matter whether copious handouts and examples, or just a few are given out; the important aspect is quality of example and handout. It does not matter whether lecturers prefer the 'classic' or the 'thesis,' or the 'comparative' or the 'sequential' or even the 'problem-centred' (Brown, 1978) style to lecturing; the important aspect is choosing a style that is appropriate for the particular course. It would appear that the effective lecturers are those who start at a point where the learners can comprehend and lead them step-by-step through the new material, and that that is far more important than the medium used. There are always improvements one can make to style and delivery of presentation, and any slight differences observed in this study between electronic and non-electronic presentations are most likely due to lecturers learning how to use the new tools effectively. However, the conclusion of this study is that the mode of presentation is not the key difference in learning outcomes.

Acknowledgments

DES would like to thank the HEA for a National Teaching Fellowship and TGH thanks Bristol ChemLabS for a School Teacher Fellowship. DES and TGH thank the HEA, Bristol ChemLabS and the School of Chemistry at Bristol, under whose auspices various elements of this research was carried out. We thank the staff and students at the School of Chemistry at Bristol University for their cooperation with this project. Bristol ChemLabS is a Higher Education Funding Council Centre for Excellence in Teaching and Learning in practical chemistry.

References

Ausubel D.P., (1968), *Educational psychology: A cognitive view*, New York, Holt, Rinehart and Winston.

Bartsch R.A.and Cobern K.M., (2003), Effectiveness of PowerPoint presentations in lectures, *Computers and Education*, **41**, 77-86.

Brown G., (1978), *Lecturing and explaining*, Methuen, London.

Brown G. and Atkins M., (1988), *Effective teaching in higher education*, London, Routledge.

Cannon R., (1988), *Lecturing*, Kensington (N.S.W.), Higher Education Research and Development Society of Australia.

Falk B. and Dow K.L., (1971), *The assessment of university teaching*, The Society for Research into Higher Education, London.

Johnstone A.H. and Percival F., (1976), Attention breaks in lectures, *Education in Chemistry,* **13**, 49-50.

Lowry R.B., (1999), Electronic presentation of lectures – effect upon student performance, *University Chemistry Education,* **8**, 18-21.

OFSTED (2002) ICT in schools: effect of government initiatives Secondary Science June 2002 HMI 715. Report from the Office of Her Majesty's Chief Inspector of Schools.

Preston C., (2004), Learning to use ICT in classrooms: teachers' and trainers' perspectives, Teacher Training Agency.

Szabo A. and Hastings N., (2000), Using IT in the undergraduate classroom: should we replace the blackboard with PowerPoint? *Computers and Education*, **35**, 175-187.

Walton A.J., (1972), *Lectures, tutorials and the like*, Oxford and Lancaster, MTP.

Formative peer and self feedback as a catalyst for change within science teaching

Simon Bedford and Serena Legg

Department of Chemistry, University of Bath, Bath, BA2 7AY
e-mail: S.B.Bedford@bath.ac.uk

Received 30 August 2006, accepted 9 January 2007

Abstract: Feedback to students is vital for effective learning; however, it is a relatively under-researched area in the UK. This study sought to use new and more effective methods of formative feedback to students within the context of Chemistry teaching in order to facilitate student learning. Emphasis was placed on the use of Student Directed Assessment, and in particular, the use of Student Self and Peer-Assessment. During semester 2 of the 2005-06 academic year, a cohort of some 100 Chemistry students and 33 Natural Sciences students attended a series of problem-based workshops designed to test Self and Peer Assessment methods. Results show that both Peer and Self Assessment were preferred over Tutor Assessment. Whilst Peer Assessment was viewed as helping to learn more on specific topics, Self Assessment was perceived as having a didactic value as they learned from their own mistakes. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 80-92]

Keywords: Feedback to students, Peer assessment, Self assessment, Tutor assessment, Formative assessment

Introduction

The number of students in Higher Education in the UK has greatly increased over the last decade. The present government has set a target to continue this growth, requiring a further 17,000 lecturers to be employed by the year 2010 to teach the extra students (Ratchford, 2006). In the current academic context, it is widely accepted that feedback is an essential component in the process of learning and in a student's development (Weaver, 2006), although it has not always been seen this way (Fritz, 2000). Unfortunately, despite the best efforts to retain a level of consistency in the quality and amount of feedback given to students, recent surveys carried out on students have highlighted their dissatisfaction with the feedback they receive.

The National Student Survey in 2005 provided a snapshot of one year within the Higher Education sector. The Survey showed that whilst most students were overwhelmingly satisfied with the quality of courses, there was a general dissatisfaction in many Higher Education institutions with the provision of assessment and feedback. Responses to the 'Assessment and Feedback' section of the survey gave 86 out of 128 (67%) participating institutions their lowest score.

At the University of Bath, dissatisfaction with feedback amongst students was highlighted by the Student Satisfaction Survey in 2003. A suggested cause of the problem was the increasing student-staff ratio, which has resulted in the decline of feedback to students (Macaskill, 2006). It is claimed that both teachers and students consider feedback to students an important process in aiding the learning and development of students' skills. When Business, and Art and Design students were asked if they thought "*that feedback is helpful to* *explain gaps in knowledge and understanding*" and "*constructive criticism is needed to know how to improve*" (Weaver, 2006), there was an overwhelming majority in strong agreement.

The Quality Assurance Agency for Higher Education stipulates the provision of effective and appropriate feedback within Higher Education. Craft (2001) and others identified feedback as one of the four key conditions for creative learning. Therefore, where there is a lack of effective feedback, the students' learning could be hindered. In response to this, the Higher Education Academy has produced a briefing paper within a project for 'Student Enhanced Learning through Effective Feedback' (Nicol, 2005). The aim of the project is to help practitioners who wish to improve feedback or be inspired with new ideas. They identified seven features of good feedback practice:

1. Facilitates the development of self-assessment (reflection) in learning.
2. Encourages teacher and peer dialogue around learning.
3. Helps clarify what good performance is (goals, criteria, standards expected).
4. Provides opportunities to close the gap between current and desired performance.
5. Delivers high quality information to students about their learning.
6. Encourages positive motivational beliefs and self-esteem
7. Provides information to teachers that can be used to shape the teaching.

These are all good ideals for feedback given to students. However, in practical terms, and with time pressures on staff, they can be merely ideals that direct the provision of feedback to a more effective path. This is recognised within this paper, and specific strategies are described that seek to improve current feedback mechanisms which are, for example, those that involve written or verbal communication between tutor and student, such as tutorials and essay transcripts. These methods of feedback are time-consuming, especially in institutions where the academic year is based on a modular system with multiple assessments per semester. There is a need to adapt current procedures, or experiment with new methods to provide for the increasing number of students who are ever more diverse. This should be done in a way that requires little extra time and commitment from the tutor, but also in a way that meets the needs of the students.

Research Methods

There are many methods of feedback, including 'group needs' led feedback, statement banks (Rust, 2001) and those implemented via computer software (Tasker, 2006) and the e-learning technologies (Price, 2006). Currently, within this Department, some of the methods for giving formative feedback to students are: tutorials (made up of ~ 6 students), problem workshops (comprising ~30 students), coursework response, and laboratory work response. Even with a wide variety of feedback opportunities, the aforementioned surveys show that students are dissatisfied with the feedback that they receive in the Department of Chemistry. This prompted the investigation into enhancing feedback processes and an exploration of mechanisms that could provide better feedback to students without placing further burdens onto academic staff. In our context, it was decided that student formative assessments in semester 2 should be explored as a possible solution to the feedback problem. One potential approach is the use of Student Directed Assessment, in particular via Peer and Self Assessment.

Peer Assessment

A University Engineering Department used the scheme of Peer Assessment within a compulsory second year module (Gibbs, 1999). Its aim was to counter the growing workload for the tutors due to larger numbers of students and to improve the feedback to students as a way to enhance their learning. The result of the new scheme was that the average marks of the students increased from 45% to 75%! The Peer Assessment workshops not only had a large impact on the average marks of the students, but were also found to teach students other key skills through marking others' work. The process allowed them to note errors of their own, and guided them to other ways in solving problems. This encouraged 'active student learning' and an element of Self Assessment. This method of Peer Assessment also provided prompt feedback to the students, which was received positively by the students.

