

# EFFECTIVENESS OF SCIENTIFIC VISUALIZATIONS IN YEAR 11 CHEMISTRY AND PHYSICS EDUCATION

David Geelan  
Griffith University, Gold Coast

## Abstract

*The use of computer-based tools in education is booming, and governments are pouring large amounts of funding into providing these tools for teachers, while the evidence for their educational effectiveness is still being gathered. This study compared the effectiveness of scientific visualizations - computer-based animations and simulations - with teacher's 'normal' teaching without visualizations, in terms of students' development of key chemistry and physics concepts. In a cross-over research design, the same classes - students and teacher - completed one concept using scientific visualizations and one concept without, while other classes reversed the sequence. Over all, there were no significant differences in conceptual development between the students taught with and without visualizations. While this is in a sense a null finding, there is a significant literature that indicates that students enjoy learning with visualizations, so given that they enhance enjoyment and 'do no harm' in terms of learning, there is a case to be made for their use in teaching. That case should probably not, though, at least based on these results, be made on the basis of extravagant claims about enhancing student learning.*

## Introduction

Computer-based scientific visualisations, including animations and simulations, are increasingly an everyday part of senior secondary school physics and chemistry education nationally and internationally (Park, et al., 2008; Phillips, Norris & McNab, 2010). They are being used increasingly in classrooms to help students gain access to phenomena and processes that are more difficult to address using teacher talk and static diagrams such as those found in textbooks or on whiteboards or presentation screens. The Australian National Curriculum for Science (ACARA, 2010) notes: "Digital aids such as animations and simulations provide opportunities to test predictions that cannot be investigated through practical experiments in the classroom and may enhance students' understanding and engagement with science." The Digital Education Revolution project (DEEWR, 2008) involves over a billion dollars of spending to enhance computer facilities and internet access in schools, in part to facilitate the use of these kinds of tools.

Books such as those by Gilbert (2005) and Gilbert, Reiner and Nakhleh (2008) have collected research around these new tools, and a body of research is developing. There is (and has been for some time) good evidence that students enjoy learning with scientific visualizations and are more engaged in their learning when using them (e.g. Annetta et al., 2009; Cifuentes and Hsieh, 2001; Delgado & Krajcik, 2010).

There is less evidence, however, in relation to the educational effectiveness – in terms of supporting students' development of the key scientific concepts that are an important goal of science education – of scientific visualisations as used in classrooms, and the evidence that exists is mixed (e.g. Frailich, Kesner & Hoffstein, 2009; Lee et al., 2010). This study used a quasi-experimental quantitative design to compare students' conceptual learning with scientific visualisations with their learning when teachers taught without visualisation tools. Data on students' sex and level of academic achievement were also collected to explore whether there were differential effects for particular students.

## Method

The overall study was quantitative in approach and quasi-experimental in design. The project used a modified crossover (Ratkowsky, Evans & Alldredge, 1993) design.

There are a number of difficulties involved in conducting experimental or quasi-experimental research in school classrooms, however we are committed to classroom-based evaluations because we believe it is essential that research in science education serves the profession as directly as possible (Hirschhorn & Geelan, 2008). These difficulties include challenges with random assignment of students to experimental and control groups when they are already in established classes, and the almost insurmountable challenges of finding classes that are well enough matched to be compared with one another in an experimental design.

Crossover designs help to meet these challenges by essentially making each class-and-teacher unit into its own control group. This is done by having each class complete one teaching sequence with and one without the innovation – in this case the scientific visualisations. Results are then compared for the same group of students between the situation when they learned with scientific visualisations and when they did not.

Under the crossover design of the study (Ratkowsky, Evans & Alldredge, 1993), students essentially serve as their own ‘controls’, since each student is represented in both the ‘experimental’ treatment group – learning with visualisations – and the ‘control’ group – learning without visualisations. The groups are therefore perfectly matched for academic ability, gender balance and other factors, because the same individuals are in each group. The students also completed both trials with the same classmates and the same teacher. Some students completed the visualisation trial first and the no-visualisation trial some months later, and others completed the trials in the reverse order, reducing the effects of maturation on the part of students.

