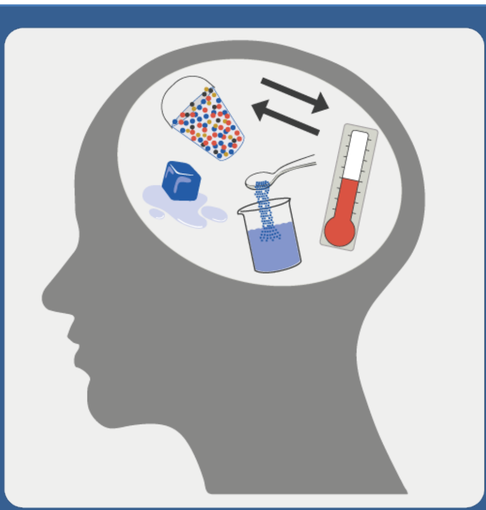


Enhancing the secondary-tertiary transition in chemistry through formative assessment and self-regulated learning environments.

Final Report 2014

The University of Queensland (Lead Institution)

Project Leader: Gwen Lawrie

Project Team: Chantal Bailey, Aaron Micallef, Anthony Wright (The University of Queensland); Tim Dargaville, Madeleine Schultz (Queensland University of Technology); Roy Tasker, Mark Williams (University of Western Sydney); Simon Bedford, Glennys O'Brien (University of Wollongong).

Author: Gwen Lawrie

< https://shire.science.uq.edu.au/chembytes/Index.html#/Home/Main>

Support for the production of this report has been provided by the Australian Government Office for Learning and Teaching. The views expressed in this report do not necessarily reflect the views of the Australian Government Office for Learning and Teaching.

2014

ISBN 978-1-76028-118-2 [PRINT]
ISBN 978-1-76028-119-9 [PDF]
ISBN 978-1-76028-120-5 [DOCX]

Acknowledgements

Dr Christopher Thompson and Mr Hayden Dickson (Monash University), while not formally a partner institution, became an integral part of the project team. They contributed extensively to the collation and intellectual development of concept items and the outcomes of the project would not have been as rich without their involvement and their students participation.

The willingness of the participant students, studying across all five institutions, to share their perceptions of their thinking, of feedback and of their motivation is acknowledged along with their patience during the development and trials of the online modules.

The ChemBytes website was inspired by the UQ-based 5 Minute Physics web activities (Margaret Wegener and Timothy McIntyre). The creative input of the learning and graphic designers based at UQ are acknowledged (Tania Ryan, Marnie Holt and Dominic McGrath), in creating the ChemBytes web platform for our online activities. Finally support from Trevor Daniels (Osiris Software) in implementing web analytics is also acknowledged.

Md Abdullah Al Mamun, whose UQ PhD project has evolved from this work, is acknowledged for his substantial contribution to the development of the online module activities. Hanisah Hassan and Ashleigh Henders, two UQ undergraduate research project students, are also acknowledged for their contribution to the collection and analysis of data that informed this project.

We thank Supriya Sant for her support in the analysis of qualitative data for the project.

Many academics across Australia and the UK have shared rich feedback in regard to the concept diagnostic items which we are grateful for and we particularly thank Professor Simon Lancaster (University of East Anglia) for engaging with and supporting our project activities through dissemination in the UK.

Finally, we would like to thank and acknowledge the expertise, advice and mentoring generously provided by our external evaluator, Professor Carmel McNaught. Carmel became a critical friend to the project team and her experience, both in evaluation and in chemistry education, had a significant impact on the activities, progress and outcomes of the project.

List of acronyms used

CCI	Chemical Concept Inventory
CI	Concept Inventory
CTT	Classical Test Theory
FCI	Force Concept Inventory
HS	High school
IRT	Item Response Theory
LMS	Learning Management System
LTi	Learning Tools interoperability
MSLQ	Motivated Strategies for Learning Questionnaire
NSW	New South Wales
Qld	Queensland
OLT	Office for Learning and Teaching
QUT	Queensland University of Technology
TLO	Threshold Learning Outcomes
UoW	University of Wollongong
UQ	The University of Queensland
UWS	University of Western Sydney
VIC	Victoria

Executive summary

The perspectives of the different stakeholders frame the transition between secondary and tertiary chemistry learning environments. From the students' perspective, they are entering a new world that is very different to their high-school environment. At many Australian tertiary institutions, incoming students experience very large lecture theatres with classes of hundreds of students, and the delivery and assessment of content is markedly different to high school. From the instructors' perspective, assumptions are made in regard to the prior learning experiences of these students and what they already know. Also, instructors hope that students have skills to navigate the learning activities in order to construct understanding with minimal guidance. What is missing is a shared awareness between instructors and students of the diversity of prior experiences and understanding and the provision of resources to support students who are at risk of limited success in their studies due to their incomplete conceptual understanding of chemistry.

This project brought together a team of academics who are deeply involved in the teaching and curriculum decisions surrounding first-year chemistry at five Australian institutions across three states, while not formally a partner institution, Christopher Thompson from Monash University became an integral part of the project team. This provided a rich lens in terms of the diversity of preparation for the secondary–tertiary transition with three separate secondary chemistry syllabi. A combination of diagnostic tools, formative feedback options and a range of strategies for delivering face-to-face or self-regulated online study modules were developed. The modules have integrated dynamic visualization and simulation tools to challenge, adjust or reconstruct conceptions in core chemistry concepts. Insights into the nature of alternate or missing conceptions, timing and delivery of formative feedback and engagement with online resources were collected through an extensive evaluation process

Recommendations for practice:

- Diagnostic concept instruments should only be used by instructors to gain insight into the range of students' understanding in core chemical concepts. This information should then be used to provide formative feedback to students and direct them to resources or activities that enable them to adjust alternate conceptions or misconceptions. Completion of the diagnostic test should not be associated directly with summative assessment as students will not engage authentically in considering their thinking but rather will seek the answers for marks.
- In contrast, summative marks are required to engage students in the online learning activities (extrinsic motivation). Without these, very few students will persist in completing these in a self-regulated manner simply with a goal of improving their existing understanding (intrinsic motivation).
- Students engage more in acting on formative feedback when it is supplied as close as possible in timing to lecture or laboratory learning activities that are based on directly related concepts.

Outcomes:

The principle outcomes of this project are:

- A profile of the existing chemistry conceptions held by first-year chemistry students as they transition between secondary and tertiary contexts. This represents benchmarking data that can be used to inform anticipated learning outcomes and curriculum design.
- Practitioner recommendations for delivering formative feedback and the implementation of online resources to address the alternate conceptions possessed by students in large, first-year chemistry courses.

Deliverables:

The deliverables for this project include:

- Two diagnostic concept instruments tailored for students dependent on whether they have completed senior high-school chemistry.
- Five web-based online learning modules, located on the project ChemBytes website: https://shire.science.uq.edu.au/chembytes/Index.html#/Home/Main. These resources are open access and are also available as printable pdf resources containing embedded hyperlinks that can be used as an alternative to the web-based modules.
- Five case studies demonstrating multiple strategies for the delivery of formative feedback and learning resources in multiple contexts.
- A set of benchmarking data for student outcomes as they transition between the secondary and tertiary contexts.