In a similar way, research by Jordan (1999) noted that Student-Student Assessment and Feedback also promoted team effort, which encouraged even the more 'reticent' students to take part in the group work.

Self Assessment

Boud (1991) defined Self-Assessment as: "*The involvement of students in identifying standards and/or criteria to apply to their work and making judgements about the extent to which they have met these criteria and standards*." This method of assessment encourages students to identify their own strengths and/or weaknesses, as well as to request the kinds of feedback they would like to receive on their work before handing it in. They can also mark and evaluate their own work, making judgements directed by a clear mark/assessment scheme.

When a Self Assessment scheme was used (alongside peer assessment) in the English Department at St Andrews University (Mallett, 1995) the benefits were perceived as follows:

- The students learnt more about their own working style.
- The students became more self aware, and were encouraged to think more carefully about the criteria, aims and objectives of a module.
- The staff were able to give more directed feedback.
- The tutor learnt from the comments made by students how to tone comments made on future assignments.

Mallett (1995) stated in this case study that it would have been feasible to contribute the marks awarded from the results from the self assessed essays to the summative mark for the course. However, colleagues in the department did not share his view. Nevertheless, these benefits are notable and suggest that it would be useful to investigate the use of Self-Assessment further.

Zoller and co-workers have studied student self-assessment and self-grading of Higher-Order Cognitive Skills (HOCS) type chemistry examinations (Zoller et al., 1997, 1999). The authors concluded that potential exists for adequate student self-assessment, but reported a gap between the students' self-assessment (overestimation) and their teachers' assessment, which was attributed to the prevailing Lower-Order-Cognitive-Skills (LOCS) orientation in the practice of science/chemistry teaching.

All these studies demonstrate that the student is at the heart of Peer and Self-Assessment. Since these processes are relatively under-researched in Higher Education, we carried out this small scale investigation aiming to shed more light on the use of Peer and Self-Assessment in a Higher Education context, and particularly in science. Additional reasons for selecting Peer and Self Assessment for study are as follows:

- Supporting literature has shown that in other institutions Peer and Self-Assessment has been successful in providing an alternative method for feedback to students where there is pressure on staff time.
- Methods of Peer and Self-Assessment can be implemented and investigated in the time that is available for the project.
- Peer and Self-Assessment is likely to increase students' confidence. This is especially important, as the aim of the workshops is to increase the confidence of the students in using skills that they have learnt.

The Higher Education Academy (HEA) gives several suggestions on how to prepare Self and Peer-assessment activities (Orsmond, 2004). Some of these are:

- Give clear instructions to the students relating to all of the stages of the assessment process, e.g. mechanisms for disagreement and the extent the marks count. (Instructions should be given in writing and verbally).
- Give the students structured written schedules, so that the marking criteria are clear.
- Make modifications where necessary.
- Repeat exercises with the same student cohort.
- Ensure that students have a clear understanding of the feedback that they can expect to receive.

Method

The student cohort involved in the investigation of Peer and Self-Assessment, some 140 in total, were those taking first year Organic Chemistry modules in the academic year 2005-06. They were chosen the following reasons:

- As new to university study, they came with no preconceived ideas on assessment and feedback at this level.
- The foundation level of this subject topic permeates many different programmes of study.
- As first years, they also offer the potential for a longitudinal study in years 2 and 3.

The students were divided into four independent teaching groups for their workshops, which were timetabled across consecutive weeks; this allowed for consistency in the investigation. These timetabled workshops placed emphasis on improving skills in drawing reaction mechanisms, rather than the frequent format of using knowledge from lectures to answer problems. This in turn made the workshops ideal to investigate both the students' reactions to different methods of feedback, and particularly those points of the 'Good Feedback practice' mentioned earlier.

Taking the above suggestions from the HEA into consideration, two problem papers were designed. The questions on the first paper sought to guide the students through practice in drawing reaction mechanisms with the aim to help them to practise these important skills. It was also essential to allow enough time during the workshop for assessment and feedback *to* the students and also feedback to be received *from* the students about the process undertaken.

Four groups of students containing roughly equal numbers (~35) took part in the investigation (three groups of Chemistry students and one group of Natural Sciences students); they each had a different form of assessment method as follows.

Group 1 – Peer assessment workshop
Group 2 – Control workshop
Group 3 – Tutor assessed workshop
Group 4 – (Natural Sciences students) – Self assessed workshop

The control workshop was included for comparison. This was run as a 'normal' workshop in that, as in general departmental workshops, it did not involve any aspect of feedback other than a tutor being available to answer questions, and to go over general group problems on the board.

The Tutor-assessed workshop was given so that comparisons could be made between the groups of the level of satisfaction with the feedback they received. The students received feedback on their answer sheets from the tutor, which they received a week later. They also had the opportunity to look over the mark scheme at this point. This allowed us to investigate how much the students valued a fast feedback response, as received by groups 1 and 4.

Groups 1 and 4 received Peer and Self Assessment feedback

The marking criteria given to students were within the answer sheets with the comprehensive answers to the questions in the workshop. The sheets also had clear directions as to how the worksheets should be marked, particularly in week 1, as this was the first week of the new feedback methods. Groups 1 and 4 received these sheets within the workshop to mark either each other's work, or their own (depending on the group). Group 2 received the answer sheet at the end of the workshop and Group 3 received the answer sheet at the beginning of the following session, accompanied by their annotated workshop sheet. At all stages, verbal instructions were given to the students on how to complete their task.

A fast and efficient way of receiving feedback from the students was required because of the limited time at the end of the workshop, and the need to obtain feedback from all of the students present. From past experience, questionnaires were not considered to be the best option, as the return rate is rarely above 50%, therefore it was decided that students would be asked to raise their hands if they agreed with statements that were shown on the Overhead projector. An assistant in the workshop recorded the number in agreement.

The workshops

There were eight 50 minute workshops given in total, two each to groups 1, 2, 3 and 4. The general structure of the workshops in week 1 was as follows**:**

- Introduction of those taking the workshop.
- Initial warm-up activity using the overhead projector (10 minutes) with a view to encourage the students to settle into the workshop, and to give them confidence in their chemistry knowledge before being given the workshop sheet. Answers to these activities were drawn on the blackboard by students in the group, and corrections were explained to the group as a whole by the tutor.
- Workshop sheets were handed out to the students and tutors were available to assist with any questions. (~25 minutes)
- Additional problems were given to students on the blackboard for those who had finished early.
- For the relevant workshops, answer sheets were given to the students 15 minutes before the end of the workshop.
- In the last 5 minutes of the relevant workshops, feedback was collected from the students.

The structure of the week 2 workshops was very similar, apart from the absence of introductions of the tutors (on the basis that by then the students knew who we were) and the activities given to the students to complete during the workshop were different.

The workshop sheets varied in character from week 1 to week 2. The workshop in week 1 was skills-based, involving practice in drawing reaction mechanisms. The workshop in week 2 was more content driven, which meant that in order to succeed the students had to have attended the lectures. The initial activity for week 2 provided the students with a helpful starting point for the first questions on the worksheet, and was designed so students would have a deeper understanding of why reactions occur as they do.

Results

Peer assessment workshops

In view of the fact that little peer assessment is carried out in the Chemistry Department, it was pleasing to see from the responses that students liked the Peer Assessment workshop method. This can be seen from 68% of students being happy for the process to continue in future workshops, and no student thought that the method should never be used again. All the students agreed that they liked the fast feedback response of peer marking. The number of those who found the peer marking method helpful to their own learning increased markedly from week 1 to week 2 by 30% and can be explained by the increased confidence or growing familiarity with the procedure (Raaheim, 1991).

As seen earlier, the workshop in week 2 was more content driven and meant that in order to succeed, the students had to have attended the lectures. This could also explain why there was a 30% increase in agreement for the statement: "*You found it helpful to look over someone else's work, in helping your own understanding of the topics covered today"*. The workshop in week 2 required that students had an understanding of the work and perhaps stretched them further than the workshop in week 1, which was skills-based. It would also seem that 100% of students (28/28 and 25/25) were positive about the fast feedback procedure over both weeks. Nonetheless, it is important to note that almost half of the students (50% in week 1 and 44% in week 2) were not happy with a fellow student marking their work. This could be due to the lack of anonymity in the marking. This lack of anonymity has ethical implications, which will be discussed later. Also, 18% of students in week 1 felt that they would rather have a tutor annotate their work.

