MEASURING LEARNING IN COMPLEX LEARNING ENVIRONMENTS

MICHAEL J. TIMMS

Australian Council for Educational Research

Dr Michael Timms directs the Division of Assessment and Psychometric Research at ACER, which develops high-quality assessments and conducts cutting-edge research in educational measurement. He is a recognised leader in the development of innovative ways to assess students in electronic-learning environments. His research is widely published in peer-reviewed journals and he was awarded the 2013 Journal for Research on Science Teaching Award by the National Association for Research on Science Teaching.

Before joining ACER, Dr Timms was Associate Director of the Science, Technology, Engineering and Mathematics (STEM) Program at WestEd, a pre-eminent educational research and development organisation in the USA. He led large-scale research studies in STEM education, with special focus on computer-based assessment projects, especially through the SimScientists research program (http://www.simscientist.org). He has been involved in the development of two assessment frameworks for the US National Assessment of Educational Progress for which he was awarded the Paul Hood award for excellence in educational research at WestEd.

Dr Timms has experience in leading evaluation research projects for other educational research grant recipients, such as universities, and has managed large-scale item development projects across many content areas. He is knowledgeable about the education systems of Australia, the USA and the UK.

ABSTRACT

Technology offers the opportunity to enhance the learning experience through providing students with learning environments that bring to them other worlds outside the classroom. For example, the use of animations, simulations and augmented reality can help to show dynamic processes such as geological events over time, virtual chemistry laboratories or events from history in deeper and richer ways than are possible in textbooks. These technological tools also offer the chance to allow students to explore and manipulate the virtual environments that are created, bringing opportunities for learners to engage in the construction of knowledge rather than just receiving facts. But, as the learning environments become more complex and the number of paths that students can take through them increases, how can teachers be assured that their students are learning what was intended? How can we measure learning in such a way that ensures students get feedback at the right time and teachers remain in touch with how their students are progressing? This session explores how learning can be traced in complex learning environments that use technology and illustrates the techniques from several projects that have been developed to do that.

HOW INTERACTIVE LEARNING ENVIRONMENTS CAN ASSIST STUDENT LEARNING

Interactive learning environments hold a lot of promise for assisting learners in ways that are tailored to the needs of each learner. Well-designed interactive learning environments combine pedagogical approaches that are based on cognitive theory of learning in interactive ways in electronic environments with methods of measuring the progress of learners and techniques for providing assistance at key moments.

This paper focuses on how interactive learning environments can support student learning in science, a curriculum area in which there is an increasing emphasis on understanding scientific concepts and on developing skills in applying science inquiry practices. In science, students often have difficulty connecting concepts to real world phenomena and in understanding how to use scientific practices in investigating those phenomena (TIMSS, 2008). Studies in the USA point to the lack of 'rigorous and excellent' instruction in US schools on science inquiry skills – those that build students' ability to form ideas or hypotheses about phenomena and to design experiments to test those ideas (Weiss & Pasley, 2004).

This paper demonstrates how three interactive learning environments, which were designed for instructing students in developing their understanding and science inquiry practices across several areas of science, dealt with the challenges of supporting learning of complex concepts in interactive ways using technology. These are the three interactive learning environments:

- ChemVlab+ (http://www.chemvlab.org) – an interactive learning environment in which secondary students work with a virtual chemistry laboratory to undertake tasks in a series of embedded assessment modules that provide them with opportunities to apply chemistry knowledge in meaningful contexts and to receive immediate, individualised tutoring. The four modules cover concentration, unit conversion, molar mass, balancing reactions and using stoichiometry.
- SimScientists (http://www.simscientists.org) – a suite of modules that use simulations to enrich science learning and assessment for students in middle school and secondary school. Science simulations can be used in curriculum activities as embedded, formative assessments and as summative assessments. The SimScientists modules cover topics in life science (ecosystems and cells; human body systems), physical science (forces and motion; atoms and molecules) and Earth science (climate; plate tectonics).
- Voyage to Galapagos – The Voyage to Galapagos provides middle school students with an interactive learning environment in which they can follow in the footsteps of Charles Darwin by doing simulated exploration of a selection of the Galapagos Islands. Students collect and then analyse data on iguanas to arrive at specific connections among the key concepts of variation, function and natural selection.