Table of Contents

Tables and Figures

Tables

Figures

1. Introduction

1.1 Rationale and objectives

Large first-year STEM courses represent a new and perhaps a complex learning environment for students as they commence their tertiary studies, particularly in terms of classes that are very large, individual access to instructors and the higher expectation that they are independent and can engage in self-regulated study. Chemistry courses (units) form part of the progression for many separate science, engineering and health-related programs of study (e.g. engineering, biotechnology, materials science, agricultural science, medicine, pharmacy, dentistry, biomedical science and health sciences). A typical Australian tertiary first-year chemistry cohort is characterised by high enrolments (ranging between 300 and 1500 students) and comprises a diverse range of academic abilities, interests, and motivations for learning chemistry. Many students will have completed senior high-school chemistry studies; however, many students also enrol in tertiary chemistry courses without this prior knowledge.

Enhancement of the secondary-tertiary transition and the first-year experience can be achieved through provision of a learning environment and curriculum based on shared expectations between the instructor and students as well as proactive and timely access to learning support (Kift, 2009; Kift et al, 2010). Constructivist learning environments are most effective when the learner and instructor are both aware of the learners' existing conceptual models enabling student to extend and apply their new understanding rather than resort to rote learning (Taber, 2001). It is also important to assist students to address conceptual deficiencies and increase their skills thereby maximising the effectiveness of student learning (Pitkethly & Prosser, 2001). This process requires measurement or diagnosis of students' existing conceptual understanding, followed by provision of formative feedback, access to remediative resources and further opportunity to apply new knowledge (Nicol & Macfarlane-Dick, 2006; Carless et al., 2010; Schraw et al., 2006).

A significant body of research exists relating to student conceptions, alternate conceptions and missing conceptions, and a number of evaluation instruments have been developed to characterise the range of conceptions that exist in any cohort (Mulford & Robinson, 2002; Potgieter et al., 2008; Pavelich et al., 2004; Treagust et al., 2011). However, there has been little attempt to apply such instruments to inform transitional curriculum reform. The template for the new national curriculum does not identify core concepts that enable constructive learning; rather, all concepts that are deemed significant across all aspects of chemistry are addressed.

One application of concept inventories that is of particular interest to us is the potential diagnostic data that can be used to identify students at risk of low achievement or failure in tertiary chemistry studies (Potgieter & Davidowitz, 2010; Potgieter et al., 2011; Bell, 2011). Having identified students that are at risk in a cohort, an appropriate learning intervention is required, one that does not place additional load on the students. There is a distinct absence of such interventions reported in the literature. A versatile, low resource-intensive intervention is required, which can be readily translated by chemistry academics into any teaching context; for example, collaborative self-directed activities (Sandi-Urena et al., 2010). In this project, the team investigated how best to enable students to self-diagnose flaws in their existing conceptual models through interactive modules, which supply formative feedback. Self-correction of these models can be achieved through activities that generate cognitive conflict (dissonance) based on the nature of their misconception. This represents a novel approach to simultaneously addressing the student diversity in conceptions in chemistry and facilitating students to develop improved learning strategies

Two key objectives of this project have been:

- To design and demonstrate a technology-assisted intervention in the form of learning modules that supply formative feedback and use cognitive dissonance to challenge students' misconceptions, thereby enabling reflection and adjustment of conceptions. These modules will be self-contained enabling an academic to select one 'off the shelf' and implement it in any one of a range of learning environments (laboratory, lecture, tutorials, self-directed study sessions).
- To change the culture of learning, assessment and feedback in first-year chemistry courses; in particular, the provision of formative assessment and feedback to enhance the transition from secondary chemistry studies

1.2 Processes

This project encompassed the development, implementation and improvement of practices informed by both research literature and project data as it emerged; hence it represents an example of design-based research. The timeframe of the project activities included three academic semesters involving five separate institutions and more than 5000 first-year chemistry students. The project team, who met weekly via Skype and face-to-face on four occasions, were fully engaged across the entire timeframe. A reference group was nominated for this project; however, as the project progressed there was substantial engagement with Australian and International academics who acted as an informal steering group. In particular, the dissemination activities engaged several international experts in related fields and they provided invaluable guidance to the team.

1.3 Deliverables

The deliverables from this project include:

- *Two diagnostic instruments* that assess students' concepts in chemistry entering tertiary studies. One is tailored towards students who have completed senior high-school chemistry (core concepts) and the second is tailored towards those who have not completed senior high-school chemistry (foundation concepts). These instruments can be delivered either as complete diagnostic tests or divided into concept clusters and delivered at the instructor's discretion in regard to timing during semester.
- *Open-access online web-based modules* designed to challenge alternate conceptions held by students by engaging them in active and interactive learning (ChemBytes URL: https://shire.science.uq.edu.au/chembytes/Index.html#/Home/Main). Modules contain dynamic simulations and animations sourced from existing high-quality educational initiatives.
- A *suite of case studies* that illustrate the multiple ways that an instructor can combine the first two deliverables into their practice, and recommendations for the provision of formative feedback. These represent practitioner exemplars.
- *Benchmarking data* from students across five institutions and three Australian states enabling comparison of cohorts and informing curriculum design.

2. Diagnostic aspects (concept item collation and development)

Chemistry concept inventories are well-validated multiple-choice tests that have been widely considered as a route to exploring students' existing conceptions (Schwartz & Barbera, 2014; Heredia et al., 2012; Barbera et al., 2013; Othman et al., 2008; Mulford & Robinson, 2002; Pavelich et al., 2004). These instruments are traditionally applied in chemical-education research to profile the nature and range of alternate conceptions that students hold. The concept-inventory community protect the instruments by not releasing the answers to students; hence, they do not represent a typical test whereby students receive the correct answer as feedback. Instructors are likely to apply concept-based diagnostic testing for a range of reasons including the preparedness of students for their tertiary studies in chemistry (Potgieter & Davidowitz, 2011; Heredia et al., 2012), and to measuring efficacy of a curriculum change (Krause et al., 2004).

A key objective that was intended for the diagnostic instrument in this project was the ability to cluster concept items that were related so that rich individual formative feedback could be provided to students. The project team established that clusters of five items would potentially be most effective and sought these by reviewing the substantial chemical concept literature.

2.1 Existing concept items

There has been sustained interest and research into the nature of student conceptions, alternate conceptions and missing conceptions in chemistry. The most common form of exploration of student conceptions is through interviews where students explain their thinking. From these studies, several evaluation instruments have been developed with the aim of characterising the range and nature of conceptions that are possessed amongst students across multiple age levels (Othman et al., 2008; Potgieter et al., 2008; Pavelich et al., 2004; Mulford & Robinson, 2002). The items that have been included in concept inventories to date examine well-known alternate and misconceptions related to the particulate nature of matter, phase changes, chemical equations, bonding and equilibrium.

There is an accepted process involved in the development and validation of authentic concept items (Schwartz & Barbera, 2014; Arjoon et al., 2013; Treagust, 1988; Treagust, 1986). Each item is evaluated through a series of steps, often compiled from multiple studies and contexts, beginning with student explanations of phenomena in interviews, followed by content validation by expert review and ending with trials of multiple-choice items to establish item response validity (Figure 2.1).