As it can be seen from the results in Tables 1 and 2 below, students became more confident in using the peer assessment method, as eight (32% of attendees) students felt they were more confident in marking work than the previous week (statement 1).

For both weeks 1 and 2, roughly half of the students felt that they found peer comments helpful (statement 3) although this method of giving feedback from model answers leaves little room for students to add subjective comments.

Table 1. Feedback results from Peer Assessment workshop 1.

	Statement	Students who agreed /28	% of students who agreed
1	You would rather mark your own work than have a friend mark it.	14	50%
2	You found it helpful to your own learning to look over someone else's work and mark it.	4	14%
3	You found any comments on your work helpful	14	50%
4	You would like this method to occur in every workshop.	1	7%
5	You would prefer to have your work commented on by a tutor and returned the following week.	5	18%
6	You like the fast feedback response of peer marking	28	100%
7	You did not like marking someone else's work.	4	14%
8	You did not like another student marking your work.	8	29%

Table 2. Feedback results from Peer Assessment workshop 2.

	Statement	Students who agreed /25	% of students who agreed
1	You found the process of marking work more straightforward this week after having done it last week.	8	32%
2	You would rather mark your own work than have a friend mark it.	11	44%
3	You found it helpful to look over someone else's work, in helping your own understanding of the topics covered today.	11	44%
4	You found any comments on your work helpful.	11	44%
5	After using peer assessment again today you feel more positive about the process than you did last week.	6	24%
6	After using peer assessment again today you would rather that it was not used again in a workshop	0	0
7	After using peer assessment again today, you would be happy for it to continue in some workshops.	17	68%

Tutor marked workshops

Students were asked to respond to the statements relating to the tutor's comments on their worksheets in workshop 2 (Table 3). Although thirty students were present at the workshop, only twenty-three of them responded to the statements, as some had attended the previous workshop. All the students who responded, agreed with the statement that they found the feedback comments easy to understand (statement 1) and 70% of students said this feedback helped them to see where they could improve (statement 2). This is very encouraging, as an aim of good feedback is to provide opportunities to "*close the gap between current and desired performance*" (Nicol, 2005). More than half of the students felt that they would like feedback during the workshop (statement 5). This reflects the opinion of those students who were involved in the peer assessment workshops, where 100% of the students liked the 'fast feedback' response.

Table 3. Feedback results from students in Tutor Assessment workshop 2.

	Statement	Number of students /23	% of students who agreed
1	You found the feedback comments easy to understand.	23	100%
2	From the comments on the sheet you see how you can improve your technique	16	70%
3	You did not pay much attention to the comments made on the sheet.	2	9%
4	You had forgotten the workshop from last week	0	0
5	You would prefer feedback during the mechanism workshop.	13	57%
6	You would have preferred to mark your own workshop from an answer sheet.	6	26%
7	You would like this method of tutor marking to occur in every workshop.	12	52%
8	You do not like this method of feedback and would rather that it was not continued within workshops.	7	30%

This style of workshop was the one example where 30% of the students agreed that they did not like the tutor marked method of feedback and would rather it did not happen again (statement 8). This finding reflects a preference for an alternative to the 'traditional' feedback approach.

Self-Assessment workshops

Students were asked to respond to the statements at the end of each workshop (Tables 4 and 5). Table 4 reveals that 100% of students agreed that they liked the fast feedback mechanism in the Self-Assessment. These results echo students' views in the Peer Assessment workshops, where several students felt that they would abstain rather than agree with statement 7 in workshop 1. It would seem that the system of receiving fast feedback appealed to them more than having more frequent workshops in the style of Self Assessment.

These views altered greatly in week 2 (Table 5) where 100% of the students were happy to receive feedback from Self-Assessment methods in future workshops (statement 4) and all found the process of marking their own work more straightforward, having done it the week before (statement 2).

When asked, only three students said that they had completed the marking without much thought over the two weeks. If this was an honest response from all participating students, this is a valuable indication that the students took the process seriously and thought about the marking that they were doing. Here again, students seemed very reluctant to having their workshop sheet annotated by a tutor. This has been the case in all of the previous workshops.

Table 4. Feedback results from Self-Assessed workshop 1.

	Statement	Number of students who agreed/33	% of students who agreed
1	You found it helpful to receive the answers during the workshop.	33	100%
2	You found it helpful to mark your own work from the answers.	33	100%
3	You would prefer the tutor to mark your work and hand it back with comments.	0	0
4	You learnt from your mistakes by going over the answers yourself.	33	100%
5	You liked receiving fast feedback from the self-assessment structure to this workshop.	33	100%
6	You did not pay much attention to marking your own work and went quickly through the marking process without much thought.	3	9%
7	You like the method of self-assessment and would like it to occur more frequently in workshops.	30	90%
8	You would like this method to occur in every workshop.	9	27%
9	You do not like this method of feedback and would rather that it was not continued within workshops	0	0

Table 5. Feedback results from Self-Assessed workshop 2.

	Statement	Number of students who agreed/26	% of students who agreed
1	You were happier with the self-marking process this week than you were last week?	0	0
2	You found it more straightforward to look through your own work and mark it because you had followed this procedure last week.	26	100%
3	You found it helpful to receive the answers during the workshop this week.	26	100%
4	You found it helpful to mark your own work from the answers.	26	100%
5	You would prefer to complete the workshop first, before getting the answers.	8	31%
6	You would prefer the tutor to mark your work and hand it back with comments.	0	0
7	You did not pay much attention to marking your own work and went through the marking process without much thought.	0	0
8	After doing this style of workshop again you would like this method to be used more often in workshops.	26	100%
9	After doing this style of workshop today you would like this method to be used in every workshop.	16	62%
10	You do not like this style of workshop and would rather that it did not continue within workshops.	0	0

Control workshop

The feedback was collected from students in the second control workshops by the method described earlier. As can be seen in Table 6, 100% of the students agreed with the statement that they were "*content receiving the answer sheet to look through in [their] own time*" (statement 1). The majority of students also seemed satisfied with the feedback they had received during the workshop, though some (23%) felt that they would have liked to receive more, and a further 13% claimed that the level of feedback was inadequate.

Table 6. Feedback results from the Control workshop.

	Statement	**Number of students/30**	**% of students who agreed**
1	You have been content receiving the answer sheet to look through in your own time.	30	100%
2	You would have found it helpful to go through the answer sheet within the workshop.	8	27%
3	You feel the feedback you received within the workshops over the past 2 weeks has been adequate.	23	77%
4	You think that the feedback given in the workshops over the past 2 weeks has been inadequate.	4	13%
5	You like the idea of swapping sheets with another student to be marked to receive feedback.	0	0

With hindsight it would seem that the way question 5 was phrased may have confused students, as it produced no response, therefore making it impossible to evaluate the result of this statement. Perhaps a simpler statement such as ‘you like peer-marking and receiving feedback from another student’ may elicit more responses in the future?

As well as the workshops, interviews were carried out with two students, one each from the peer and tutor assessment workshops. The aim of these interviews was to clarify the feedback received from students, and also to analyse more deeply student perception of feedback. As the number of responses is very small, they should be treated with caution. The main points from the interviews were as follows:

(Student A attended the Peer Assessment workshop and Student B attended the tutor assessed workshop)

- Both students were unsure of the feedback that they could expect from the University and neither of the students had read the student handbook for their course. They felt that this could reflect a relatively accurate picture for the majority of students on their course.
- Both considered their respective workshops to have provided adequate feedback. Student A felt that more time to annotate their peer’s work would have enabled the students to learn more from the procedure. He also commented that although feedback from a tutor would have been better, peer feedback was better than none at all.
- Both students agreed that they expected less feedback at University compared to school. Student A acknowledged that “*the focus is supposed to be more on you getting the work done yourself*”.

Discussion

The results from the feedback received from students highlight the fact that:

Tutor comments were not preferred to self or peer-assessment by the students.