THE CHALLENGE OF PROVIDING ASSISTANCE IN INQUIRY SCIENCE INSTRUCTION

The goal of inquiry learning is to allow students to induce the characteristics of a domain through their own experiments and exploration (de Jong, 2006). But, even in curricula with hands-on laboratories and the opportunity to engage in inquiry learning, students are typically asked to replicate standard experiments rather than perform their own inquiries. Critics of such approaches say they are limited to 'transmitting' science rather than teaching its practices (Duschl, Schweingruber & Shouse, 2007). This pedagogical approach is likely to contribute to the reported difficulties students have in designing and conducting scientific experiments; for instance, by varying more than one variable at a time (Keselman, 2003), by incorrectly interpreting data (Lewis, Stern & Linn, 1993) and by sticking with preconceived ideas in the face of contradictory data (Chinn & Brewer, 1993, 2001).

On the other hand, a variety of research has suggested that, *with appropriate guidance*, students can learn about science and successfully engage in scientific inquiry, including taking the well-established steps followed by professional scientists, such as making hypotheses, gathering evidence, designing experiments and evaluating hypotheses in light of evidence (Chen & Klahr, 1999; de Jong & van Joolingen, 1998; Klahr & Dunbar, 1988; Lehrer & Schauble, 2002; Njoo & de Jong, 1993). Theory about how best to scaffold inquiry learning has also emerged (Edelson, 2001; Quintana et al., 2004). Building on these fundamental findings and theory, a variety of researchers have developed simulations, cognitive tools and scaffolding to support the kind of reasoning that underlies inquiry learning in science. Research on scaffolded inquiry learning suggests that teaching the critically important skills associated with scientific inquiry can be greatly improved if supported by the right kind of guidance (Linn & Hsi, 2000; Sandoval & Reiser, 2004; Slotta, 2004; van Joolingen, de Jong, Lazonder, Savelsbergh & Manlove, 2005; White & Frederiksen, 1998).

But what exactly is the right amount and type of guidance? While past work with inquiry learning environments makes clear that *some* guidance is necessary, it doesn't fully answer this question, which in the learning sciences more generally has been variously investigated under the guise of 'desirable difficulty' (Schmidt & Bjork, 1992), the 'assistance dilemma' (Koedinger & Aleven, 2007) and 'productive failure' (Kapur, 2009). Essentially the issue is to find the right balance between, on the one hand, full support and, on the other hand, allowing students to make their own decisions and, at times, mistakes. There are cost benefits associated with each end of this spectrum. *Assistance giving* allows students to move forwards when they are struggling and to experience success, yet can lead to shallow learning, non-activation of long-term memory and the lack of motivation to learn on their own. On the other hand, *assistance withholding* encourages students to think and learn for themselves, yet can lead to floundering, frustration and wasted time when students are unsure of what to do. Advocates of direct instruction point to the many studies that show the advantages of giving assistance (Kirschner, Sweller & Clark, 2006; Mayer, 2004), but this still does not acknowledge the subtlety of exactly how and when instruction should be made available, particularly in light of the differences in domains and learners (Klahr, 2009).

Grappling with the assistance dilemma requires, at least in part, an understanding of the human cognitive architecture. It is well established in cognitive science that humans have both a working memory, where conscious processing occurs, and a long-term memory, where our extensive experience and knowledge resides (Atkinson & Shiffrin, 1968). Long-term memory is critical to what we 'know' – unless an educational activity changes long-term memory, we have not learned anything. Further, learning is subject to the severe limitations of working memory (Sweller 2003, 2004), both in capacity (estimated to be a very small number of elements: three to seven) and duration (unrehearsed information disappears within 30 seconds). When students are confronted with new content in an unfamiliar environment, such as an inquiry-learning tool, their working memory is easily and quickly overloaded unless strong guidance is provided to focus them on relevant information and tasks. As students become more familiar with the material and environment, through transfer of information to long-term memory, they are typically able to focus on the right content and choose the correct steps to take without as much guidance and without experiencing cognitive overload.

Not surprisingly, in light of this theory, studies of how human tutors deploy both the frequency and the nature of assistance have shown that effective tutors adapt their support based on the ability level of the student. Katz, Allbritton and Connelly (2003) found differences in the feedback tutors gave to students who had (unknown to the tutors) scored low on a pretest versus those who scored well. The differences in the frequency and nature of the assistance provided was based on the tutor's perception of the relative abilities – and therefore needs – of each student.

