Measuring the effectiveness of computer-based scientific visualizations for conceptual development in Australian chemistry classrooms

David Geelan AND Michelle Mukherjee
University of Queensland
Australia
d.geelan@uq.edu.au
m.mukherjee@uq.edu.au

Abstract: Visual modes of representation have always been very important in science and science education. Interactive computer-based animations and simulations offer new visual resources for chemistry education. Many studies have shown that students enjoy learning with visualizations but few have explored how learning outcomes compare when teaching with or without visualizations. This study employs a quasi-experimental crossover research design and quantitative methods to measure the educational effectiveness - defined as level of conceptual development on the part of students - of using computer-based scientific visualizations versus teaching without visualizations in teaching chemistry. In addition to finding that teaching with visualizations offered outcomes that were not significantly different from teaching without visualizations, the study also explored differences in outcomes for male and female students, students with different learning styles (visual, aural, kinesthetic) and students of differing levels of academic ability.

Introduction

Scientific visualizations – visual representations of scientific data as well as of objects and interactions – are an increasingly important set of tools used by scientists in their work. Visualizations are also increasingly being used in science teaching. While there are both extravagant claims (e.g. Bell, Park & Toti, 2004; Kozhevnikov & Thornton, 2006) and some encouraging results (e.g. Cifuentes & Hsieh, 2001; Dori & Belcher, 2005; Hakerem, 1993; Hinrichs, 2004; Royuk & Brooks, 2003; Williamson & Abraham, 1995) in relation to the educational effectiveness of the pedagogical use of such visualizations, there is little formal research work, particularly quantitative research, which specifically addresses this issue, particularly at the secondary school level.

Millions of dollars are being spent on the development of Learning Object Repositories (Koppi, Bogle & Bogle, 2005) in numerous jurisdictions, in the absence of much more than anecdotal evidence for the educational effectiveness of such teaching approaches. The study described in this paper is intended to begin to provide such evidence through allowing direct comparison of conceptual development on the part of students taught using scientific visualizations with that of the same students when taught using traditional classroom teaching methods.

We have chosen the narrower term ‘conceptual development’ over the broader term ‘learning’ for use in this project both because our interest is specifically in students’ development of well-elaborated understandings of scientific concepts (as opposed to retention of scientific facts and data or other forms of learning) and because there is a well recognised literature on conceptual development in science and well validated instruments for measuring students’ conceptual development that can be used as models for the development of our own instruments.

Research Question

The research question for the present project can be stated as:

Is teaching with scientific visualizations more effective than traditional classroom teaching for supporting students’ conceptual development of specific concepts in chemistry?

The independent variable is the ‘treatment’ – the teaching of the science concepts using visualisation or ‘traditional’ methods. The dependent variable is conceptual development, conceptualised as change in conceptual understanding between pre-instruction and post-instruction situations, measured using conceptual knowledge tests developed by the research team in an approach very similar to that adopted in the Force Concepts Inventory (Hestenes, Wells & Swackhamer, 1992) and the Chemistry Concepts Inventory (Mulford & Robinson, 2002).
Scientific Visualizations and Learning

Our focus in this project is on the use of a particular set of technologies, which we broadly label ‘scientific visualizations’, for teaching in science. The term ‘visualizations’ in science teaching is broadly applied to children’s drawing in the exploration of scientific ideas (Brooks, 2009) or to external visualizations of scientific concepts such as gestures or paper drawn diagrams (Subramaniam & Padalkar, 2009). In this project, we define it as computer-based animations or simulations. All the selected visualizations are available at no cost from the internet.

Numerous authors (e.g., Copolo & Hounshell, 1995; Gordin & Pea, 1995; Kali & Orion, 1997; Pea, 1994; Wu, Krajick & Solloway, 2001) have argued that visualizations make information that might otherwise remain opaque perceptible and cognitively tractable. Moreover, several researchers (e.g., Wu et al., 2001) have confirmed in experimental studies that visualizations convey a clear benefit in some forms of learning.

Visualizations influence greatly the ways in which scientists, mathematicians, and engineers practice their respective fields. They become integrated into the practice of a scientific discipline, and in turn engender new ways of thinking about relevant information (e.g., Kaput, 1999; Nemirovsky, 1994; Suwa and Tversky, 1996). For a practicing scientist, the distinction between a visualisation of reality and reality itself may become blurred: the visualizations become a primary tool of practice (Hutchins, 1995). The use of multiple representations can be valuable in reminding users that our best understandings of scientific phenomena are often models of reality rather than reality itself.