A perhaps surprising finding from the data collected was that students overwhelmingly felt that they would not like to have a tutor annotate their work individually. This was especially clear for several reasons:

The observed reaction of the students in the tutor-assessed workshop could have been due to the fact that they had to hand in their worksheets at the end. Those students who had not completed the work would be reluctant to hand it in.

All students in the Self and Peer Assessment workshops agreed that they would not like to have a tutor mark their work, but would instead prefer receiving an answer sheet and marking their own work. The students marked the worksheets during the workshop, which meant that they also received fast feedback from the peer and self assessment mechanisms.

Another reason could be that the thought of handing work in to be commented on unsettled the students, as there seemed to be a belief that tutors will think less of them if they did not do well. This perception concerning the tutor-assessed workshop may be linked to the ‘fear of failing’ as described by Stiggins (1999), who goes on to say that “*the trick is to help students understand that failure holds the seeds of later success.*”

Results from later exam feedback tutorials came as a surprise when they revealed that feedback from the tutor was deemed to be very valuable. The tutor’s feedback in these tutorials was generally in the form of a verbal feedback with no peer assessment involved. Self-assessment may have occurred, but this would have been at the discretion of the individual tutors and the way that they chose to run their feedback tutorials. The interview analysis with our students indicated that they would rate tutor feedback very highly. When probed further, it seemed that there is a difference in the minds of the students between having annotated comments from a tutor and having face-to-face feedback with a tutor. Students claim they would prefer to have either feedback from a peer or the opportunity to go through their own work with a view to self critique and learn from their mistakes, rather than receive written feedback on their work from a tutor.

Thus the ranking in order of preference for written feedback from the results seems to be *Peer Assessment, followed by Self Assessment, followed by Tutor Assessment*. This observation is largely based on how the students would like the procedure to run in future workshops. This is a very interesting result as it suggests that there is great value in exploring Peer and Self-Assessment as a method of feedback to students and in doing so to move away from the more conventional use of tutor written feedback.

Fast feedback appreciated

Students who received feedback in the workshop (those in the peer and self assessment workshops) were unanimous in agreement that fast feedback was good. Over half of the students in the tutor marked workshop would have preferred faster feedback. With regards to the tutor marked workshops, students were concerned with the fact that they had received feedback several weeks after the workshop. The University Quality Assurance specification advises that feedback is given within 3 weeks (QA, 2002), therefore any feedback given within this time period is within that of the Quality Assurance specifications. However, this is perhaps too long for students to wait in order to be able to implement what they have learnt in the workshops and feedback process. A similar issue with laboratory reports came up during interviews. These are produced weekly, but marking and feedback can take several weeks and the feedback comments were often of little use at this point for improvement.

The promptness of feedback is one of the conditions for the success of a student centred approach adopted by Gibbs (1999) in assessment. A Student Centred learning approach is where the focus is on the student's learning, as opposed to being on the teacher transmitting knowledge. Harden (2000) says that the focus is on student learning and "*what students do to achieve this, rather than what the teacher does"*. Yorke (2003) notes from this study of Peer Assessment with Engineering students that the "*improvement in student's end-of-course outcomes was very marked"* and relates the success to fast feedback that the students received.

Self Assessment – students learnt from their mistakes

It was interesting to see that all students taking part in the Self-Assessment workshop agreed that they learnt from their mistakes through marking their own work from the answer sheet. This is very pleasing, as a good feedback practice should "*provide opportunities to close the gap between current and desired performance*" (Nicol, 2005). However, the feedback would seem to be more helpful to the students if they had more time to complete the workshop sheet and answers. This is reflected by 31% of students in week 2 agreeing that they would "*prefer to complete the workshop first, before getting the answers"*.

In an interview, Student A's responses about this workshop highlighted the need to increase the time allocated to 'Peer-Assessment' to increase its potential to provide good feedback to students. Boud (1991) stressed the need to allow time for students to assimilate the process of giving and receiving feedback.

As Peer Assessment is relatively unknown by students and tutors in the Chemistry Department, an investment of time is required to both prepare and undertake assessment in this way. Our results show that time spent this way has paid off, with 30% of students increasing in confidence with giving and using peer feedback. However, time may not resolve the fact the some students are not happy with marking other student's work, or having their work marked by peers. An issue that arose from the Peer-Assessment workshops was that of ethics. Indeed, students did not appear completely satisfied with exchanging worksheets in a way that was not anonymous. Because the workshop was run in this way, the possibility of marking work objectively was greatly reduced and may have hindered the learning process. An anonymous process would reduce this element and improve student learning by removing potential awkwardness from students knowing who their marker was and whom they were marking.

Anonymity would be crucial if the process of Peer-Assessment were to move beyond giving feedback in a formative assessment context to that of summative assessment. This would avoid bias influenced by factors such as: over-marking friends, gender bias (as noted in Langan, 2005) and dominant individuals receiving higher marks.

A better-designed study would have to be developed in order to dispel concerns over the validity and fairness of peer marked work compared to tutor assessment (Falchikov, 1989).

Conclusion

The study of peer- and self-assessment workshops has shown convincingly that students valued this fast feedback approach and that they appreciated the quality of feedback received from their peers or from the Self Assessment exercise. Interestingly, the study also revealed that students viewed feedback from Peer- and Self-Assessment more favourably than tutor feedback. Thus the dual aim of giving quality feedback to students, but without adding more time pressures on to tutors was achieved. This will lead to greater enhancement of feedback mechanisms within the programmes of study offered by the Department of Chemistry at the University of Bath. It is also hoped that in the near future more studies will follow with a view to explore further the potential of Self and Peer Assessment in order to improve the quality of feedback to students throughout the programmes of study.

References

Boud D., (1991), Implementing student self-assessment, *HERDSA Green Guide, 2nd ed.,* Sydney, Higher Education Research and Development Society of Australasia.

Craft A., (2001), *Creativity in education*, ISBN 0826448631 and Craft A., (2005) Continuum Higher Education Funding Council Press release *'Most Students overwhelmingly satisfied with the quality of courses'* Accessed online 15/03/06, http://www.hefce.ac.uk/news/hefce/2005/nss3.asp .

Falchikov N. and Boud D., (1989), Student self assessment in higher education: a meta-analysis, *Review of Educational Research*, **59**, 395-430.

Fritz C., Morris P. and Bjork R., (2000), When further learning fails: stability and change following repeated presentation of text, *British Journal of Psychology*, **91**, 493-511.

Gibbs G., (1999), Using assessment strategically to change the way students learn in assessment matters in higher education, *The Society for Research into Higher Education,* 41-53.

Harden, R.M., and Crosby J (2000), AMEE Guide No 20: The good teacher is more than a lecturer-the twelve roles of the teacher, *Medical Teacher*, **22**, 334-347.

Jordan S., (1999), Self assessment and peer assessment, in assessment matters in higher education, *The Society for Research into Higher Education*, 172-182.

Langan M.A., (2005), Peer assessment of oral presentations: effects of student gender, university affiliation and participation in the development of assessment criteria, *Assessment and Evaluation in Higher Education,* **30**, 21-34.

Macaskill A., (2006), IMPACT article, Feedback is unacceptable, 6th March 2006, edition 3.

Mallett P., (1995), Self- and peer assessment of written work in English literature, *Case study for Higher Education Academy*.

Nicol D., (2005), Rethinking formative assessment in HE: a theoretical model and seven principles of good feedback practice, *SENLEF briefing paper for Higher Education Academy*.

Orsmond P., (2004), Self- and peer- assessment. Guidance on practice in the biosciences, *The Higher Education Academy,* 41-43.

Price G.J., (2006), Computer aided assessment and feedback: can we enhance students' early experience at university?, *New Directions in Teaching Physical Sciences*, **2**, 35-40.

Raaheim K., (1991), Helping students to learn: teaching, counselling, research, The Society of Research into Higher Education and Open University Press.

Ratchford J. (2006), Higher education: universities as employers, accessed online 04/04/06, http://www.prospects.ac.uk/cms/ShowPage/Home_page/Explore_job_sectors/Higher_education/Universities_as_employers/p!eXfedm.

Rust C., (2001), A briefing on assessment of large groups; statement banks, *LTSN Generic Centre.*

Stiggins R, (1999) Assessment, student confidence, and school success, *Kappan Professional*,**81**, 191-198.

Tasker. R., 21st University of Western Sydney, March 2006, informal interview via email.