It would be ideal from an experimental perspective if the students could be taught the exact same content in each instance, but this is impossible both in terms of human learning – once something has been learned once, learning it again is a dramatically different experience – and due to the constraints of honoring teachers and students’ time in class. For this reason different concepts – of comparable conceptual difficulty – were used, but under the crossover design some groups of students studied each concept using visualisations and some studied it without visualisations. Each possible combination of conditions and topics was therefore addressed.

One consequence of the wide range of levels of support and availability for the use of ICTs in schools was that students accessed the visualisations in a variety of different ways. We had suggested to teachers that the ideal approach in most cases was 2-3 students to a computer, interacting with the visualisations and each other and recording results. In some schools the computer labs were arranged in such a way that it was much easier to have students work on one computer each. In others it was impossible to get a computer lab so the teacher displayed the visualisation on a data projector screen at the front of the classroom and the class worked through the activities as a whole. The groups were not large enough for us to be able to conduct quantitative analyses of the differences between these different modes of delivery. Stephens, Vasu and Clement (2010) studied the specific issue of differences between small-group and whole-class use of visualisations in physics learning, and found no significant differences between the situations. One future avenue for research will be to focus in a more naturalistic, qualitative way on the ways in which teachers and students work and learn with visualisations in their own particular contexts, given their own particular sets of interests and constraints.

Typically the visualisations were not particularly complex or ‘high tech’, but involved students in actively manipulating variables and exploring the effect of these changes on the motions being demonstrated. The present study was quantitative in approach, and did not look closely at issues like the complexity and ‘distraction value’ of particular visualisations, only at their educational

effectiveness.

When the teachers were *not* teaching with scientific visualisations, they used the normal range of teaching strategies available to them, with the exception of visualisations (which a number of the teachers in the study already used consistently in their ‘normal’ teaching). This was not limited to teacher-centred lecturing but included classroom discussions, diagrams on the board, teacher explanations, demonstrations using apparatus or other approaches to introducing students to the identified scientific concepts.

Approaches to using the visualisations were not dictated to the teachers, and as discussed above, technical and infrastructure constraints would have made it difficult to standardize approaches anyway, however general suggestions were made about possible students activities using the computer-based tools. These suggestions focused on interactivity and exploration, including longer interactions.

Teaching sequences for each topic, with or without visualisations, involved about one week of instruction, typically 3-5 lessons. These are quite brief ‘interventions’, but were focused quite tightly on teaching single concepts, and the tests focused on students’ understanding of these concepts.

Conceptual development on the part of students was measured using conceptual knowledge tests based on the Chemistry Concept Inventory (CCI)(Mulford & Robinson, 2002) and the Force Concept Inventory (FCI)(Hestenes, Wells & Swackhamer, 1992). The tests were designed to distinguish the extent to which students developed the ‘correct’ scientific concept in relation to a topic, rather than any of a number of possible ‘misconceptions’. The tests for each of the three topics in each subject area – used as both pre-test and post-test – comprised 12 multiple choice items, with four possible answers, and the distractors focused on the common misconceptions as identified in the Chemistry Concept Inventory and Force Concept Inventory.

While evaluating the effectiveness of learning with scientific visualisations for all students is valuable, it is also plausible that this teaching approach might be more or less effective for particular students. Two additional characteristics of students were identified anonymously by the participating teachers for the research team: the sex of the students and their academic rank within the class. On the second measure, teachers were asked to rank students on a class list in terms of whether they fell in the top, middle or bottom third of their class. The class lists were then anonymised so that only test results, sexes and ranks were given to the research team, not student names.

Teachers gave their informed consent to participate, and students and parents (the students were minors) also signed consent forms to participate after being informed about the research project.