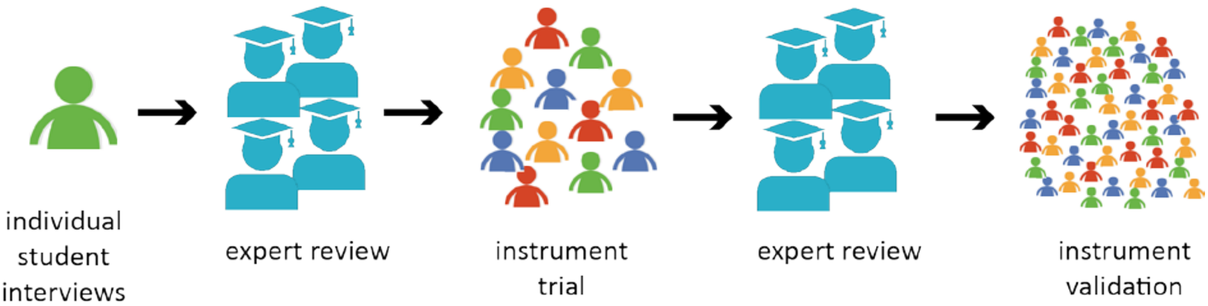


Figure 2.1 Schematic representation of an example process for the content and item response validation of concept items.

An example of the process of item evolution can be demonstrated for the famous item that explores students' conceptions of what exists within the bubbles of boiling water (Mulford

& Robinson, 2002). Interviews with children by Osborne and Cosgrove (1983) revealed the existence of a number of alternative conceptions where students believed that the bubbles contained air, oxygen or hydrogen. Bodner (1991) applied this question as a multiple-choice item in a test delivered to general chemistry students confirming that these alternate conceptions were held. These conceptions were explored further through a questionnaire and interviews by Johnson (Johnson, 1998) before the question appeared in Mulford and Robinson's 2002 multi-item inventory for alternate conceptions, now commonly referred to as the Chemical Concept Inventory (CCI). The reliability and validity of Mulford and Robinson's inventory has recently been explored further through classical test theory and Rasch analysis (Barbera, 2013). Multiple items were sourced from this CCI and also from other literature instruments and these are summarised in Table 2.1 below for the final instruments:

Table 2.1 Sources of concept items for the two diagnostic instruments used in this study.

Core Chemistry Instrument Items (Appendix C)	Foundation Chemistry Instrument Items (Appendix D)	Literature Source
1,2,7,11,12,13, 14,15,17,22	5,10,11,12,22,25	Mulford & Robinson (2002)
3,6,10,21	16,17	Pavelich et al. (2004)
4	13	Griffiths & Preston (1992)
	14	Osborne & Cosgrove (1983)
	2,3	Hadenfeldt et al. (2013)
	9,21	Othman et al. (2008)
	1,4,15,20	Yezierski & Birk (2006)
	18	Prince et al. (2012)
	6,7	Stains (2013)
5,8,9,16,18,19,20, 23,24,25	**8,19,23,24**	**Developed by project team: (**Bailey, Bedford, Dargaville, Dickson, Lawrie, Micallef, O'Brien, Schultz, Tasker, Thompson, Williams & Wright)**.**

After the first two semesters of trialling the instrument had been completed, it became evident that the core chemistry diagnostic was not appropriate for students who had not completed senior high-school chemistry. The project team began the process of developing a second related instrument (Table 2.1) to explore students existing understanding of concepts that are a foundation for senior high-school chemistry and these were based in the topics of behaviour of matter, nature of particles, states of matter, temperature and heat, and aqueous solutions. This latter instrument was only trialled in semester 1, 2014 and requires further validation.

2.2 New concept items

2.2.1 Item validity

The original intent of the project team was to identify and apply existing concept-inventory items but it became quickly evident that several of these did not fit the aims of the study and ten new items were required (Table 2.1). The process of developing new items was carefully considered and was informed by literature (Figure 2.1). Published studies were sought relating to the targeted concept to establish the nature of the alternate conceptions that are possessed by students; then the project team collectively, and iteratively, formulated a series of four to five potential responses. This process represented 'expert review' due to the number and experience of the project team members. The multiple-choice items were then trialled by students and the project team considered the distribution of responses, any item that did not discriminate student understanding (for example, where

>90% students gained the correct response or only two responses were opted for including the correct response) was revised or rejected.

The selection and formulation of questions emerged as a significant component of the project because the process stimulated rich discussion amongst the team members in regard to their pedagogical practices, how they explain concepts and a comparison of their students' general understanding of chemical concepts.

2.3 Instrument validation

The pilot core chemistry diagnostic instrument, comprising 25 concept items, was trialled across four institutions in semester 1, 2013, along with multiple additional demographic questions (refer to Appendices A and C for the full item sets). The resulting data was filtered for consent and completion, and compiled into a single data set to explore factors relating to the transition between secondary chemistry and tertiary chemistry concepts (there was no intent to explicitly compare students between individual institutions). Validation of the instrument included statistical analysis to gain descriptive statistics and application of item response theory and Rasch analysis (Pentecost & Barbera, 2013; Wei et al., 2012).

2.3.1 Item reliability

The use of an instrument to measure student conceptual understanding provides deeper information than simply whether the student has mastered a concept or not. The alternate answer that a student selects provides information in relation to their existing mental model; hence, the questions are carefully constructed. Classical test theory (CTT) and item response theory (IRT), based on whether a student's answer is correct or not, are methods to that enable consideration of the utility of a collection of items in testing students' understanding.

2.3.2 Rasch analysis

Rasch analysis extends the evaluation of the instrument to consider the degree of difficulty of questions in parallel with the number of items that an individual student got correct. The application of a probabilistic model enables the display of both the person response and item difficulty on the same linear scale known as a logit or log-odds scale, often referred to as a Wright map (Figure 2.2, next page). The logit scale is displayed between -4 and + 4 typically and the spread of items between the extremes of the range indicates the relative item difficulty (easiest questions have the lowest logit value). Therefore items that are discovered to be too difficult or too easy can be revised or discarded (Pentecost & Barbera, 2013; Wei et al., 2012). The distribution of student abilities can also be inferred with lower-ability students represented by lower logit values.

The pilot core concept diagnostic instrument (Appendix A) items were found to be distributed in the range of -2 to +2 (Figure 2.2) and this distribution was centred about zero indicating that the instrument had good reliability. The person response covered a wider range (+4 to -3) and was also centred about zero, confirming the spread of abilities amongst the participants; however, the Cronbach alpha value of 0.8 indicated that the instrument had good internal consistency (a value of 0.65 is widely regarded as the boundary value below which the internal consistency becomes less reliable).

The first implementation of the foundation concepts instrument (Appendix D) was completed in semester 1, 2014 at the University of Wollongong and Rasch analysis indicated that the instrument had good item and person response reliability with a Cronbach alpha value of 0.71.

After the trial of the pilot instrument in 2013, several items appeared not to have functioned in terms of discriminating students' alternate conceptions. These items were rephrased or substituted to develop the core concepts instrument, used in 2014 (Appendix C). This instrument was delivered either at a single time-point at the beginning of semester (UoW) or as a concept cluster that was delivered at the point of teaching related concepts during the semester.

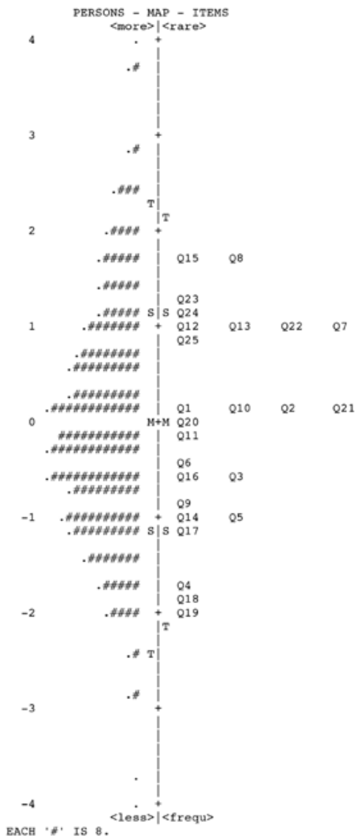


Figure 2.2: Wright map displaying the outcomes of Rasch analysis for the person response and item response for the trial core concept instrument (Semester 1, 2013). Each # symbol represents 8 individual students.