Weaver M.R., (2006), Do students value feedback? Student's perceptions of tutor's written responses, *Assessment and Evaluation in Higher Education*, **31**, 379–394.

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

Zoller U., Tsaparlis G., Fatsow M. and Lubezky A., (1997), Student self-assessment of higher-order cognitive skills in college science teaching, *Journal of College Science Teaching,* **27**, 99-101.

Zoller U., Fatsow M., Lubezky A. and Tsaparlis G., (1999), Students' self-assessment in chemistry examinations requiring higher- and lower-order cognitive skills, *Journal of Chemical Education*, **76,** 112-113.

Perspective

Strengthening conceptual connections in introductory chemistry courses

George M. Bodner
Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA
e-mail: gmbodner@purdue.edu

Received 5 November 2004, accepted 8 December 2006

Abstract: Both research on learning across a wide range of disciplines and common theories of learning recognize the importance of learning through making connections between new concepts to which one is exposed and existing cognitive structures or schemata. This paper considers examples of underappreciated cognitive connections that our experience has shown can facilitate students' learning of chemistry in the introductory course. The first deals with the question of whether the 'common-ion effect' is limited to discussions of solubility product equilibrium, as many textbooks seem to indicate. The second example questions why certain traditional approaches to teaching the chemistry of conjugate oxidizing agents and reducing agents are not applied to discussions of the chemistry of conjugate Brønsted acids and bases. [*Chem. Educ. Res. Pract.,* 2007, **8** (1), 93-100]

Keywords: Brønsted acids and bases, oxidation-reduction reactions, conjugate acids/bases, conjugate oxidizing/reducing agents, common-ion effect

Introduction

When I entered the field of chemical education, one of the dominant perspectives was that of Piaget (Herron, 1975; 1978; Good et al., 1979), which focused on stages of intellectual development and questioned whether it was possible for students to learn certain ideas or concepts until they had reached an appropriate level of epistemic development. In recent years, there has been a growing appreciation of the view of learning proposed by Vygotsky (1986) and others, which focuses on learning through interactions with different people, objects, events, experiences, and contexts. Both of these perspectives on learning are consistent with theoretical models proposed by Ausubel (Ausubel et al., 1978) and Gagné and White (1978) that emphasize the importance of learning through making connections between the new concepts to which one is exposed and existing cognitive structures or schemata.

The importance of cognitive connections in learning has been invoked across a wide range of disciplines, from art (Koroscik, 1996) to physics (Robertson, 1990). Major and Palmer (2001) argued that cognitive connections play an important role in problem-based learning, and Mastropieri and Scruggs (1996) have argued that the failure to make cognitive connections between already known and to-be-learned information was a primary characteristic of students referred to special education. Cross (1999) has gone so far as to assert that "learning is about making connections".

This paper considers examples of seldom-appreciated connections that I have found can facilitate students' learning of chemistry. One of these questions the conventional wisdom that the common-ion effect is limited to solubility product equilibria. The other revolves around the question: considering the ubiquitous presence of tables of redox half-reactions in introductory chemistry textbooks, why don't analogous tables appear in discussions of Brønsted acids and bases? As evidence for the power of these tables, the reader might consider the following question before continuing with this paper: Would you expect the following acid-base reaction to proceed as written? I.e., would you expect the overall equilibrium constant for this reaction to be larger than 1?

$$HSO_4^-(aq) + H_2PO_4^-(aq) \rightleftharpoons SO_4^{2-}(aq) + H_3PO_4(aq)$$

The common-ion effect

The shelves that line one wall of my office complex contain more than 200 introductory general chemistry textbooks, enough books to fill more than 9 m of shelf space. It doesn't matter which book one pulls from the shelf, the index at the back contains an entry on the 'common-ion effect'. As might be expected, many of these texts use the same example, the decrease in the solubility of AgCl in a solution that contains either the Ag^+ or Cl^- ion.

$$AgCl(s) \overset{H_2O}{\rightleftharpoons} Ag^+(aq) + Cl^-(aq)$$

$$NaCl(s) \overset{H_2O}{\rightleftharpoons} Na^+(aq) + Cl^-(aq)$$

Perhaps it is not surprising to find the common-ion effect defined as the "*decrease in solubility of an ionic salt, i.e., one that dissociates in solution into its ions, caused by the presence in solution of another solute that contains one of the same ions as the salt*" (Lagasse et al., 2000).

If one agrees with Ausubel (Ausubel et al., 1978) that "*The most important thing influencing learning is what the learner already knows*", it is somewhat disconcerting to note that relatively few introductory texts note that a common-ion effect also occurs in buffer solutions. Consider a traditional 0.10 *M* CH_3CO_2H/0.10 *M* CH_3CO_2Na buffer, for example.

$$CH_3CO_2H(aq) + H_2O(l) \rightleftharpoons H_3O^+(aq) + CH_3CO_2^-(aq)$$

$$CH_3CO_2Na(aq) \overset{H_2O}{\rightleftharpoons} Na^+(aq) + CH_3CO_2^-(aq)$$

Many, if not most, instructors invoke LeChátelier's principle to explain why the addition of sodium acetate raises the pH of the solution by decreasing the extent to which acetic acid dissociates. But it is rare to find either instructors or textbooks that explicitly make the connection between the common-ion effect in this buffer and the common-ion effect in their discussion of the solubility of silver chloride. It is interesting to note that one of the places where this connection is made is the Wikipedia (2006), which uses the example of an acetic acid/sodium acetate buffer solution to introduce the idea of the common-ion effect.

I have no objection to invoking LeChátelier's principle in discussions of buffer solutions. If one believes the concept of a common-ion effect is important for students to learn, however, this might be better achieved by noting that it is the presence of a common ion that inhibits both the dissociation of a weak acid in a buffer solution and the dissociation of an 'insoluble' salt in a solution that contains a source of a common ion.

Indeed, if one accepts the notion that making connections between analogous situations can facilitate learning, one might conclude that the buffer example cited above is not the only example of a common-ion effect one might invoke. Consider one of the first examples students encounter when they are exposed to discussions of equilibria of aqueous solutions of weak acids such as acetic acid.

$$CH_3CO_2H(aq) + H_2O(l) \rightleftharpoons H_3O^+(aq) + CH_3CO_2^-(aq)$$
$$2\ H_2O(l) \rightleftharpoons H_3O^+(aq) + OH^-(aq)$$

The total concentration of the H_3O^+ ion in a 0.10 *M* CH_3CO_2H solution from the dissociation of both the weak acid and water is presumed to be 0.0013 *M*. The contribution to the total concentration of the H_3O^+ ion from the dissociation of water, however, must be equal to the concentration of the OH^- ion from the dissociation of water. If the pH of this solution is 2.9, this means that the contribution to the concentration of the H_3O^+ and OH^- ions from the dissociation of water has been decreased by four orders of magnitude.

0.10 *M* CH_3CO_2H (pH 2.9): $\left[H_3O^+\right]_w = \left[OH^-\right]_w = 7.7\times10^{-12}\ M$

This is an obvious example of LeChátelier's principle, but it also another example of a common-ion effect. The addition of a second source of the H_3O^+ ion leads to a significant decrease in concentration of this ion from the original source of this ion.

Tables of oxidation-reduction half reactions

It is difficult to trace the history of the use of tables of electromotive force back to determine the date at which they first appeared in chemistry textbooks. A table of the 'electromotive series' can be found, however, as early as 1913, in W. A. Noyes's *A Textbook of Chemistry* (Noyes, 1913). According to a footnote in Noyes's text, this table was based on data reported by Wilh. Palmaer in Nernst's Festschrift in 1907.

By 1925, tables of electromotive force had sufficient explanatory and predictive power in the minds of textbook authors that Deming (1925) began his chapter on electrochemistry as follows: "*The electrochemical series, in its simplest form, is a list of oxidizable substances (metals). Read from below upward, it gives the order of increasing ease of oxidation. Metals nearest the top are the most readily oxidized, that is, they part most readily with electrons to form cations ...*" Deming then presented a table written in the form of oxidation half-reactions accompanied by the corresponding half-cell potentials in volts, as shown in Table 1. On each side of this table, Deming placed a label with an arrow pointing up that was accompanied by the caption: "*Order of increasing ease of oxidation or increasing activity as reducing agents.*"

Table 1. Deming's table of 'oxidizable' and 'reducible' substances.