## Chemistry Study

A total of 129 Year 11 Chemistry students from 11 different classes in 7 different Brisbane area secondary schools, some public and some private, participated in the study.

Specific concepts that appear in the Queensland Year 11 Chemistry syllabuses were chosen for the study. The three concepts chosen were Le Chatelier’s Principle (and dynamic chemical equilibria more broadly), Intermolecular Forces (and other interparticle forces) and Thermochemistry.

One or more web-based visualisations were chosen for each concept. We chose to use existing resources that were available on the net. This may have led to the use of less directly comparable visualisations in terms of approach and style, but we felt that it allowed us to model more closely what really happens in school classrooms. Here are the links to the visualisations used:

### *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://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](http://schools.matter.org.uk/Content/Reactions/BE_enthalpyHCl.html)

## Physics Study

Students in six Year 11 physics classrooms (students aged 15-17) in four Brisbane-area high schools in Australia participated in the physics section of the overall study. There were six teachers and a total of 80 students involved. Two of the four schools were co-educational government schools and the other two were private girls' schools. There were 39 male and 41 female students in the sample.

The three topics chosen were Newton's First Law, Straight Line (Accelerated) Motion and Momentum.

Examples of the kinds of visualisations used include:

[http://phet.colorado.edu/simulations/sims.php?sim=The\\_Ramp](http://phet.colorado.edu/simulations/sims.php?sim=The_Ramp) (for Newton's First Law – from the PhET group at the University of Colorado)

[http://kcv.s.ca/nonpublic/kinematics/motion1d/motion\\_1d.swf](http://kcv.s.ca/nonpublic/kinematics/motion1d/motion_1d.swf) (for Straight Line Motion - from the King's Centre for Visualisation in Science)

[http://qbx6.ltu.edu/s\\_schneider/physlets/main/momenta3c.shtml](http://qbx6.ltu.edu/s_schneider/physlets/main/momenta3c.shtml) (for Momentum - from Lawrence Technological University)

The results from both studies were also pooled and these combined data are discussed below.

## Results

It is important to note throughout these results that the 'visualisation' ('treatment') and 'no-visualisation' ('control') groups are the same students, with the same teachers, being tested on different topics with the different treatments. A comparison of topic difficulty was made for the three chemistry topics, however students in different classes were treated with both treatments on all three topics, so systematic differences in the difficulty of the topics are negligible.

All scores are in marks out of 12 and are in terms of conceptual knowledge gains (posttest minus pretest).

## Chemistry Study

The three chosen topics were considered by the participating teachers to be of approximately equal conceptual difficulty. There were 99 students who completed the Thermochemistry topic, 111 who completed Equilibrium and 48 who completed Intermolecular forces (this is a total of 258, since each of the 129 students completed two topics). Table 1 shows the gain scores for the three topics.

Table 1 – Difficulty comparison of chemistry topics

Topic	Gain	
	Mean	SD
Thermochemistry (n=99)	1.72	2.76
Equilibrium (n=111)	2.04	2.79
Intermolecular Forces (n=48)	1.60	2.08

A one-way ANOVA shows that the differences are not statistically significant ( $F(257)=.594$ ,  $p=.55$ ), suggesting that in fact the topics are not significantly different in terms of their difficulty for student learning.

Table 2 shows the means for the students under the visualisation and no-visualisation teaching conditions.

Table 2 – Visualisation vs no-visualisation, all chemistry students/topics

Treatment	Gain	
	Mean	SD
No visualisation (n=129)	1.74	2.67
Visualisation (n=129)	1.92	2.65

It is almost unnecessary after looking at those results, but a two-tailed independent-samples t-test shows no significant difference in the learning gains between the two treatments ( $t(256)=-.538$ ,  $p=.59$ ).

Table 3 shows the mean gain scores for male and female students learning with and without visualisations. By inspection the means for female students are almost identical. The means for male students are different to look at, but a t-test shows that the differences are not statistically significant ( $t(54)=-1.35$ ,  $p=.18$ ).