Visualizations play an equally central role in education. Visualizations have been used to extend the reach of instruction by overcoming the limitations of traditional ways of representing information (Horwitz, 2002; Tinker, 1999). In many fields of science education, acquiring an understanding of visualizations is critically important for mastering relevant concepts. Treagust and Harrison (2000) describe the processes by which explanations in science teaching support development of “a dynamic and fluid mental model” on the part of the learner, and it seems plausible that various forms of visualisation can powerfully extend the teacher’s ‘toolkit’ for helping students in this process (Edelson, Gordin, & Pea, 1999).

Visualizations are especially important for teaching concepts in chemistry, which studies the world that is too small to see and the ways in which structures and properties at the molecular scale influence macroscopic properties. It is, therefore, not surprising that funding bodies and other organisations have devoted tremendous resources to the development and use of visualizations in science and science education. However, relatively little research has evaluated the effectiveness of visualisation use. “At the moment, most of our information on how to use simulations and visualizations in the classroom is based on anecdotal evidence” (Horwitz, 2002). There are, of course, important exceptions to this claim (see e.g. Gobert & Pallant, 2004; Wu, Krajcik, & Soloway, 2001).

Conceptual Development and Misconceptions

Posner, Strike, Hewson and Gerzog (1982) suggest that, by analogy with the social processes of paradigm shift in the scientific community at large described by Kuhn (1970), individuals learn new scientific schemes through a process of ‘conceptual change’. This four-part scheme - dissatisfaction with a current conception, dealt with by the development of a new conception which is intelligible, plausible and fruitful - is the theoretical heart of conceptual change perspectives on learning (e.g. Smith, Blakeslee & Anderson, 1993).

It seems plausible to suggest that computer-based scientific visualizations might have the potential to support teachers and students in each of these dimensions: demonstrating the shortcomings of students’ existing conceptual frameworks and helping them to develop models of the new conception that are intelligible, plausible and promise to be fruitful. This study, however, will not directly yield information about the mechanism by which visualizations yield improved conceptual understanding (if indeed they do). The results will show only the extent of any differences (the ‘what is happening’) – a later qualitative study involving interviews with students, classroom observations and ‘think aloud’ protocols would be required to explore more deeply the specific learning mechanisms associated with visualizations.

An extensive literature has grown up in chemistry education around the conceptual change notion, focused on exploring the ‘misconceptions’ that students bring to class, and the processes of teaching and learning involved in changing students’ conceptions of scientific phenomena from these ‘misconceptions’ to the ‘correct’ scientific concept. (It should be noted that, for a variety of reasons, some traditions within educational research prefer the terms ‘naïve conceptions’, ‘alternativ(e) conceptions’, ‘prior conceptions’ or ‘children’s science’ over the term ‘misconceptions’, but the latter has been the dominant term.)

Hestenes, Wells and Swackhamer (1992) developed the Force Concepts Inventory (FCI) in order to allow physics teachers to measure the extent to which students’ conceptions around the concept of ‘force’ fit the received scientific
conception. Each of the 29 multiple choice items on the FCI presents one correct (i.e. Newtonian) answer and four answers derived from various known misconceptions from the science education literature. A decade later, Mulford and Robinson (2002) developed the Chemistry Concepts Inventory (CCI - sometimes also called the ‘Chemical Concepts Inventory’). Also a multiple-choice instrument focused on distinguishing students’ correct conceptions from their misconceptions, the CCI has 22 items, some of them linked such that the second question elicits from students an explanation of their response to the first. The CCI has also been used to explore the chemical conceptions of chemistry teachers (Kruse & Roehrig, 2005).

The Chemistry Concepts Inventory is more broadly focused than the Force Concept Inventory: the latter is based around one, albeit complicated, set of concepts around force and Newton’s laws, whereas the former attempts to address many of the key concepts covered in an entire first year university chemistry course.