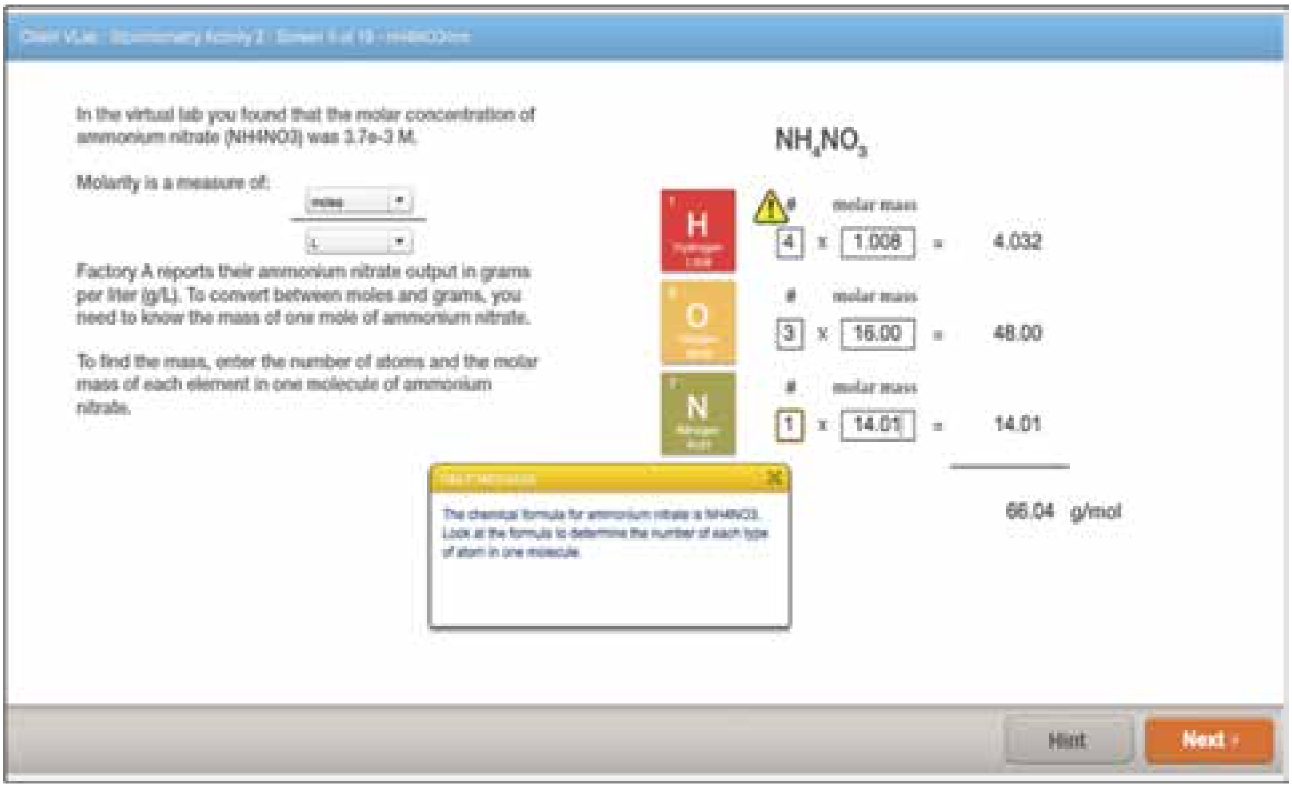

Figure 1 Screenshot that shows how ChemVlab+ provides feedback and coaching to students

EXAMPLES OF INTERACTIVE LEARNING ENVIRONMENTS

The three interactive learning environments employ two different techniques to detect students' need for help and to deliver assistance as they complete the tasks they are set. The ChemVlab+ and SimScientists projects use contingent-based modelling in which the systems are designed to detect when students are making errors or behaving in ways that are known to be unproductive. When these contingent behaviours are detected, the system is designed to flag the error and offer a sequence of hints that lead the student to a productive solution. An example from the ChemVlab+ is shown in Figure 1.

The feedback that student receive is differentiated based on their needs. When a student makes a response and clicks on the 'Next' button in the bottom right of the screen, the system evaluates the student's work on that screen through applying a logic structure that determines the correctness and, if incorrect, the nature of the misconception that the student has. Figure 1 shows how the system provides a symbol (! in a triangle) where a hint is available, and the hint text that the student has been given. A student may also call for a hint by pressing the 'Hint' button, but only receives it when the system judges that a hint is needed.