It is not considered good practice to compile data that includes clusters of questions that were collected at a number of separate time-points for Rasch analysis; the team investigated whether this was an issue by analysing the small separate clusters that were delivered at four institutions in semester 1, 2014. The data for the different Rasch measurements explored are provided in Table 2.2.

Table 2.2: Rasch statistical data for the semester 1, 2014 implementation of the core concepts instrument as a complete instrument and as separate clusters.

Item clusters included	UQ	UoW	Monash	UWS	Cronbach Alpha
All clusters (multiple weeks)	✓				0.72
Aqueous Solution + Equilibrium Clusters	✓	✓	✓	✓	0.47
Phase Change Cluster	✓				0.66
Heat & Energy Cluster	✓				0.37
Conservation of Matter Cluster	✓				0.33

Aqueous Solutions Cluster	✓	0.15
Equilibrium Cluster	✓	0.44

The internal reliability measured as a Cronbach alpha value for each of the separate clusters based on UQ data (Table 2.2) indicates that only the phase-change item cluster demonstrates a weak internal consistency; the remaining four clusters exhibit low internal consistency. This is not surprising because the phase-change items contain a common thread in the sense that they all relate to water, and hence are strongly interconnected concepts in a single context. For the remaining clusters, it was not possible to easily introduce a single context and thus this measure of validity should not be used in isolation. To complete a scale reliability analysis, the different responses need to be assigned numerical values and the alternate responses in concept items are assigned as zero. Further work is required to identify whether the student responses to alternate conceptions can be numerically ranked.

2.3.3 Comparing student outcomes from concept diagnostic instruments

Concept diagnostic instruments are often used as pre/post tests to assess students learning gains as a result of a teaching intervention or initiative. In this project, the aim was to use the student responses to provide formative feedback and then engage students in self-regulated learning. Hence, post-tests were not routinely conducted as part of this project, although there are some instances where an item was re-presented to students at a later time-point or in the summative exam.

The implementation of a single instrument across multiple institutions at the same time-point in 2013 offered the opportunity to explore a number of possible variables, including curricular differences, display of concept items (graphical vs textual) and other demographic variables that may potentially reveal differences (e.g. gender). Comparison of student outcomes was achieved through: t-tests (independent pairs); one-way ANOVA tests; and chi-squared tests. The outcomes are discussed further in section 5.2 below.

3. Module design and delivery

Once alternative conceptions have been identified, the challenge is to place students in learning environments where they can challenge their conceptions through alternative frameworks (Chi et al., 1994). The first step in the process for instructors to be able to develop constructive activities where learners actively develop sound mental models or adjust their conceptions requires the evaluation of students' existing conceptions and prior learning experiences. The introduction of cognitive conflict is a well-known strategy to address alternate, missing or misconceptions (Linenberger & Lowery Bretz, 2012). The design of activities that introduce cognitive conflict is an emerging chemical-education research field; the strategy taken in this project was to involve students in self-directed active learning using existing multimedia online visualisation resources.

3.1 Activity design

Multiple excellent web-based open-access resources that have been carefully designed to support students in their construction of concepts are available online. Three of these resources (PhET, Molecular Workbench and Vischem, URLs provided in reference list below) were integrated into online modules designed by the project team. The key elements of the activities that were identified by the project team were:

- Concise descriptions of concepts supported by graphical representations with directions for students to engage in an exploration of a simulation of the concept using Socratic statements.
- Multiple-choice review questions so that students could check their understanding and re-attempt the activities.

An example storyboard is provided in Appendix E, illustrating the composition that was typical of the web pages used in this study in terms of the balance of text and graphical representation. An attempt was made to introduce cognitive conflict through embedded Socratic questions and also through prompts to students in regard to the type of observations that they could make by interacting with the simulation. In each concept module, an option was also provided for students who possessed correct conceptions but were still seeking access to activities to challenge their thinking.

Two separate web-based initiatives were trialled during the module design during the course of the project. Student feedback was collected to inform subsequent adaptation of both the focus and content of these modules.

3.1.1 ReSOLv

The first design attempt sought to guide students through activities in an adaptive manner (Figure 3.1) using a series of web pages, each with a separate learning objective and content. Each student is directed to one of four sequences in the online activity based on the formative feedback that they received according to the incorrect answer that they has selected for concept items. These were arranged according to the following categories:

- Concept Builder (negligible understanding)
- Concept Fix (significant alternate conception)
- Concept Shift (minor alternate conception)
- Concept Quest (well-formed conception)

The University of Western Sydney hosted these web pages and students from all institutions were provided with login credentials to access the activities as part of their emailed feedback. Web analytics enabled tracking of student activity and progress.

Evaluation of the implementation across two semesters in 2013 (through questionnaires and focus-group interviews) provided substantial data in regard to how students accessed and persisted in the activities.

3.1.2 ChemBytes

Based on the evaluation of ReSOLv, several changes were made to the learning modules to formulate ChemBytes. They were moved to an open-access website and the activities were divided into self-contained pages addressing a specific conception. The 5 Minute Physics lecture preparation modules that existed at UQ informed the graphical design of ChemBytes and the expertise of a learning designer (Tania Ryan) was engaged in page design and graphics. Google Analytics was embedded on these web pages to collect information in regard to usage and other demographic data (individual access could not be linked to specific students). Five concept modules were released during semester 1, 2014 and are linked to the diagnostic concept clusters.

3.2 Mode of delivery: Online vs face-to-face

Through the experiences in the implementation of the modules at each institution, it became evident that a 'one size fits all' approach to delivering the learning modules was not viable; instructors required the ability to tailor activities to their own learning contexts and students. Therefore, the learning modules were evaluated in the context of delivery in both online vs face-to-face modes; the following were explored.

- Entirely online and independent-of-course teaching activities.
- Embedded within face-to-face teaching contact sessions (lectures or tutorials), either online or in the form of pdf instruction sheets.

The case studies in section 4.2 (below) provide several exemplars of the multiple ways in which an instructor could combine the diagnostic items and learning activities in their own contexts. There is no single recommendation for practice because this depends on the learning environment and the prior experiences of the student cohort.

3.2.1 Online self-directed modules

As part of the formative feedback that they received from the diagnostic concept instruments, students were directed to online modules based on which items they had incorrect and the nature of their alternate conceptions. Students could work through activities independently to suit their needs.

3.2.2 Face-to-face learning environments

The learning modules were integrated into learning activities in lectures and tutorials where students were supported through the guidance of the instructor (lecturer or tutor). Feedback was provided at the time of the activity to enable students to have a personalised link and direction in the activities.

3.3 Platforms

In the application of technology-based learning environments, there are two potential barriers to their success: transfer of resources between instructors and their learning management systems (LMS) and the platforms that students use to access the activities.

3.3.1 Challenges in the translation of different online modules between LMS platforms

The institutional LMSs that were considered were Blackboard™ and Moodle™ and the process of translating both diagnostic tests and the online resources were investigated. Initially, the web-based activities were not SCORM-compliant, nor able to be embedded as a tool with Learning Tools interoperability (LTi). It was quickly apparent that instructors would need support from their institutional IT personnel to embed the web activities within the LMS and this was beyond the scope of the project. The alternative was to place URLs within the LMS and then track whether students accessed these links in parallel with activity on the web pages monitored through Google Analytics.