Table 3 – Visualisation vs no-visualisation, chemistry, related to student sex

Treatment	Sex	Gain	
		Mean	SD
Male (n=28)	No visualisation	1.75	2.08
	Visualisation	2.54	2.27
Female (n=101)	No visualisation	1.74	2.82
	Visualisation	1.75	2.74

Table 4 shows the gain scores for students in the lowest, middle and highest achieving thirds of their classes, learning with and without visualisations.

Table 4 – Visualisation vs no-visualisation, chemistry, related to student achievement

Treatment	Achievement Rank	Gain	
		Mean	SD
Lowest (n=34)	No visualisation	1.24	2.13
	Visualisation	1.26	2.87
Middle (n=72)	No visualisation	1.89	2.91
	Visualisation	1.82	2.44
Highest (n=23)	No visualisation	2.04	2.60
	Visualisation	3.22	2.63

Means for the lower and middle thirds are very similar across the treatments, particularly given the size of the standard deviations. Results in the highest-achieving group appear to exhibit a larger

difference, however a t-test shows that the difference is not statistically significant ( $t(44)=-1.522$ ,  $p=.14$ ). Cohen's  $d$  yields a score of .45, which is a medium-sized effect, however the small sample size (23 students) and lack of statistical significance means this result should be treated with considerable caution. It is plausible that it is the most able students who can most effectively make use of multiple representations, but further research is required to explore this issue.

## Physics Study

An initial comparison – and the ‘headline’ finding of this project – can be made between the learning gains for the students when learning the physics concepts with and without visualisations.

Table 5 shows this comparison for the physics students. Scores are in marks out of 12. While the two means look quite different by inspection, the standard deviations are large, indicating a broad spread of knowledge gains. A two-tailed t-test shows that the difference is not statistically significant ( $t(158)=-1.58$ ,  $p=.116$ ). That is to say, neither teaching ‘treatment’ is significantly better than the other, when all participating students are considered together.

Table 5 – Visualisation vs no-visualisation, all physics students/topics

Treatment	Gain	
	Mean	SD
No visualisation (n=80)	.95	2.22
Visualisation (n=80)	1.53	2.38

Table 6 lays out the results analysed in terms of the sex of the participants.

Table 6 – Visualisation vs no-visualisation, physics, related to student sex

Treatment	Sex	Gain	
		Mean	SD
No visualisation (n=80)	Male (n=39)	1.00	2.52
	Female (n=41)	.91	1.90
Visualisation (n=80)	Male (n=39)	2.15	1.81
	Female (n=41)	.93	2.71

A t-test comparing male and female students within the visualisation group shows a difference significant at the .05 level ( $t(78)=2.37$ ,  $p=.02$ ). That is to say, male students benefited equally with female from the no-visualisation case but benefited significantly more than female students from learning with visualisations.

Statistical significance is only one measure of the effectiveness of a teaching innovation, however. Effect size measures such as Cohen's  $d$  give some sense of the magnitude of the learning gains achieved. A modified form of Cohen's  $d$  can be calculated by dividing the difference between the means of the two groups by the mean of the standard deviations of the groups. This gives a sense of ‘by how many standard deviations’ the innovation has improved learning. For the visualisation groups, for male students vs female students, this form of  $d$  is equal to .54. This is a medium effect size.

In terms of the degree to which students at differing levels of academic ability learned with and without visualisations, Table 7 summarises the learning gains (out of 12) for the three ranked groups.

Table 7 – Visualisation vs no-visualisation, physics, related to student achievement

Treatment	Achievement Rank	Gain	
		Mean	SD
No visualisation	Lowest (n=15)	.67	2.35

(n=80)	Middle (n=40)	.98	2.36
	Highest (n=25)	1.08	1.96
Visualisation (n=80)	Lowest (n=15)	2.07	2.76
	Middle (n=40)	1.27	2.26
	Highest (n=25)	1.60	2.36

A one-way ANOVA for the three groups learning with visualisations shows no significant difference between the mean gain scores in this group ( $F(79)=.615$ ,  $p=.54$ ). Similarly, for the no-visualisation group there is no significant difference ( $F(79)=.165$ ,  $p=.85$ ). Learning with or without visualisation yielded no significant learning differences between the academic achievement groups.