The Voyage to Galapagos is a more open-ended learning environment and employs a more complex system to detect a students' need for help. It uses a Bayesian network to represent the contingent-based model, which is a way of keeping tally of actions that the student takes which suggest that he or she is not on a productive learning path. For example, in Figure 2 the student is part-way through a task in which he or she has to photograph a sample of iguanas that show variation in their physical traits. The panel on the lower right shows a map view of the path the student is taking around the island and the main panel on the left shows the view as the student follows that path. An iguana is in the bottom of the view. If the student needs more iguanas in the sample, but moves on without taking a photo, the system detects this and passes that data to the Bayes Net. Each such incident increases the probability that the student needs help with data collection and, if the student continues to pass by iguanas, the system eventually will prompt the student by flagging the missed iguana and indicating that it needs to be added to the sample. Our research study in Voyage to Galapagos is looking at what mixture of assistance is best for which kind of learner and the Bayes Net system can be used to trigger a range of levels of help.

Figure 2 Screenshot that shows how a student task in Voyage to Galapagos can provide data on when a student needs help.

Table 1 Gaps in total performance between English learners or students with disabilities and the general population

Group	Ecosystems post-test (%)	Force and motion post-test (%)	Ecosystems benchmark (%)	Force and motion benchmark (%)
English learners	24.0 (n = 123)	27.4 (n = 50)	10.6 (n = 126)	13.6 (n = 50)
Students with disabilities	20.2 (n = 183)	15.7 (n = 153)	8.4 (n = 189)	7.0 (n = 153)

IMPACT

Results from trials of the SimScientists and ChemVlab+ modules indicate that the kinds of feedback built into the systems are producing learning gains and, more interestingly, that they might benefit particular students.

In a study of two of the SimScientists modules, the use of interactive assessments produced higher outcomes compared to performance on traditional multiple-choice assessment items (Quellmalz, Timms, Silberglitt & Buckley, 2012). Overall, students performed better on the interactive assessments than on the multiple-choice post-test, and performance gaps between both English-language learners and students with disabilities compared to other students were reduced on the interactive assessment. Table 1 compares performance gaps of both these student groups to a reference group of all other students.

The gaps between the focal groups and the reference group are comparatively smaller than for the post-test. This indicates that the multiple representations in the simulations and active manipulations may help English-language learners and students with disabilities to understand the assessment tasks and questions and to respond.

In a study of the ChemVlab+ modules, we were interested in whether the activities produced learning overall, as well as whether the schools with differing student demographics benefited similarly from the instructional activities. School A was in a low-income area in which almost half the students qualified for free or reduced-price lunches and only 26 per cent of students had scored at proficient level on the state science test. School B had 20 per cent of students eligible for free or reduced-price lunches and 40 per cent of students were proficient on the science test. School C was in a wealthier area in which only 8 per cent of the students qualified for free or reduced-price lunches and 70 per cent were proficient in science. Students took a pre- or post-test that comprised 15 multiple-choice and open-ended items with a maximum score of 30 points.

Figure 3 shows that for schools A and B, post-test scores were significantly higher than pre-test scores. At School A, where a higher proportion of students were disadvantaged, overall scores were lower, but the change from pre- to post-test was higher. The average of the post-tests was 13.4 while the pre-test average was 9.4 ($p < 0.001$, $t = 9.86$, $n = 102$ students), which represents an effect size of 0.68 (Cohen's d). The second-highest gains were at School B, which had a moderate proportion of disadvantaged students. At School B, the average of the post-tests was 15.6 while the pre-test average was 13.0 ($p < 0.001$, $t = 6.75$, $n = 147$ students), an effect size of 0.48. For School C, where there were hardly any disadvantaged students, there was a gain from 15.84 at pre-test to 16.4 at post-test ($p < 0.2$, $t = 1.1$, $n = 81$), but the difference was not significantly different. This indicates that the interactive learning materials seemed to have an increased effect for disadvantaged students.