3.3.2 Platforms that students used to access the ChemBytes web pages

Google Analytics enables insight into the types of technology that students adopted to access the online modules, including platforms, devices, operating systems and browsers that students used. Table 3.1 provides a summary for semester 1, 2014 and includes students across three institutions and two states.

Table 3.1: Platforms, devices and browsers that students used to access the online learning modules.

Category	Number of users	Distribution
Platform	2060	Desktop (95%) Tablet (3%) Mobile (1%)
Operating Systems	2067	Windows (72%) Mac/iOS (27%) Android (1%)
Browsers	2060	Chrome (53%) Safari (15%) Firefox (19.5%) Internet explorer (11%)

Of the mobile devices that were used to access the web pages, 78% were iPads or iPhones. According to Google Analytics, across the semester there were 6,948 sessions and 25,191 page views (some of these will be the instructors also accessing the resources but it was not possible to discriminate the identity of the user).

4. Instructors

4.1 Instructor rationale

An instructor sets about planning the learning activities in their courses based on the level and the perceived learning needs of their student cohort. Available resources including the learning environment and technological facilities typically restricted them. Armed with the suite of diagnostic and intervention resources, the instructors amongst the project team implemented a range of pedagogical strategies that were specific to their contexts in their courses in semester 1, 2014. This has resulted in a suite of teaching exemplars.

4.2 Exemplars

These different strategies have been illustrated graphically in the case study summaries below. The key to these graphics is provided in figure 4.1. The timing of delivery of the diagnostic questions, online activities and feedback is illustrated in the bottom right-hand diagram in each graphic.

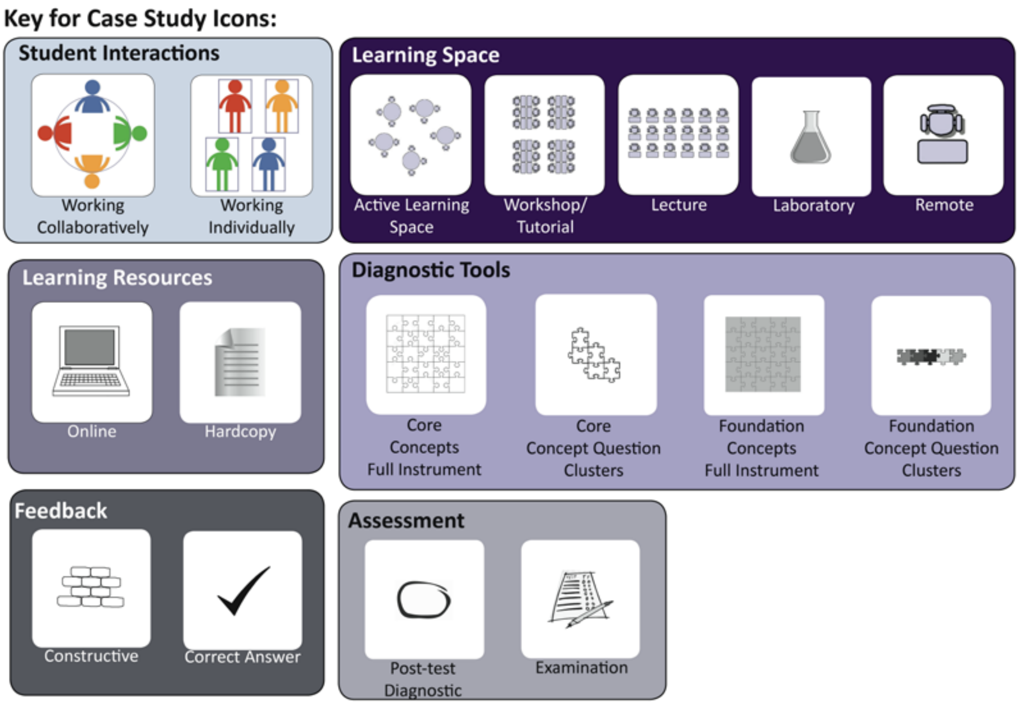


Figure 4.1: Key to the icons that are used to summarise the exemplars below.

4.2.1 The University of Queensland

The context was a large first-year chemistry service course where the instructors were faced with a challenging lecture timetable in which the sequence for all of the four lecture streams (sections) was Wednesday/ Thursday/ Friday. This impacted on the instructor's ability to set lecture preparation/ homework tasks, and so weekly online review modules (dubbed CROMs) were formulated, each containing one cluster of concept diagnostic items with formative feedback, links to online activities and finally a summative quiz.

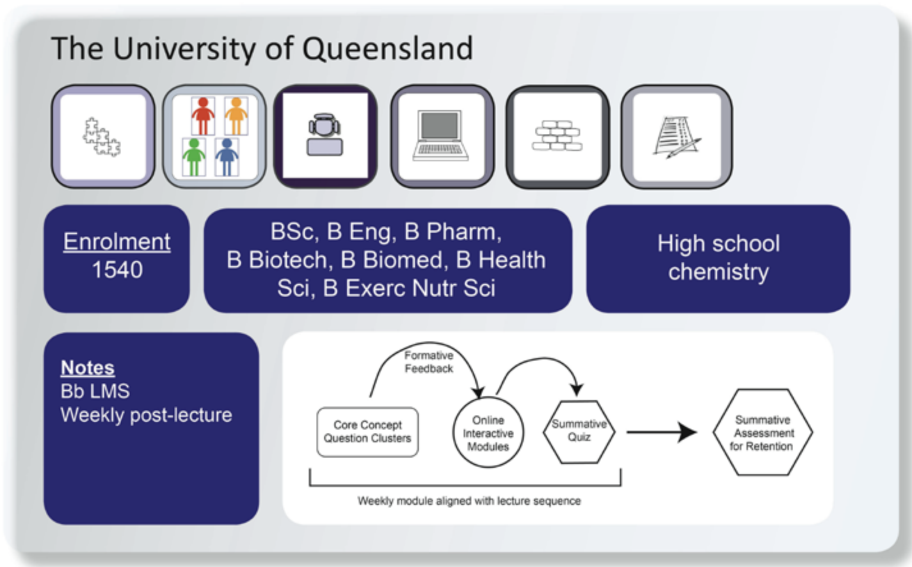


Figure 4.2: Graphical summary of the implementation at UQ in semester 1, 2014.

The CROMs were released every week across the semester in alignment with the lecture content, and hence seven concept review clusters, in addition to those developed through this project, were required. The instructor compiled these by sourcing items from the foundation chemistry instrument and literature. Evaluation of these activities included LMS and web analytics, concept-item outcomes, an end-of-semester questionnaire and focus-group interviews.

4.2.2 University of Wollongong

Two separate initiatives were implemented at UoW in two separate first-year courses; in Case 1, students had studied high-school chemistry as a prerequisite, and in Case 2 they had no prior high-school chemistry.