### Combined Data

In terms of ‘clean’ (no missing data) cases from the combined studies there were 157 participating students (34 male, 123 female). 79 students’ data appears from the physics study along with that of 78 students from the chemistry study (yielding closely balanced student numbers from the two subject areas. Since each student completes one topic with visualisations and one without, that yielded a total of 314 data points.

Table 8 shows the mean scores for all students across both subject areas when learning without and with visualisations.

Table 8 – Visualisation vs no-visualisation, all students, chemistry and physics

Treatment	Gain	
	Mean	SD
No visualisation (n=157)	1.19	2.26
Visualisation (n=157)	1.58	2.39

Given the size of the standard variations, it seems likely that these differences would not be statistically significant, and a t-test ( $t(512)=-1.48$ ,  $p=.14$ ) bears out this impression.

### Discussion and Conclusion

In a sense these are rather disappointing findings: while for the purposes of the research project I entered the study agnostic as to the educational effectiveness of scientific visualisations, they are teaching tools I value and have been working with for some time. Friends and colleagues have invested a lot of energy in developing and promoting visualisations for learning, and the claims that have been made in relation to their effectiveness have not always been modest.

Having said that, the claims made on the basis of this study also need to be modest – the interventions were quite short, and there were a number of variables – such as the quality of the visualisations used the specific ways in which students interacted with them – that needed to be better controlled in order to make stronger and more generalizable claims. There will also need to be further, qualitative, research into the specific ways in which students and teachers use scientific visualisations and the pedagogies that are most effective for their use.

With those caveats in mind, however, the results are clear: in these classrooms, students gained no more benefit from scientific visualisations than from more traditional modes of classroom teaching involving teachers explanations, static diagrams on the board and class discussions. They also learned no less, however, and on the Hippocratic principle of ‘first do no harm’, this suggest that at least the evidence does not make a case **against** using visualisations in classrooms. And, with studies showing that while Australian students achieve very well on international measures of science knowledge, they are among the least enthusiastic and engaged science students in the world (Fensham, 2006), the evidence that students find learning with visualisations enjoyable and engaging (Annetta et al., 2009;

Delgado & Krajcik, 2010) may in itself provide a solid motivation for teachers to use them.

The finding in the chemistry study, which was very weak and needs considerable further investigation – that the students achieving at the highest level in their studies also demonstrated the highest learning gains when using visualisations – is intriguing. In terms of equity of educational opportunities this wouldn't be an especially encouraging finding: visualisations seem to offer more help to those already doing best. It is possible (if the finding is real) that it is a result of the extent to which the more academically capable students already understand the concepts and are able to use the visual representations to test and enhance their understanding. Less able students who are still developing an understanding of the concepts may misinterpret features of the visualisations, or not be able to mentally bridge between the representation and their developing mental models. Much closer analysis is required to distinguish between these possibilities, and to explore whether different pedagogical approaches might not, after all, allow scientific visualisations to enhance the learning of students who are struggling.

Similarly, although girls have made significant gains in the past decade, their achievement and participation in physics in Australia still lags that of boys (Rennie, 2010). The finding, which is more robust, that boys studying physics gained more benefit from learning with visualisations than girls seems similarly to give more to those who already have most. This is another finding, however, that needs considerable further investigation in order to explain the cause(s).

Further research is required to support or challenge the present findings in more tightly-controlled studies and a wider range of contexts, however these results suggest that the educational use of scientific visualisations – and the costly infrastructure required to deliver them – should possibly not be 'sold' to governments and funders on the basis of extravagant claims about learning gains, but on their other affordances.