These two inventories have been used as models for the development of the conceptual tests used in the present study. Each test – the same tests are used as both pre- and post-test – contains 12 multiple-choice items, each with four possible responses; one scientifically correct response and three responses representing common student misconceptions in relation to the concepts taught. Here are a few sample items:

**Le Chatelier’s Principle**

Question 10 relates to the reversible reaction of iron (III) ions, Fe$^{3+}$, with thiocyanate ions, SCN$^-$ to produce iron thiocyanate, FeSCN$^{2+}$, ions in accordance with the equation:

$$\text{Fe}^{3+}_{(aq)} \text{(pale yellow)} + \text{SCN}^-_{(aq)} \text{(colourless)} \rightleftharpoons \text{FeSCN}^{2+}_{(aq)} \text{(red)}$$

10. If colourless solid potassium thiocyanate, KSCN$_{(s)}$, is added to the solution, it will dissolve producing thiocyanate, SCN$_{(aq)}$, ions according to the reaction $\text{KSCN}^{(s)} \rightleftharpoons \text{K}^+_{(aq)} + \text{SCN}^-_{(aq)}$. As it comes to its new equilibrium the colour of the solution will:

a. become more red  
b. become paler  
c. stay the same  
d. there is not enough information to tell

**Intermolecular Forces**

9. Although the water molecule has no overall electric charge (it is neutral), a stream of water will be attracted to a charged rod. This attraction is due to:

a. an induced dipole in the water molecule  
b. the water molecules separating into charged H$^+$ and OH$^-$ ions  
c. the existing dipole (charge separation) between the O and H atoms in water molecules  
d. electrons being removed from the water by the charged rod to create H$_2$O$^+$ ions

**Thermochemistry**

1. The reaction between octane and air is very exothermic, and yet an open container of octane can be left at room temperature for several days without catching fire (i.e. reacting) (although it will evaporate). This is because:

a. octane is naturally in a liquid state  
b. energy must be supplied to start the reaction  
c. there is not enough oxygen in the air to start the reaction  
d. energy must be removed from the system to break the bonds in the octane before it can reac

**Significance**

Much of the published literature in the field of educational technology still tends toward what might be described as ‘technoboosterism’ – a relatively uncritical belief that information technology based approaches to teaching and learning will yield improvements in students’ attitude to and engagement with learning as well as in their
understanding and achievement. This effect is exacerbated by the fact that often papers are written by the originators of the particular technological application being described, so that many reports are of the ‘I made it, I used it, it was great!’ genre. There certainly have been critical studies and reviews of the literature on the effectiveness of ICT-based teaching innovations (e.g. Clements & Sarama, 2003; Cordes & Miller, 2000; Kompf, 2005; Reeves, 1995) but there is still a dearth of well-designed studies that measure the educational effectiveness (defined more narrowly as conceptual development effectiveness in this study) of various forms of ‘technologies for teaching and learning’.

Several good examples of experimental and quasi-experimental studies of conceptual development in science education supported by various forms of educational technology do exist, including Dori & Belcher’s (2005) work on electromagnetism with undergraduates, Hinrichs’ (2004) work on his ‘system schema’ tool, Williamson and Abraham’s (1995) work on the particulate nature of matter and Kozhevnikov and Thornton’s (2006) study in relation to spatial visualisation ability. These studies are all at the university undergraduate level, however, rather than the high school level. There are also a number of studies, like those of Cifuentes and Hsieh (2001), focused on student engagement, and Robblee et al. (2000), focused on teacher attitude, that relate to issues surrounding educational technology but do not directly address students’ conceptual development.

The present study is intended to continue the process, which is in its early stages, of contributing to the literature studies that do not assume the superiority of computer-based visualizations for learning, but rather seek evidence of the relative benefits for conceptual development of teaching approaches in science employing scientific visualizations vis a vis more traditional science teaching approaches.

Approach and Methodology

While there are quantitative experimental or quasi-experimental studies conducted in non-classroom settings e.g. Shepard and Metzler’s study of the mental rotation of three dimensional objects (Shepard & Metzler, 1971), and qualitative classroom case studies (e.g. Subramaniam & Padalkar, 2009), there are very few high quality quasi-experimental studies of the ‘real world’ classroom use of visualisation technologies in teaching.

Crossover research design, although it has a long history in clinical trials in medicine, agriculture and other scientific fields, is a methodology that has not been common in educational research. This is surprising in some ways, since its features offer significant benefits in conducting quantitative research within the set of constraints offered by school classrooms. This study uses an adapted form of crossover design that ‘fits’ with the constraints of the classroom while continuing to support quasi-experimental quantitative research.