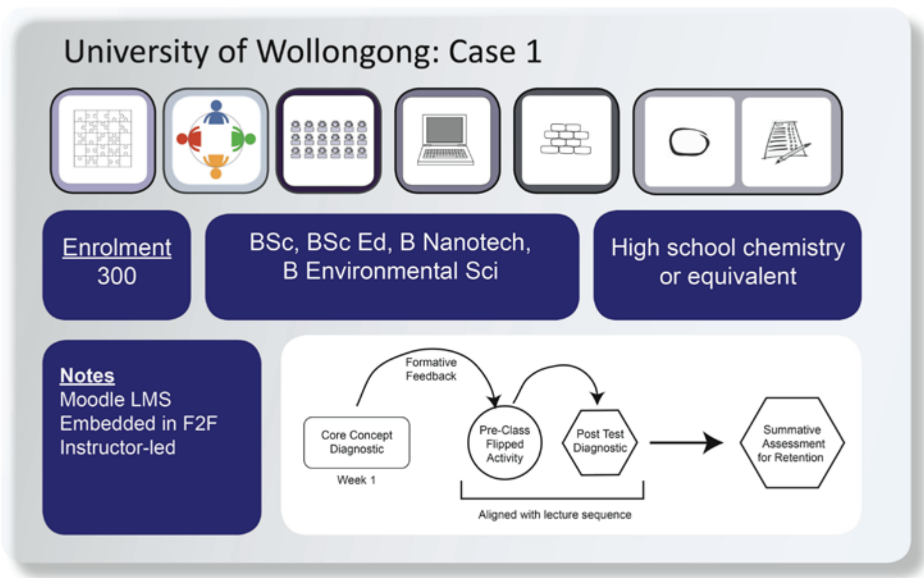


Figure 4.3: Graphical summary of the first example of implementation at UoW in semester 1, 2014.

The two instructors took different approaches to using the diagnostic instruments and providing formative feedbacks with one (case 1) providing the core chemistry instrument at the beginning of semester then delivering feedback at the point when related concepts were taught in lectures across the semester. In the second (case 2) the foundation chemistry instrument was delivered at a point mid-semester to catalyse lecture discussion. Evaluation of these initiatives has included concept-item outcomes, student engagement with in-class and online activities and focus-group interviews.

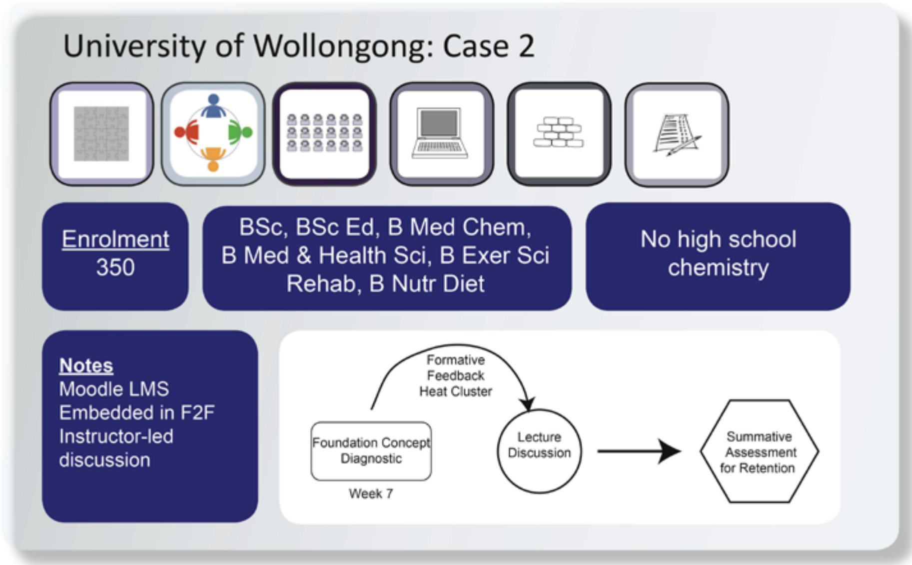


Figure 4.4: Graphical summary of the second example of implementation at UoW in semester 1, 2014.

4.2.3 University of Western Sydney

In this example, selected concept clusters from the core chemistry instrument were delivered in technology-enhanced tutorials as a pre-test; then related questions delivered as a post-test after students had completed online learning activities.

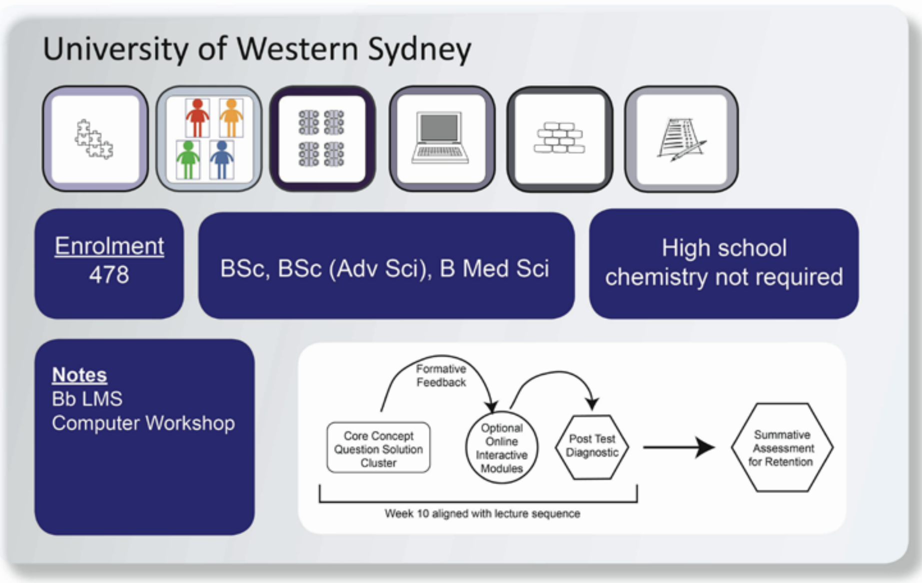


Figure 4.5: Graphical summary of the first example of implementation at UWS in semester 1, 2014.

Evaluation included concept-item outcomes and student engagement in online activities through analytics; focus-group interviews were also conducted.

4.2.4 Monash University

The diagnostic concept instrument was embedded as preparation for chemistry laboratory activities and clusters were selected to align with concepts that would be encountered in experiments.

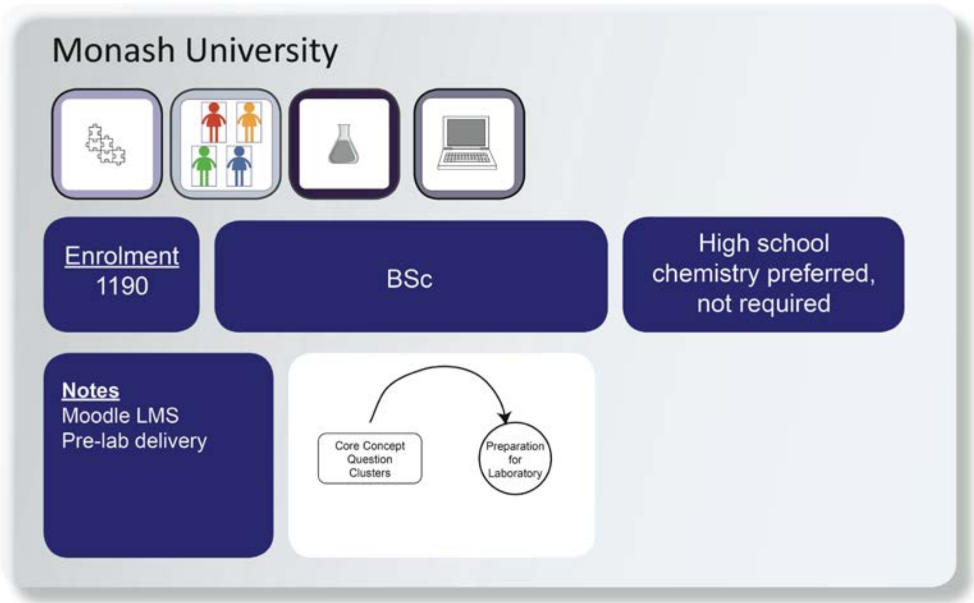


Figure 4.6: Graphical summary of the first example of implementation at Monash in semester 1, 2014.