Oxidizable substances (Reducing agents)		*Reducible substances (Oxidizing agents)*	*Potential (in volts)*
Li	$\rightleftharpoons$	$Li^+ + e^-$	+2.96
K	$\rightleftharpoons$	$K^+ + e^-$	+2.92
Na	$\rightleftharpoons$	$Na^+ + e^-$	+2.72
Mg	$\rightleftharpoons$	$Mg^{++} + 2e^-$	+1.55
Al	$\rightleftharpoons$	$Al^{+++} + 3e^-$	+1.35
Zn	$\rightleftharpoons$	$Zn^{++} + 2e^-$	+0.75
Fe	$\rightleftharpoons$	$Fe^{++} + 2e^-$	+0.45
Cd	$\rightleftharpoons$	$Cd^{++} + 2e^-$	+0.40
Ni	$\rightleftharpoons$	$Ni^{++} + 2e^-$	+0.21

By 1944, introductory chemistry textbooks often linked reducing agents with oxidizing agents in tables that showed the relative strengths of the reducing agents as proceeding from 'strong' to 'weak' as one went down the column of half-reactions, while the corresponding oxidizing agents went from 'weak' to 'strong' (Timm, 1944). Tables of the electrochemical series had sufficient explanatory power that many chemistry classrooms once had a table of oxidation half-cell potentials prominently posted on one of the walls of the room. With time, the convention for half-cell potentials was changed, so that these tables eventually appeared as lists of reduction half-cell reactions.

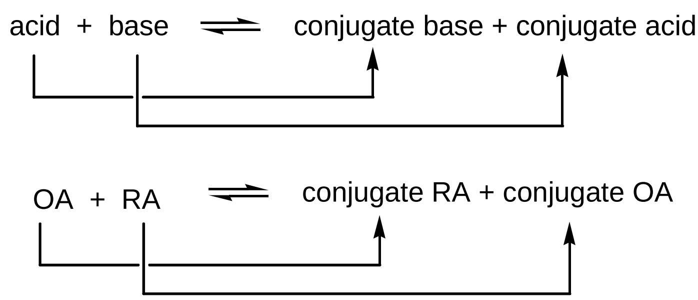
The first edition of the textbook by Nebergall and Schmidt (1957), which eventually went through ten editions and sold more than a million of copies, noted that the electromotive series of the elements helps us: (1) identify metals that are most easily oxidized and are therefore good reducing agents; (2) identify metal ions and nonmetals in their elemental states that are good oxidizing agents; and (3) predict which redox reactions should occur.

It is only recently that discussions of the electrochemical series have made explicit use of the term *conjugate* to link a given oxidizing agent to the corresponding reducing agent in a given half-reaction. Thus, whereas the notion that the relative strength of a given reducing agent is linked or coupled to the relative strength of the corresponding oxidizing agent has long been a feature of introductory chemistry textbooks, it is only in the last few decades that one finds explicit mention of the fact that strong reducing agents (such as Na) have weak conjugate oxidizing agents (such as the Na^+ ion), whereas strong oxidizing agents (such as F_2) are linked by the reduction half-reaction to a weak conjugate reducing agent (such as the F^- ion).[1]

Brønsted acid-base reactions

The ubiquitous presence of tables of redox half-reactions in introductory chemistry textbooks raises an interesting question: why don't analogous tables routinely appear in discussions of Brønsted acids and bases? The philosophical and conceptual bases of the phenomena are similar, as can be seen from the descriptions of similarities between the two systems given in below.

1. Both systems can be described in terms of coupled or linked pairs of reagents, i.e., conjugate pairs.



2. Both systems involve the transfer of a particle. In one case, the transfer of a proton.

$$HSO_4^-(aq) + H_2PO_4^-(aq) \rightleftharpoons SO_4^{2-}(aq) + H_3PO_4(aq)$$

In the other case, the transfer of one or more electrons,

$$2\ Ag^+(aq) + Cu(s) \rightleftharpoons 2\ Ag(s) + Cu^{2+}(aq)$$

or one or more atoms.

$$CO_2(g) + H_2(g) \rightleftharpoons CO(g) + H_2O(g)$$

3. In both systems there is a relationship between the relative strengths of the two components of the half-reaction: a strong oxidizing agent is linked to a weak reducing agent; a strong acid is linked to a weak conjugate base.

[1] I have noticed that it is the rare individual, indeed, who enters one of his chemistry courses who has not been asked 'to conjugate' a verb. I therefore find it interesting to note that it is even rarer to find anyone in these courses who has been told by one of their language teachers that 'to conjugate' means 'to link or couple'.

4. In both systems, one can predict whether a reaction should occur by asking: On which side of the equation do the stronger agents appear? The stronger of a pair of reducing agents and the stronger of a pair of oxidizing agents should react to form the weaker reducing agent and the weaker oxidizing agent.

$2\,Ag^{+}(aq)$	+	$Cu(s)$	⇌	$2\,Ag(s)$	+	$Cu^{2+}(aq)$
stronger oxidizing agent		*stronger reducing agent*		*weaker reducing agent*		*weaker oxidizing agent*

In a similar manner, the stronger of a pair of Brønsted acids and the stronger of a pair of Brønsted bases should react to form the weaker Brønsted acid and the weaker Brønsted base.

$HSO_4^{-}(aq)$	+	$H_2PO_4^{-}(aq)$	⇌	$SO_4^{2-}(aq)$	+	$H_3PO_4(aq)$
stronger acid		*stronger base*		*weaker base*		*weaker acid*

Thus, in theory, a table of Brønsted acid-base reactions has the potential to serve functions that are analogous to those filled by tables of redox half-reactions. Paraphrasing the words of Nebergall and Schmidt (1957), such a table can help us: (1) identify substances that lose a proton with relative ease, which are therefore good acids; (2) identify compounds that have a relatively high affinity for a proton and are therefore good bases; and (3) predict acid-base reactions that should occur.

Conclusion

For some time, I have had considerable success convincing the students in both my introductory chemistry courses and the physical chemistry course I teach for students from the life sciences of the benefit of recognizing an explicit connection between the concepts of conjugate acid/base pairs and conjugate oxidizing/reducing agent pairs. In the absence of a controlled experiment, or even a quasi-experimental design, I merely report an increase in the percentage of students who seem to be able to handle questions such as the following, which was answered correctly by 95% of the students on a recent introductory chemistry exam.

$NaHCO_3$ can be used to neutralize strong bases, such as NaOH. What conclusion can be drawn from the fact that the following acid-base reaction proceeds to the right as written?

$HCO_3^{-}(aq) + OH^{-}(aq) \rightleftharpoons CO_3^{2-}(aq) + H_2O(l)$

(a) HCO_3^{-} is a stronger acid than H_2O
(b) HCO_3^{-} is a stronger base than CO_3^{2-}
(c) HCO_3^{-} is a stronger base than OH^{-}
(d) CO_3^{2-} is a stronger base than OH^{-}
(e) H_2O is a stronger acid than HCO_3^{-}

Table 2. Relative strengths of typical Brønsted acids and bases.