The focus of this research project is specifically on a quantitative comparison between the effectiveness of purpose-developed computer-based scientific visualizations and ‘traditional’ classroom teaching methods for the purpose of helping high school students to develop particular scientific concepts.

Conceptual development on the part of students was measured using a conceptual knowledge test based on the Chemistry Concept Inventory (CCI)(Mulford & Robinson, 2002). This instrument was designed to distinguish the extent to which students have developed the ‘correct’ scientific concept in relation to a topic, rather than any of a number of possible ‘misconceptions’. The Inventory has been used extensively internationally and is well validated (Kruse & Roehring, 2005). Each subject test comprises 12 multiple choice items, with four possible answers, and the distractors focus on the common misconceptions as identified in the Chemistry Concept Inventory.

Specific concepts that appear in the Queensland Year Eleven Chemistry syllabuses were chosen for the study. Groups of students in a number of purposively chosen Brisbane area government high schools were taught these concepts in their normal science classes, and the conceptual knowledge tests were used before and after each teaching sequence to measure students’ conceptual development. Classes at schools with relatively large class sizes in Year Eleven Chemistry were chosen for the study.

The teachers using the visualizations were provided with teaching points to include, but were left to structure the lesson in their individual style, using their personal professional judgements. Due to the possible variations in the presentation of the material across different classes providing test results for the same topic, we conducted classroom observations. The primary intent was to look at the teaching style – did the teacher demonstrate the visualisation on a projector screen, in a more transmissive style, or were the students interacting personally with the visualisation – was there groupwork and discussion during the learning, how large were the groups and were all members engaged? We were also determining the amount of prior knowledge the students had before completing the pre-test – although the pre-test was taken at the beginning of the unit, due to the overlaps and interconnectedness of topics in the syllabus, students frequently had experienced some previous exposure to the topic before it was formally studied. We also noted the gender breakdown of the students, the classroom layout and facilities, the number of absentees that day and the number of English as a Second Language students, to gauge whether the literacy demands of the items might be
influencing the outcome. The classroom observations allowed us to qualitatively determine the extent of the effectiveness of the visualisation as a teaching tool.

**Crossover Design**

A ‘crossover’ (Ratkowsky, Evans & Alldredge, 1993) research design has been chosen in order to yield strong quantitative results, including the ability to calculate effect sizes, within the constraints of the high school science classroom situation. These constraints, including the difficulty of truly random assignment of students to experimental and control groups, constraints on the concepts that can be taught due to the syllabus and the difficulty of matching teaching style variables between classes, have tended to make quasi-experimental designs difficult to carry out in classroom settings.

One benefit of crossover designs is that individual participants are essentially their own controls, since they receive both the ‘treatment’ of interest in the study (in this case the pedagogical use of scientific visualizations) and the ‘control’ situation (in this case ‘traditional’ classroom teaching). In an educational situation, where the teaching style of the teacher as well as his/her relationships with the students has the potential to influence the results of a study using multiple teachers, the crossover design also in a sense allows each teacher to be his/her own control. Statistical analysis then compares conceptual growth for all students under each condition.

It should be noted that ‘traditional’ classroom teaching is used here as a shorthand term to denote all the features of the way in which the participating classroom teacher would usually teach these concepts. ‘Traditional’ teaching methods will likely include some lecturing, demonstrations, experiments, diagrams, calculations, class discussions and other activities. The use of the term ‘traditional’ here is explicitly not used as a contrast with constructivist teaching, or as shorthand for lecture-and-notes only teaching. Teachers were asked not to use other scientific visualizations during the ‘control’ (non-visualisation) teaching sequences even if they would usually use them for that topic (many teachers in the study reported that they already use visualizations in their teaching to various extents). The comparison is therefore essentially one between ‘teaching with visualizations’ and ‘teaching without visualizations’.

Workshops were conducted for the participating teachers. These focused on supporting the teachers’ understanding and pedagogical use of the developed scientific visualizations. They also helped the participating teachers to compare and discuss their own understanding of the scientific concepts and elaborate their understandings. All of the participating teachers taught their students both ‘traditionally’ and using scientific visualizations, and the crossover design allows differences due to teacher personal style to be taken into account in a way direct experimental comparisons of the classes of different teachers does not.