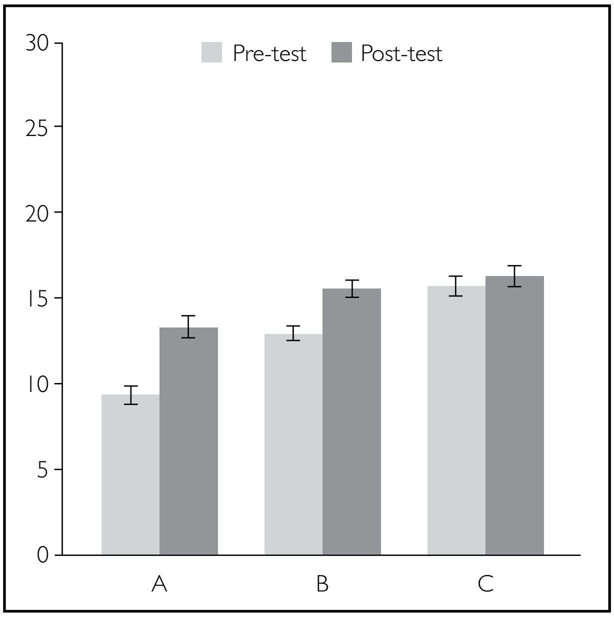

Figure 3 Comparison of the pre- to post-test learning outcomes for three schools in the ChemVlab+ pilot study (error bars indicate one standard error) (Davenport, Rafferty & Timms, 2013)

At the time of writing this paper, we have not yet pilot tested the Voyage to Galapagos learning environment.

Overall, the use of interactive learning environments appears to have differential effects that enable students who are disadvantaged, are not native English speakers or have disabilities that affect their learning to improve their performance relative to their peers.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence, (Eds.), *The psychology of learning and motivation* (vol. 2, pp. 89–195). New York. NY: Academic Press.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70*, 1098–1120.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1–49.

Chinn, C. A., & Brewer, W. F. (2001). Models of data: A theory of how people evaluate data. *Cognition and Instruction, 19*, 323–393.

Davenport, J., Rafferty, A., & Timms, M. J. (2013). *Exploring the usability and feasibility of a virtual chemistry lab tutor for high school science*. Paper presented at the Annual Conference of the American Educational Research Association, San Francisco, CA.

de Jong, T. (2006). Computer simulations – Technological advances in inquiry learning. *Science, 312*, 532–533.

de Jong, T., & van Joolingen W. R., (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179–201.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K–8*. Washington, DC: National Academies Press.

Edelson, D. (2001) Learning-for-use: A framework for the design of technology-supported inquiry activities. *Journal of Research in Science Teaching, 38*(3), 355–385.

Kapur, M. (2009). The role of productive failure in mathematics teaching and learning. In B. Kaur, B. H. Yeap & M. Kapur (Eds.), *Mathematical problem solving* (pp. 43–68). Singapore: World Scientific.

Katz, S., Allbritton, D., & Connelly, J. (2003). Going beyond the problem given: How human tutors use post-solution discussions to support transfer. *International Journal of Artificial Intelligence in Education, 13*, 79–116.

Keselman, A. (2003), Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching, 40*, 898–921.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Klahr, D. (2009). 'To every thing there is a season, and a time to every purpose under the heavens'; What about direct instruction? In S. Tobias and T. M. Duffy (Eds.), *Constructivist theory applied to instruction: Success or failure?* (pp. 291–310). New York, NY: Taylor & Francis.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1–55.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*, 239–264.

Lehrer, R., & Schauble, L. (2002). Symbolic communication in mathematics and science: Co-constituting inscription and thought. In E. D. Amsel & J. Byrnes (Eds.), *The development of symbolic communication* (pp. 167–192). Mahwah, NJ: Lawrence Erlbaum Associates.

Lewis, E. L., Stern, J. L., & Linn, M. C. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, 33*, 45–58.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Hillsdale, NJ: Erlbaum.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*, 14–19.

Njoo, M., & de Jong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching, 30*, 821–844.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching, 49*(3), 363–393.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G. et al. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*, 337–387.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*, 345–372.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217.

Slotta, J. (2004). The web-based inquiry science environment (WISE): Scaffolding knowledge integration in the science classroom. In M. Linn, E. A. Davis & P. Bell (Eds.), *Internet environments for science education* (pp. 203–233). Mahwah, NJ: Erlbaum.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (vol. 43, pp. 215–266). San Diego, CA: Academic.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science, 32*, 9–31.

TIMSS. (2008). *Technical Report*. J. F. Olson, M. O. Martin, & I. V. S. Mullis, (Eds.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

van Joolingen, W. R., de Jong, T., Lazonder, A.W., Savelsbergh, E., & Manlove, S. (2005). Co-lab: Research and development of an on-line learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*, 671–688.

Weiss, I. R., & Pasley, J. P. (2004). What is high quality instruction? *Educational Leadership, 61*, 24–28.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*, 3–11.