	acid	⇌	H^+ + conjugate base		K_a (at 25°C)	
Best	HI	⇌	$H^+ + I^-$		3×10^9	
Brønsted	$HClO_4$	⇌	$H^+ + ClO_4^-$		1×10^8	
Acids	HCl	⇌	$H^+ + Cl^-$		1×10^6	
	H_2SO_4	⇌	$H^+ + HSO_4^-$		1×10^3	
	$HClO_3$	⇌	$H^+ + ClO_3^-$		5×10^2	
	H_3O^+	⇌	$H^+ + H_2O$		55	
	HNO_3	⇌	$H^+ + NO_3^-$		28	
	H_2CrO_4	⇌	$H^+ + HCrO_4^{2-}$		9.6	
	HSO_4^-	⇌	$H^+ + SO_4^{2-}$		1.2×10^{-2}	
	$HClO_2$	⇌	$H^+ + ClO_2^-$		1.1×10^{-2}	
	H_3PO_4	⇌	$H^+ + H_2PO_4^-$		7.1×10^{-3}	
	HF	⇌	$H^+ + F^-$		7.2×10^{-4}	
	CH_3CO_2H	⇌	$H^+ + CH_3CO_2^-$		1.8×10^{-5}	
	H_2CO_3	⇌	$H^+ + HCO_3^-$		4.5×10^{-7}	
	H_2S	⇌	$H^+ + HS^-$		1.0×10^{-7}	
	$H_2PO_4^-$	⇌	$H^+ + HPO_4^{2-}$		6.3×10^{-8}	
	HClO	⇌	$H^+ + ClO^-$		2.9×10^{-8}	
	H_3BO_3	⇌	$H^+ + H_2BO_3^-$		7.3×10^{-10}	
	NH_4^+	⇌	$H^+ + NH_3$		5.8×10^{-10}	
	HCO_3^-	⇌	$H^+ + CO_3^{2-}$		4.7×10^{-11}	
	HPO_4^{2-}	⇌	$H^+ + PO_4^{3-}$		4.2×10^{-13}	
	HS^-	⇌	$H^+ + S^{2-}$		1.3×10^{-13}	
	H_2O	⇌	$H^+ + OH^-$		1.8×10^{-16}	([2])
	CH_3OH	⇌	$H^+ + CH_3O^-$		1×10^{-18}	
	HC≡CH	⇌	H^+ + HC≡C$^-$		1×10^{-25}	
	NH_3	⇌	$H^+ + NH_2^-$		1×10^{-33}	
	H_2	⇌	$H^+ + H^-$	Best	1×10^{-35}	
	$CH_2{=}CH_2$	⇌	$H^+ + CH_2{=}CH^-$	Brønsted	1×10^{-44}	
	CH_4	⇌	$H^+ + CH_3^-$	Bases	1×10^{-49}	

On the basis of this success, I have added the information in Table 2 to my most recent textbook (Spencer et al., 2005). It summarizes the relative strengths of Brønsted acids and bases by organizing the reactions in which a given acid is converted into its conjugate base such that the strongest Brønsted acids are in the upper-left corner of this table; the strongest Brønsted bases in the bottom-right corner.

This table can be used to predict whether certain acid-base reactions should occur. For example, it predicts that acetylene should react with sodium amide because the NH_2^- ion is a strong enough base to remove a proton from acetylene to form the acetylide ion.

HC≡CH	+	NH_2^-	⇌	HC≡C$^-$	+	NH_3
stronger acid		*stronger base*		*weaker base*		*weaker acid*

[2]The value of K_a for water is equal to the value of K_w divided by concentration of water in moles per liter at 25°C.

It also predicts that methanol should react with sodium hydride because the H^- ion should be a strong enough base to remove the acidic proton from methanol.

CH_3OH	+	H^-	$\rightleftharpoons$	CH_3O^-	+	H_2
stronger acid		*stronger base*		*weaker base*		*weaker acid*

This table, therefore, provides a basis upon which our colleagues who teach organic chemistry can build their discussions of Brønsted acid-base concepts as they extend this concept to reactions that do not occur in aqueous solution. This is not a problem, however, because electrode potentials have been used in a similar fashion for decades to analyze oxidation-reduction reactions that do not always occur in aqueous solutions.

It should be noted that many contemporary chemistry textbooks contain tables that provide lists of Brønsted acids and their corresponding Brønsted bases (Brown et al., 1999; Atkins and Jones, 2002; Kotz and Treichel, 2003), and that virtually all current texts discuss the implications of the existence of conjugate acid/base pairs in the Brønsted theory. What is missing from these textbooks, however, is an explicit connection between discussions of the proton transfer that lies at the heart of Brønsted acid/base reactions and discussions of electron transfer that can occur in oxidation-reduction reactions. What is also missing is a mechanism for quantifying the relative strengths of components of the tables of conjugate acid/base pairs found in many textbooks.

The difference between tables of redox half-reactions and Brønsted acid-base reactions historically has been the ability to quantify the relative strengths of oxidizing and reducing agents by adding cell potentials to the table of half-cell potentials. A similar result can be achieved by adding acid-dissociation equilibrium constants to the table of Brønsted acid-base reactions.

Those who question the use of K_a as a means of quantifying the relative magnitude of Brønsted acids or Brønsted bases might wish to look at Pauling's general chemistry textbook (Pauling, 1970). Next to each half-reaction in his table of redox half–reactions he lists not only the value of the half-cell potential but also the equilibrium constant for that half-reaction that would be extracted from that half-cell potential. The entry in his table for the Na/Na^+ half-reaction, for example, is written as follows.

	E^o	K
$Na \rightleftharpoons Na^+ + e^-$	2.712	4.0×10^{45}

The values of K_a in the proposed table of Brønsted acids and their conjugate bases serves the same function as the value of K in Pauling's table of redox half-reactions.

I have had considerably less experience incorporating the connection between weak acid, buffer solution and solubility product calculations in introductory and physical chemistry courses. Anecdotal conversations with students in these classes when this connection has been made, however, have led me to believe that the students understand and appreciate the connection being made. Students in the physical chemistry class also seem to appreciate the difference between this approach and the dogmatic approach certain textbooks take, which would restrict, by definition, application of the term 'common-ion effect' to only those situations that involve the solubility of an ionic substance in the presence of another solute.

References

Atkins P. and Jones L., (2002), *Chemical principles: the quest for insight*, New York, NY, W.H. Freeman.

Ausubel D.P., Novak J.D. and Hanesian H., (1978), *Educational psychology: a cognitive view*, 2nd ed., New York, NY, Holt, Rinehart, Winston.
Brown T.L., LeMay H.E. and Burstein B.E., (1999), *Chemistry: the central science*, 8th ed., Upper Saddle River, NJ, Prentice-Hall.
Cross K.P., (1999), *Learning is about making connections. The Cross papers number 3*, Mission Viejo, CA,League for Innovation in the Community College and Educational Testing Service.
Deming H.G., (1925), *General chemistry: an elementary survey, emphasizing industrial applications of fundamental principles,* New York, NY, John Wiley & Sons, p. 437.
Gagné R.M. and White R.T., (1978), Memory structures and learning outcomes, *Review of Educational Research,* **48**, 187-222.
Good R., Kronhout R.A. and Mellon E.K., (1979), Piaget's work and chemical education, *Journal of Chemical Education*, **56**, 426-430.
Herron J.D., (1975), Piaget for chemists: explaining what 'good' students cannot understand, *Journal of Chemical Education*, **52**, 146-150.
Herron J.D., (1978), Piaget in the classroom: guidelines for application, *Journal of Chemical Education*, **55**, 165-170.
Koroscik J.S., (1996), Who ever said studying art would be easy? The growing cognitive demands of understanding works of art in the information age, *Studies in Art Education*, **38**, 4-20.
Kotz J.C. and Treichel P.M., (2003), *Chemistry and chemical reactivity*, 5th Ed., Thompson/Brooks Cole.
Lagasse P., Hobson A., Norton S. and Goldman L., (2000), The Columbia encyclopedia, 6th ed., New York, NY, Columbia University Press.
Major C.H. and Palmer B., (2001), Assessing the effectiveness of problem-based learning in higher education: lessons from the literature, *Academic Exchange Quarterly*, **5**, 4-8.
Mastropieri M.A. and Scruggs T.K., (1996), Reflections on 'promoting thinking skills of students with learning disabilities: effects on recall and comprehension of expository prose', *Exceptionality*, **6**, 53-57.
Nebergall W.H. and Schmidt F.C., (1957), *College chemistry*, Boston, MA, D.C. Heath and Co., p. 282.
Noyes W.A., (1913), *Textbook of chemistry*, New York, NY, Henry Holt and Co.
Pauling L., (1970), *General chemistry*, 3rd ed., San Francisco, CA, W.H. Freeman and Company, p. 528.
Robertson W.C., (1990), Detection of cognitive structure with protocol data: predicting performance on physics transfer problems, *Cognitive Science*, **14**, 253-280.
Spencer J.N., Bodner G.M. and Rickard L.H., (2005), *Chemistry: structure and dynamics*, 3rd ed., New York, NY, John Wiley and Sons.
Timm J.A., (1944), *General chemistry*, New York, NY, McGraw-Hill.
Vygotsky L., (1986), *Thought and language*, translated and edited by A. Kosulin, Cambridge, MA, MIT Press.
Wikipedia (2006), *The Free Encyclopedia*, http://en.wikipedia.org/wiki/Common-ion_effect, last accessed 27 October 2006.