For a simple crossover design, two groups would be used, would receive the two treatments in opposite orders. That is, if teaching using scientific visualizations is designated as V and traditional classroom teaching is designated T, some students would receive the teaching sequence V → T and others T → V.

This sequence is not appropriate for the present study, however, because in order to make the comparisons valid it would be necessary to find two concepts with exactly equal difficulty (since a further constraint of the classroom context is that the same students cannot be taught the same content twice using the different teaching methods).

Since it would be very difficult if not impossible to exactly match two scientific concepts in terms of their level of difficulty for students, two different concepts were chosen, and a modified crossover design used to take into account the different concepts. If the concepts are designated ‘a’ and ‘b’, then the four different treatment conditions can be summarised as follows: TaVb, TbVa, VaTb, VbTa. Since the same students cannot be taught the same concepts twice, the four possible ‘XaYa’ and ‘XbYb’ conditions are not included in the study (that is, having the same teachers teach the same students the same concepts twice using different methods). It would be desirable in a larger scale study to include the four ‘TxTy’ and ‘VxVy’ conditions (that is, having a teacher teach his/her students both concepts using traditional methods or both concepts using visualizations), however it is felt that this would unnecessarily complicate and expand the scope of the present study.

The crossover design also has the potential to allow ‘order effects’ to be analysed, addressing questions about the preferred sequence of concepts and teaching modes, and whether there are ‘carryover effects’ from one method to another (Ratkowsky, Evans & Alldredge, 1993). In the present study, however, the interest is in the relative effectiveness of the different teaching modes. For this reason some weeks (with other intervening teaching) will be allowed to elapse between the treatment and testing sequences for each class. This is seen as the equivalent of a ‘washout’ phase in a drug trial, and means that it will be assumed that any effects from the prior treatment have been submerged in ‘normal’ teaching and learning, allowing direct comparisons between the scores for each group on the two trials.
In addition, results were analysed against the sex of participating students, their score on a simple learning styles inventory (adapted from Dunn, Dunn & Price, 1984) and a teacher assessment of whether a particular student is in the top, middle or bottom third of the class in terms of academic ability.

An ANOVA of score increases (post-test – pre-test scores) on the single factor of teaching mode was used to analyse results, and effect sizes calculated.

The situation was made slightly more complex by the fact that the Queensland chemistry curriculum is not very prescriptive in terms of which topics should be covered in Grade 11 and Grade 12, so different schools cover the course in different orders. The initial intention was to choose two different topics of similar difficulty and conduct a simple crossover study as described above, however it soon became clear that more than two topics in total were needed if each school was to be able to work with two topics taught in Grade 11 (where the study was focused). We ended up choosing three topics, which makes the analysis more complex but will still allow quite robust comparisons to be made. It is important to note that despite the three chemistry ‘content’ topics, what is being ‘crossed over’ is the visualizations/no visualizations condition, so the analysis is still a crossover study with that one independent variable.

The three concepts chosen were Le Chatelier’s Principle (and dynamic chemical equilibria more broadly), Intermolecular Forces (and other interparticle forces) and Thermochemistry. These were linked to teaching sequences intended to take three to four lessons, or about one week of normal Grade 11 chemistry lessons. One or more web-based visualizations were chosen for each concept – links to the visualizations are included below.

**Le Chatelier’s Principle**

http://www.mhhe.com/physsci/chemistry/essentialchemistry/flash/lechv17.swf

**Intermolecular Forces**

http://www.kentchemistry.com/links/bonding/bondingflashes/bond_types.swf

http://faculty.washington.edu/dwoodman/IntrFrcs/dswmedia/IntrFrcsW.html

http://www.chm.davidson.edu/ronutt/che115/Phase/Phase.htm

**Thermochemistry**

http://www.bravus.com/visual/bondenthalpy.mov

http://schools.matter.org.uk/Content/Reactions/BondEnergy.html

http://schools.matter.org.uk/Content/Reactions/BE_enthalpyHCl.html

We chose to use existing resources that were available on the net. This may have led to less directly comparable visualizations in terms of approach and style, but we felt that it allowed us to model more closely what really happens in school classrooms.

