One Approach To Finding Evidence For The Effectiveness Of Scientific Visualisations In High School Physics And Chemistry Education

David R Geelan and Michelle Mukherjee
School of Education
The University of Queensland
Brisbane QLD 4072, Australia
d.geelan@uq.edu.au

Abstract: Enormous amounts of money and energy are being devoted to the development, use and organisation of computer-based scientific visualisations (e.g. animations and simulations) in science education. It seems plausible that visualisations that enable students to gain visual access to scientific phenomena that are too large or too small, occur either too quickly or too slowly to be seen by the naked eye, or to scientific concepts and models, would yield enhanced conceptual learning. When the literature is searched, however, it quickly becomes apparent that there is a dearth of quantitative evidence for the effectiveness of scientific visualisations in enhancing students’ learning of science concepts. This paper outlines an Australian project that is using innovative research methodology to gather evidence on this question in physics and chemistry classrooms.

Introduction
Scientific visualisations – 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. Visualisations 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 such visualisations, there is little formal research work that specifically addresses this issue.

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 visualisations 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.

Scientific Visualisations and Learning
Our focus in this project is on the use of a particular set of technologies, which we broadly label ‘scientific visualisations’, for teaching in science. The term ‘visualisations’ in science teaching is sometimes applied more broadly to a range of visual media from childrens’ drawing in the exploration of scientific ideas (Brooks, 2009) to external visualisations of scientific concepts such as gestures or paper drawn diagrams (Subramaniam & Padalkar, 2009). In this project, we have focused more narrowly on (interactive) computer simulations. All the selected visualisations used in the study 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 visualisations make perceptible and cognitively tractable information that might otherwise remain opaque. Moreover, several researchers (e.g. Wu et al., 2001) have confirmed in experimental studies that visualisations convey a clear benefit in some forms of learning.

Visualisations 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 visualisations 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).

Visualisations are especially important for teaching concepts in chemistry and physics, which study the world that is too small to see, forces that cannot be felt, or electromagnetic radiation outside the visible spectrum. It is, therefore, not surprising that funding bodies and other organisations have devoted tremendous resources to the
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 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

**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 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

**Conceptual Development and Misconceptions**

Posner, Strike, Hewson and Gerzog (1982) suggest that 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 visualisations might have the potential to support teachers and students in each of these dimensions. This study, however, will not directly yield information about the mechanism by which visualisations yield improved conceptual understanding. The results will show only the extent of any differences – a 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 visualisations.

An extensive literature has grown up in chemistry and physics 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.

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. The 29 items are divided to yield six scales relating to particular subconcepts relating to forces.

Huffman and Heller (1995), using factor analysis, have challenged this division of the items on the FCI on the basis of their analysis of student scores, but Hestenes and Halloun (1995) have defended the FCI on the basis that the categories are based on an expert understanding of physics, rather than on students’ responses. In fact, Hestenes and Halloun claim that Huffman and Heller’s results actually support their contention that students who have not yet developed strongly Newtonian views on the nature of force will tend to demonstrate scattered and inconsistent knowledge of the items on the FCI. The FCI has been extensively validated, and has been used, for example, to study the use of a representational scheme in a university physics course (Hinrichs, 2004) and to compare ‘cookbook’ with multimedia labs in physics courses (Royuk & Brooks, 2003).

A decade later, Mulford and Robinson (2002) developed the Chemistry Concepts Inventory (CCI). 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 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.
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 visualisations for learning, but rather seek evidence of the relative benefits for conceptual development of teaching approaches in science employing scientific visualisations \textit{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 visualisations and ‘traditional’ classroom teaching methods for the purpose of helping high school students to develop particular scientific concepts. The research question can be stated as: \textit{Is teaching with the use of scientific visualisations more effective than traditional classroom teaching for supporting students’ conceptual development of specific concepts in (a) Physics and (b) Chemistry?}

Conceptual development on the part of students will be measured using conceptual knowledge tests based on the Force Concept Inventory (Hestenes, Wells & Swackhamer, 1992) and the Chemistry Concept Inventory (Mulford & Robinson, 2002). These instruments will be 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 two Inventories have been used extensively internationally and are well validated (Hestenes & Halloun, 1995). Each subject test comprises 12 multiple choice items, with four possible answers, and the distractors focus on the common misconceptions as identified in the Force Concept Inventory and the Chemistry Concept Inventory.

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

The teachers using the visualisations will be provided with teaching points to include, but are 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 will conduct classroom observations. The primary intent is to look at the teaching style – does the teacher demonstrate the visualisation on a projector screen, in a more transmissive style, or are the students interacting personally with the visualisation – is there groupwork and discussion during the learning, how large are the groups and are all members engaged? We will also be determining the amount of prior knowledge the students have before completing the pre-test – although the pre-test is to be taken at the beginning of the unit, due to the overlaps and interconnectedness of topics in the syllabus, students frequently have some previous exposure to the topic before it is formally studied. We will also note 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 will allow us to qualitatively determine the extent of the effectiveness of the visualisation as a teaching tool.