Evaluation was completed through concept-item outcomes and focus-group interviews.

4.2.5 Queensland University of Technology

In semester 1, students are completing a core Science course and so delivery of the diagnostic concept instrument is not appropriate. However, the instructor sought to apply the underlying processes of this project in developing an innovative approach to engaging students in their thinking of chemistry concepts. Students were provided with the question stem for several concept items and interviewed their peers to assemble a set of alternate conceptions. They then created a video to address any alternate conceptions that arose.

A substantial amount of evaluation data has been collected across all five case studies and data analysis is ongoing, our aim is to publish recommendations for practice linked to downloadable resources through the project website.

5. Student engagement and learning outcomes

5.1 Evaluation framework

The evaluation framework applied in this project was developed in consultation with our external evaluator, Professor Carmel McNaught, and is based on the LEPO evaluation framework (Phillips et al., 2012). This framework separately examines the learning environment, learning process and student learning outcomes from the context of all the stakeholders (project team, instructors and students). The full plan and data that was collected throughout the project to inform the evaluation is summarised in Appendix F along with the final external evaluator's report. Data was collected according to the timeline displayed in Figure 5.1 below:

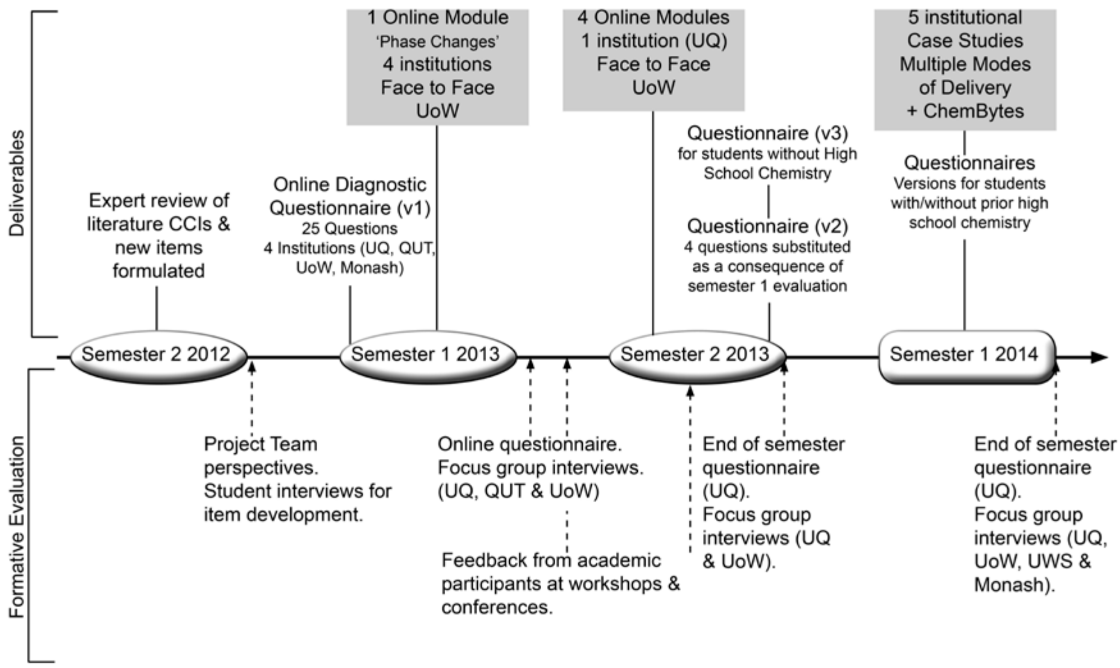


Figure 5.1: Timeline for the evaluation of activities and student outcomes in the project.

Ethics approval was gained for all of the project activities and data collection at all participating institutions.

5.2 Student conceptions and benchmarking

As we move towards measuring student outcomes across a program of study in terms of threshold learning outcomes (TLOs), it is important to assess what students' prior understanding of chemistry is as they enter tertiary studies. The diagnostic instrument developed in this project has provided data that can inform this benchmarking process since it derives from five institutions across three states. The aim was to develop a profile of students' conceptions as they entered tertiary chemistry studies. In all five institutions there were students enrolled in the participating first-year chemistry or science courses who had not completed senior high-school (HS) chemistry, even when this was a recommended prerequisite.

The data that was collected in weeks 1 and 2 of semester 1, 2013 enabled an exploration of several factors relating to students existing conceptions to develop this profile, including:

- The impact of not completing high-school chemistry (non-HS chemistry);
- Any impact of the state curriculum for high-school chemistry;
- Whether English as a second language impacted on completion of concept items;
- Any differences in students' conceptions based on gender.

These factors will each be discussed separately below.

5.2.1 The impact of high-school chemistry preparation

Students who had completed secondary chemistry as a senior subject achieved a higher percentage correct in all items in the instrument (Figure 5.2). There were two concept items where the level of achievement was not significantly different (Appendix B) between the students who had or had not completed high-school chemistry (Q7 and Q22).

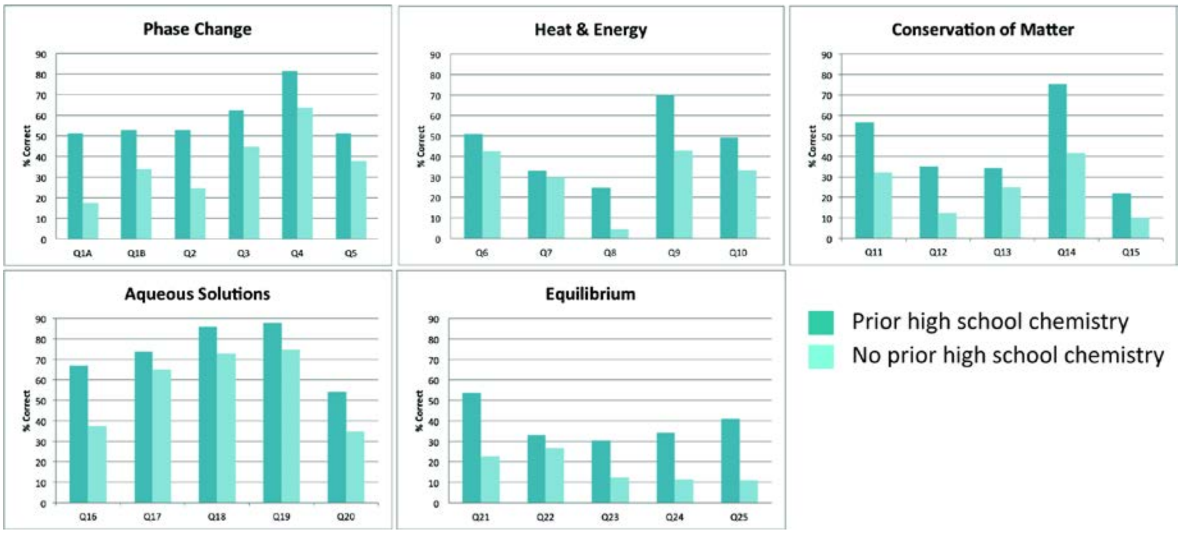


Figure 5.2: A comparison of the percentage of correct responses between students who had completed high-school chemistry and those who had not for the core concept diagnostic items in semester 1, 2013.