Some of the teachers were concerned that the students would be disadvantaged in their learning if they learned a particular concept without visualizations (many of the teachers already routinely used scientific visualizations in their teaching), however the short teaching sequence meant that after the post-test (in the non-visualisation teaching sequence) the students could then revise the concept using the visualizations.

We had realised that it is often difficult in schools to use web-based resources, due to restrictions such as limited bandwidth, very strict content filters, malfunctioning computers and lack of plug-ins such as Flash and Java, and in fact these difficulties arose, to a greater extent than we had expected. As far as possible we moved the materials offline, and a class set of laptops was bought in order to be taken into schools where it was impossible to run the visualizations. A combination of strategies made it possible to complete the research study, but it is clear that there are still significant challenges in allowing all teachers to have the option to use visualizations, even if the evidence suggests that it enhances learning.

One extra layer of analysis was added in order to further explore the ways in which visualizations support learning: data were analysed by students’ sex, learning style (using an adapted version of the Visual-Auditory-Kinesthetic (VAK) test) and academic ability (ranked by the teacher as to whether each student was in the top,
middle or bottom third of the class). This enabled us to explore whether visualizations were supportive generally, and whether they were more supportive for some students than others.

Analysis and Findings

Eighty-seven (87) students from 7 Queensland schools participated in the study. Both state (public) and private schools were included in the study. Sixty-four (64) of the students were female and twenty-three (23) male: this sex distribution is fairly typical for Year 11 Chemistry classes in Queensland.

Whole group

Conceptual tests of 12 items were given to all students before and after each instructional sequence. The same test was used as the pretest and posttest for each sequence. For all students, the mean scores and differences are summarised in Table One.

<table>
<thead>
<tr>
<th></th>
<th>Pretest mean</th>
<th>Posttest mean</th>
<th>Mean difference</th>
<th>SD of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without visualisation</td>
<td>4.943</td>
<td>6.253</td>
<td>1.310</td>
<td>2.354</td>
</tr>
<tr>
<td>With visualisation</td>
<td>5.529</td>
<td>7.230</td>
<td>1.701</td>
<td>2.703</td>
</tr>
</tbody>
</table>

Table One – Mean scores, score increases and standard deviations, pretest to posttest, for teaching without visualisations and with visualisations

A two-tailed t-test of the mean differences (increase in conceptual understanding) shows that the effect is statistically significant at the 0.000 probability level. This provides relatively robust evidence that, in this study, teaching with visualisations did lead to significantly better conceptual development outcomes than teaching without visualisations. Because the sample is reasonably large, however, statistical significance can be attained relatively easily, and does not give a clear sense of the effect size. A number of measures of effect size are used for this purpose, and Cohen’s d was approximated by dividing the difference between the mean differences (1.701 – 1.310 = 0.391) by the simple average of the standard deviations of the two conditions (since n for each is 87 and they are the same individuals under different treatments 2.529). This gave an approximate value for d of 0.154. This is a quite small effect size, though not negligible. The extent to which visualisations enhance learning is likely smaller than many of the other influences on learning that occur in classrooms. See the Conclusion for further discussion of this issue.

Sex

The mean difference scores achieved by male (n=23) and female (n=64) students in the with- and without-visualisation cases is shown in Table Two.

<table>
<thead>
<tr>
<th></th>
<th>Mean difference (male, n=23)</th>
<th>Standard deviation</th>
<th>Mean difference (female, n=64)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without visualisation</td>
<td>1.826</td>
<td>2.059</td>
<td>1.125</td>
<td>2.440</td>
</tr>
<tr>
<td>With visualisation</td>
<td>2.652</td>
<td>2.497</td>
<td>1.359</td>
<td>2.710</td>
</tr>
</tbody>
</table>

Table Two – Mean score differences and standard deviations, pretest to posttest, for teaching without visualisations and with visualisations, male and female students

The difference between the mean differences without and with visualisations for male students was 0.826 while the difference for female students was much smaller at 0.234 marks. This coarse measure suggests that boys benefited to a greater extent than girls when learning with visualisations. A t-test bears out this impression, showing significance at the 0.05 confidence level. Using the same approximation for Cohen’s d, the effect size for male students was 0.36, which is a low-to-medium effect size, while the effect size for girls was only 0.09, which is almost nonexistent. While the small size of the male sample means this result should be treated with care, it does seem as though learning with visualisations might be particularly effective for boys.