Letter

A secondary School Teacher Fellow within a university chemistry department: the answer to problems of recruitment and transition from secondary school to University and subsequent retention?

Dudley E. Shallcross and Timothy G. Harrison

Bristol ChemLabS, School of Chemistry, University of Bristol, BS8 1TS, UK.
e-mail: d.e.shallcross@bris.ac.uk

Received 28 November 2006, accepted 4 December 2006

In the UK the changes that have taken place in secondary school science education over the last 20 years are considerable. The national curriculum for sciences has once again been changed and has just been introduced to the current Year 10 (fifteen year olds) in September 2006. The most recent AS/A level (exams at 18) syllabuses (now called specifications) were unitised in 2000 and are to be changed again for September 2008. The International Baccalaureate and School Diplomas are mooted to be the way that secondary education is heading. However, changes to the scheme of work in all science subjects has been a moving target where topics within a subject have been in and out and even changed from one science discipline to another during this turbulent period. A quandary indeed for a University Science Department to maintain congruence at the A level – year 1 undergraduate interface and allow students to have a smooth transition from secondary school or college to University (e.g. Rynne and Lambert, 1997). The School of Chemistry at Bristol University, like many others in the UK, have set up a teaching advisory board (TAB) comprising secondary school teachers, academics and other interested parties. The TAB has proved to be a very helpful mechanism for exchange of ideas but is limited in that secondary school teachers can usually only commit to one or two (unpaid) meetings a year, and the focus of each meeting must be narrow for it to achieve depth of investigation.

School-university transition is not the only problem that all University Science departments struggle with in the UK, retention and recruitment are also major items on the agenda. The well-publicised demise of several Chemistry Departments has highlighted the danger, and in recent years the need to raise the profile of Science, and in particular Chemistry and Physics, has been paramount (Woods and Morris, 2005). What can be done in response to such problems? Barnes (1999) suggested that academics should return to the classroom. Bristol ChemLabS has taken the step to recruit a School Teacher Fellow (STF), and here we outline the potential benefits to all concerned of recruiting a School Teacher Fellow.

The desirable characteristics of a School Teacher Fellow

Ideally, School Teacher Fellows (STFs) should be qualified and experienced secondary school teachers who would work in the University department of their specialism. Such persons should have already held some position of responsibility in the school's middle management so that the formality of meetings, giving presentations to peers, report writing and bid writing are not alien to them. They would also be used to working with groups of teachers other than those in their own schools. Their experience would, of course, be essential when advising on changes to course structures and how this may affect subject knowledge and transferable skills for a range of undergraduates and even postgraduates in a University department. It would be quite important, but not essential, that such teachers should already have strong links with the outreach activities already going on in the university department where they are to work. The STFs would be seconded from their schools for at least a year (see later) and would maintain strong links with those schools during the tenure of their post.

The benefits of a School Teacher Fellow to the University Department

School-university transition and retention

A School Teacher Fellow will know at least one of the examination boards' specifications and their own school's scheme of work derived from it in their subject particularly at A/AS level, and can advise departments on congruence between first year undergraduate courses and A level. A particular emphasis on advising on the practical skills base of new undergraduates is vital. The STF can also work closely with technical staff on matters of experiment design, and appropriateness of equipment and labelling. Ausubel (1968) encouraged us to ascertain what the learner already knows and to teach accordingly, the STF is the ideal source of this information.

A STF will draw the attention of university staff to differences in approach to the use of IUPAC nomenclature between schools and universities that may lead to serious difficulties for new undergraduates, and will be able to advise on labelling and appropriate naming.

A STF will know what modes of assessment will be familiar to students. Feedback to a teaching committee will ensure that students have ample training where required for new forms of assessment. It may be that the school-level methods of assessment, particularly in practicals, may be built on at the University.

STFs may act as supplementary tutors to students who are struggling with basic concepts from A level and those students recruited via access routes without A level, thereby supporting widening access. They may even take on a pastoral role, being more familiar with the students' prior experience and potential problems than any University lecturer. A STF may also be one of the first points of contact by undergraduates or postgraduates considering a career in teaching by providing accurate and up-to-date information about the profession and its working conditions, and may assist in arranging some relevant work experience/classroom observation.

Recruitment

STFs will know:

- the secondary school year and in particular when examinations occur. They will also know the difficulties in taking students out of school, and the timescale required to plan to do so. Therefore, recruitment and outreach events can be arranged to have maximum impact and appeal in terms of timing.
- what secondary school teachers are looking for during an outreach event. This is also true for CPD events for teachers.

STFs will themselves have wide experience with public engagement and may know about sources of funds and resources, particularly those that will fund public engagement activities. They can work with higher education colleagues in the preparation of articles on leading edge research at the appropriate levels for use in schools, for example articles prepared for journals such as Chemistry Review or Education in Chemistry. This also applies to postgraduates who wish to write or give presentations to secondary schools where the use of appropriate language level for age group can make or break an event. They can also work closely with research staff in the development of lectures and workshops for different aged groups of young people or the public in general.

Ultimately, a major role of the STF would be finding ways in which the many resources of the university and department can be used to stimulate an active interest in the subject from school students of all ages and from the general public. This may be through the organisation of conferences, workshops, competitions, the training of less experienced secondary school teachers of that subject, and in working with academic staff on how to present their research to the wider public. The School Teacher Fellow will have the time to prepare and front numerous activities for which academic staff do not have time because of their teaching and research commitments. They will know the possible pitfalls and the correct level of language appropriate to the courses being taken by the students. In activities that engage teachers as part of professional development it is helpful to have a School Teacher Fellow at least present, if not leading the courses, to put the training into a classroom context. For summer schools a far more ambitious and well thought out approach can be adopted (our experience).

The benefits to a School Teacher Fellow

There are many secondary school teachers who want to teach, and yet as they progress in seniority the level of teaching, contact with students and the time available to devote to innovative lessons will diminish. Teachers may have ideas they wish to explore, but lack adequate resources to do so effectively. Having the resources, time, technical expertise and equipment allows the development of these ideas. Being a STF can recharge professional batteries. The teachers may develop new skills such as the presenting of lectures to very large audiences. The opportunity to write articles for fellow teachers or for students is also a benefit. STFs can experience new practical techniques and learn new aspects of science, both of which can be used to augment their lessons when back in schools.

The benefits to a school in releasing a School Teacher Fellow

The benefits are not only to the university or to the teacher concerned. There will be a much stronger relationship between the school and the university which could be of great benefit to a number of the school's teachers and not just the STF. Via the STF the school may be able to access potential funding streams for future school-based projects that were previously unknown. There will also be a substantial amount of Continuing Professional Development (CPD) for one of their 'middle managers' at no cost to the school. The school's own students may be used to pilot ideas, perhaps introducing some students to the possibility of higher education for the first time. Finally, the schools will benefit from more highly skilled teachers on their return.

Acknowledgments

DES thanks the Higher Education Academy UK for a National Teaching Fellowship and both DES and TGH thank Bristol ChemLabS under whose auspices this work was carried out. Bristol ChemLabS is a HEFCE (Higher Education Funding Council for England) CETL (Centre for Excellence in Teaching and Learning) see http://www.chemlabs.bris.ac.uk/overview.htm for more details. DES is Outreach Director for Bristol ChemLabS and TGH is the School Teacher Fellow.

References

Ausubel D.P., (1968), Educational psychology: a cognitive view, New York, Holt, Rinehart and Winston.

Barnes N., (1999), Switching places: why college teachers should teach high school students, *Curriculum Inquiry*, **29**, 293-313.

Rynne E. and Lambert D., (1997), The continuing mismatch between students' undergraduate experiences and the teaching demands of the geography classroom: experience of pre-service secondary geography teachers, *Journal of Geography in Higher Education*, **21**, 65-77.

Wood J. and Morris A., (2005), Report of the working group on physical science and maths education, *NERF Working Paper 5.4*, December [http://www.nerf-uk.org/word/5.4PhysicalScience.doc?version=1]