Further research is also required – and is currently being planned by the research team that conducted this project – that uses qualitative measures such as interviews with teachers and students, classroom observations and 'think-aloud' protocols to gain deeper insights into the pedagogies associated with the use of these tools. It may be that there are particular approaches that **do** yield significant conceptual learning gains but that these were combined with less-effective approaches in the present quantitative study to yield the no-difference findings reported here.

Acknowledgements: Professors Brian Martin and Peter Mahaffy at The King's University College, Canada, and Dr Michelle Mukherjee at Queensland University of Technology were invaluable colleagues in this project, which was also supported by ARC Discovery grant DP0878589.

## References

- Annetta, L.A., Minogue, J., Holmes, S.Y. & Cheng, M-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers and Education*, 53(1), 74-85.
- Australian Curriculum, Assessment and Reporting Authority (ACARA). (2010). *Australian National Curriculum: Science* (Foundation to Year 10).
- Cifuentes, L. & Hsieh, Y-C.J. (2001). Computer graphics for student engagement in science learning. *TechTrends*, 45(15), 21-23.
- Delgado, C. & Krajcik, J. (2010). Technology Supports for Science Learning. In P. Peterson, B. McGaw & E. Baker (Eds.), *International Encyclopedia of Education*. Elsevier: Atlanta, GA.
- Department of Education, Employment and Workplace Relations (DEEWR) (2008). Success through partnership: Strategic Plan to guide the implementation of the Digital Education Revolution initiative and related initiatives.
- Fensham, P.J. (2006). Student interest in science: The problem, possible solutions, and constraints. Paper presented at the Research Conference of the Australian Council for Educational Research. [Online: (accessed 5 May 2012) [http://www.acer.edu.au/documents/RC2006\\_Fensham.pdf](http://www.acer.edu.au/documents/RC2006_Fensham.pdf)]
- Frailich, M., Kesner, M. & Hoffstein, A. (2009). Enhancing students' understanding of the concept of chemical bonding by using activities provided on an interactive website. *Journal of Research in Science Teaching*, 46(3), 289-310.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher* 30: 141-166.
- Hirschhorn, M. & Geelan, D. (2008). Bridging the research-practice gap: research translation and/or research transformation. *Alberta Journal of Education*, 54(1): 1-13.
- Gilbert, J.K. (2005). *Visualization in Science Education*. Springer: Dordrecht.
- Gilbert, J.K, Reiner, M. & Nakhleh, M. (2008). *Visualization: Theory and practice in science education*. Springer: Surrey.
- Lee, H-S., Linn, M.C., Varma, K. & Liu, O.L. (2010). How do technology-enhanced inquiry science units impact classroom learning? *Journal of Research in Science Teaching*, 47(1), 71-90.
- Mulford, D. R., & Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education*, 79(6), 739-744.
- Park, J., Sandhu, M., Ritt, H., Pipes, D. & Myers, J. (2008). Developing, testing, and implementing dynamic visualizations: integrating multiple representations of video, data, and graphs while investigating physical science concepts. In K. McFerrin et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2008* (pp. 4759-4765). Chesapeake, VA: AACE.
- Phillips, L.M., Norris, S.P. & McNab., J.S. (2010). Visualisations and science. *Models and Modelling in Science Education*, 5(2): 63-74.

Ratkowsky, D.A., Evans, M.A. & Alldredge, J.R. (1993). *Cross-over experiments: design, analysis and application*. New York: Marcel Dekker.

Rennie, L.J. (2010). Gender Still Matters in Australian Schooling. *International Journal of Gender, Science and Technology*, 2(1). [Online: (accessed 5 May 2012)  
<http://genderandset.open.ac.uk/index.php/genderandset/article/view/79/76>)

Stephens, A.L., Vasu, I. & Clement, J.J. (2010). Small group vs. whole class use of interactive computer simulations. Comparative case studies of matched high school physics classes. Paper presented at the annual conference of the National Association for Research in Science Teaching, Philadelphia, March 21-24, 2010.