Learning Style
We have our reservations about the simplistic ways in which learning styles inventories are sometimes used in educational practice, and recognise that any scheme of this kind is useful rather than true (and sometimes dangerous, too). We still thought, however, that it was worthwhile to explore the question of whether students who were particularly strong in particular learning styles benefited to a greater or lesser extent from learning with visualisations. Students were asked to complete a short learning styles inventory and scored on visual, auditory and kinesthetic learning styles. Students were not assigned to one dominant group (i.e. identified as ‘a visual learner’). Rather the scores on all three dimensions were retained. Correlation between each of the three learning styles and learning with visualisations was calculated, and is shown in Table Three. More sophisticated analysis around this issue will be conducted in the next few months for formal publication.

<table>
<thead>
<tr>
<th></th>
<th>Correlation with mean difference (teaching with visualisations) (n=87)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>-0.183</td>
</tr>
<tr>
<td>Auditory</td>
<td>-0.017</td>
</tr>
<tr>
<td>Kinesthetic</td>
<td>0.223*</td>
</tr>
</tbody>
</table>

Table Three – Correlation (Pearson’s r) between learning styles and mean score difference for teaching with visualisations (* = significant at the 0.05 level)

Very weak negative correlations, not statistically significant, were seen between students with high inventory scores in relation to visual and auditory learning styles (not ‘visual and auditory students’) and learning with visualisations. Perhaps a little surprisingly – it would seem plausible on its face that strong visual learners would benefit most from visual tools for learning – the only statistically significant correlation was the positive one seen for students strong on the kinesthetic learning style.

Academic Achievement

Teachers were asked to divide their students (confidentially) into three groups based on their demonstrated academic achievement – the upper, middle and lower thirds of the class. Table Four shows the mean score differences – defined by subtracting the difference in mean posttest and pretest scores on the conceptual development test for teaching without visualisation from that for teaching with visualisation – and the effect sizes that these differences represent for each of the three groups.

<table>
<thead>
<tr>
<th></th>
<th>Mean score difference</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower third (n=18)</td>
<td>-0.389</td>
<td>-0.143</td>
</tr>
<tr>
<td>Middle third (n=45)</td>
<td>0.356</td>
<td>0.148</td>
</tr>
<tr>
<td>Upper third (n=16)</td>
<td>1.687</td>
<td>0.684</td>
</tr>
</tbody>
</table>

Table Four – Learning gains and effect sizes for students versus (teacher reported) academic achievement (The participating teachers obviously interpreted the notion of ‘thirds’ of the class rather freely)

Interestingly, the strongest gains in terms of effect size – at 0.684 this would be described as a medium-strong effect – was for the strongest students. It seems that visualisations are particularly helpful for enhancing the learning of the students who are already successful in chemistry. The weakest group of students actually learned better (achieved larger learning gains – all groups achieved gains, but the negative value of the result above shows that the gains were greater without visualisations than with) when taught in traditional ways. While these results should also be treated with caution due to the relatively small sample sizes in the highest and lowest groups, they may at least suggest that if the goal is to raise the achievement of students who are struggling, more traditional approaches such as teacher explanations may yield better results than more complex representations such as scientific visualisations.

Conclusion

The effect sizes measured in this study are quite small. The results are telling us that scientific visualisations do add something worthwhile to the teaching of science, when measured in terms of the conceptual development of students. This is important, since there was little real evidence one way or another available prior to this study. It is
important to remember, though, that there is a quite large body of evidence to support the idea that there are affective and motivational benefits to the use of visualisations (e.g. Cifuentes & Hsieh, 2001; Dori & Belcher, 2005; Royuk & Brooks, 2003). In a sense, given these other benefits, simply a ‘first do no harm’ finding in relation to the learning gains would have been sufficient. That is, given that students enjoy learning with visualisations, enjoy their science lessons more and are more engaged with their learning, even if teaching with visualisation only yielded the same outcomes as teaching without them, their use could be justified on the grounds of the other benefits. It’s encouraging, then, to find that there are also modest but real increases in student conceptual learning when visualisations are used in classrooms. The findings warrant further study into the specific issues of sex and ability/achievement level. Qualitative research focused on the specifics of students’ experience of learning with visualisations in chemistry classrooms and the personal and social meaning they make of such learning would also be valuable.

References