**Crossover Design**

A ‘crossover’ (Ratkowsky, Evans & Allldredge, 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 visualisations) 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 not used as a contrast with constructivist teaching, or as shorthand for lecture-and-notes only teaching. Teachers have been asked not to use other scientific visualisations 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 visualisations in their teaching to various extents). The comparison is therefore essentially one between ‘teaching with visualisations’ and ‘teaching without visualisations’.

Workshops will be conducted for the participating teachers. These will focus on supporting the teachers’ understanding and pedagogical use of the developed scientific visualisations. They will also help the participating teachers to compare and discuss their own understanding of the scientific concepts and elaborate their understandings. All of the participating teachers will teach their students both ‘traditionally’ and using scientific visualisations, 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 visualisations 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 will be 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 visualisations), 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). Our interest, however, 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.

This set of sequences will occur for both Physics and Chemistry students in 2009, allowing some comparison of the effectiveness of visualisations in these two science subject areas. Ethics clearance will be obtained from the Research Ethics Board at the University of Queensland and ‘stakeholder permission’ from the state and private school systems involved in the study sought. Parents and students will be given information letters and asked to sign consent forms, as will participating teachers and the principals of participating schools.

In addition, results will be analysed against the sex of participating students, their score on a simple learning styles inventory 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. It is hoped that these analyses will yield more finely detailed information about particular student groups for whom particular modes of teaching may be more or less effective.

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

**A Practical Constraint**

The original intention was to choose, and directly compare, two concepts in Physics and two concepts in Chemistry. The new Queensland Physics and Chemistry syllabuses, however, focus on teaching scientific knowledge in real
world contexts, and do not prescribe a particular order in which concepts should be taught. Teachers do not typically use a common textbook or anything else that would tend to dictate the order of topics, and particular topics are not specific to Year 11 or Year 12. This meant that it was very difficult – in fact impossible – to find two Physics concepts that all of the participating teachers would be teaching to Year 11 students during the year, and the same issue applied to Chemistry. The crossover design absolutely requires that the comparison be between the same teacher-and-students group across the two conditions (visualisation and no-visualisation). This meant it was impossible to study the same teacher with a Year 11 class for one topic and a Year 12 class for the other.

As a response to this constraint it became necessary to add a third concept in each of the two subject areas. It became possible in each subject to at least find 2 of the 3 concepts that teachers were teaching to their Year 11 students. This makes the statistical analysis of the data somewhat more difficult and reduces the n for each of the conditions, but it will still be able to develop good quality quantitative evidence for the relative effectiveness of science teaching with and without visualisations. Table One shows the concepts studied in Physics and Chemistry and the visualisations chosen.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Concept</th>
<th>Visualisations</th>
</tr>
</thead>
</table>
| Physics | Straight line accelerated motion | http://kcvs.ca/kinematics/motion1d/motion_1d.swf  
                        |                                                                          | http://techtv.mit.edu/videos/831-strobe-of-a-falling-ball  
                        |                                                                          | http://www.bravus.com/visual/strobes.htm  
                        |                                                                          | http://www.bravus.com/visual/strobe2.htm |
| Newton’s first law | http://www.bravus.com/visual/puckkick.htm  
                        | http://www.bravus.com/visual/satellite.htm  
                        | http://phet.colorado.edu/simulations/sims.php?sim=The_Ramp |
                        | http://dev.cpo.com/home/portals/2/Media/post_sale_content/bounce.swf  
                        | http://www.science-animations.com/support-files/explosions.swf |
| Chemistry | Le Chatelier  | http://www.mhhe.com/physsci/chemistry/essentialchemistry/flash/techv17.swf |
| Thermochemistry | http://www.bravus.com/visual/bondenthalpy.mov  
                        | http://schools.matter.org.uk/Content/Reactions/BondEnergy.html |
| Intermolecular forces | http://www.kentchemistry.com/links/bonding/bondingflashes/bond_types.swf  
                        | http://faculty.washington.edu/dwoodman/LondonForces/dswnmedia/LondonForcesW.html  
                        | http://faculty.washington.edu/dwoodman/IntrFrcs/dswnmedia/IntrFrcsW.html  
                        | http://www.chm.davidson.edu/romitt/che115/Phase/Phase.htm |

Table One – Selected science concepts and visualisations

Conclusion
Scientific visualisations are believed to have the potential to improve students’ learning of science concepts, potentially improving students’ access to and outcomes in university science programs and careers in science. This project will evaluate that potential to see whether it is realised in students’ understanding of scientific concepts.

The software, hardware and teaching skills required to use scientific visualisations in teaching represent a significant investment of money and time on the part of the community. The results from this project will help in analysing the educational benefits of visualisations, providing important information for cost-benefit analyses.

If the findings of the research are that scientific visualisations do yield significant benefits in supporting students’ development of scientific concepts, the adoption of visualisations by science teachers offers potential advantages in the recruitment, retention and achievement of science students, an area of significant recent community concern.

References