It is evident that a higher number of students possessed correct conceptions in items across the concept clusters of phase change and aqueous solutions, regardless of their prior learning in chemistry. These ideas had likely encountered as part of the junior science curriculum and form the foundation of senior high-school chemistry, whereas equilibrium involves a complex set of concepts that are linked together.

Item Q1 was prepared in two formats, text only and including a graphical representation of the processes to explore the potential impact of visualisation on students' conceptual understanding. There was no significant difference between the two formats for the outcomes for students who had completed high-school chemistry. However, an interesting outcome from this comparison was that a significantly higher number of students who had not completed high-school chemistry identified the correct answer for the graphical version of the item, indicating that the representation was more likely to have aligned with their conceptual model. There were only four items (Q1, Q2, Q12 and Q17) that contained a graphical display of the molecular level information in the core instrument. There was no other indication that this graphical format had enhanced students' ability to answer these items correctly as opposed to the text-based items.

5.2.2 Impact of state high-school chemistry curriculum

The students participating in this project were enrolled in tertiary institutions representing three separate Australian states: Queensland, New South Wales and Victoria. Australian students typically enrol in institutions in their state of origin and only a very low number of students travel to another state to study. Students who self-reported as having previously studied high-school chemistry where asked to identify which state they had completed their studies and an option of 'other' was provided for students who had studied outside these three states (including international students). A significant difference was found in comparing the overall test outcomes for NSW, Qld, Vic and students who identified 'other' as where they completed their high school chemistry (Table 5.1). Students who had studied in Victoria had the highest outcomes (ANOVA F statistic = 4.502; $p < 0.01$).

Table 5.1: Comparison of outcomes for the core concept instrument for students based on the state in which they had completed their high-school studies.

State	Number of Students	Mean	Standard Deviation
Qld	385	14.63	4.853
NSW	246	13.82	4.706
Vic	220	15.27	5.181
Other	246	13.95	4.873

These results should be treated with caution as they do not represent a comprehensive sample across all the tertiary institutions within each state and additional data collection is required.

5.2.3 The role of language in the function of concept items

Concept-inventory items often involve question stems and answers that are textual and require familiarity with terminology. It is possible that international students for whom English is a second language (ESL) may find the language used in a question difficult to interpret. This has impact on the utility of a concept diagnostic test in terms of whether the questions are testing a conception or the student's language skills. A comparison of the outcomes for students who self-reported as ESL with the rest of the population was completed to identify any items where there was a significant difference in achievement. It was found that there were significant differences in achievement for several items (Table 5.1). However, the majority of these items demonstrated that students with ESL achieved more highly than the students who identified as having English as their first language.

Table 5.2: Comparison of outcomes for the core concept instrument for students who have English as a second language.

Concept Item	Question topic	English as first language	English as second language	t statistic	Significance
5	Nitrogen gas and water intermolecular forces	0.74 (0.439)	0.82 (0.391)	-2.194	0.029
6	Heat can be described as	0.60 (0.491)	0.74 (0.442)	-3.752	<0.001
7	Hydrogen gas burning in air	0.30 (0.459)	0.49 (0.501)	-4.582	<0.001
8	Dinitrogen tetroxide bond breaking	0.22 (0.413)	0.40 (0.491)	-4.515	<0.001

9	Iron railway expanding	**0.71 (0.452)**	**0.63 (0.485)**	2.212	0.028
10	Plastic and metal temperature	0.47 (0.499)	0.59 (0.493)	-2.837	0.005
14	Nitroglycerine stoichiometry	0.74 (0.439)	0.83 (0.377)	-2.807	0.005
15	Burning match fate of matter	**0.23 (0.421)**	**0.16 (0.371)**	2.122	0.035
16	Calcium chloride solution dissolving	0.65 (0.477)	0.76 (0.425)	-3.157	0.002
17	Sugar solution dilution	**0.76 (0.424)**	**0.60 (0.492)**	4.176	<0.001
18	Salt solution evaporation	**0.87 (0.333)**	**0.78 (0.417)**	2.836	0.005

A comparison using an independent pairs t-test, revealed only 4 items where it was significantly likely that students who reported as having ESL achieved a lower outcome due to the construct of the question and these were Items Q9, Q15, Q17 and Q18. The text composition in items these questions do highlight some potential issues in the construct of concept items. In Q9 and Q15, the questions involve terminology, including ‘track’, ‘gap’, ‘fate’ and ‘ash’, which may not be immediately evident in their meaning to a student for whom English is a second language; this issue requires further investigation. In Q17 and Q18, the descriptive text framing the interpretation of the graphic is complex and, while a high % correct for both groups of students was achieved compared to other items, there may be an issue in clarity of meaning for this item.

5.2.4 Gender differences in outcomes for concept items

Students were asked to indicate their gender as part of the demographic question sets linked to the concept inventory (a response was not required). Analysis of achievement in each question on the basis of gender (Figure 5.3) revealed substantial differences with males gaining a significantly higher average % correct in all items except five (Q5, Q7, Q8, Q17 and Q20) where there was an insignificant difference or an equivalent performance.

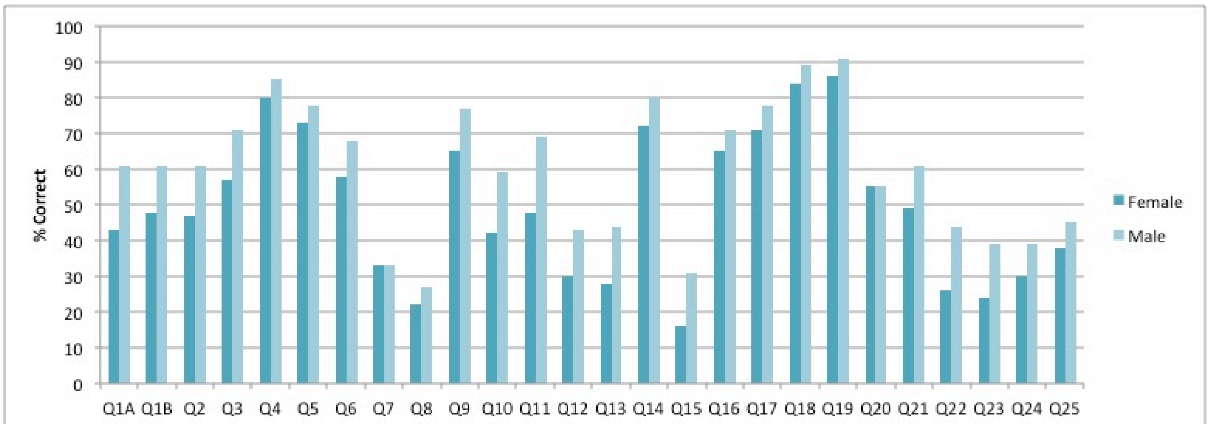


Figure 5.3: A comparison of the percentage of correct responses by gender for the core concept diagnostic items in semester 1, 2013.

There is no explanation for this lower performance evident from consideration of the composition, format, or display of the concept items. The literature on differences in achievement on concept tests by gender is not extensive. McCullough (2004) has analysed performance on the Force Concept Inventory (FCI) and found that men achieve higher outcomes than women; changing the context of the items from cannonballs, rockets and steel balls to babies, slides and jewellery led to different performance on some questions but no overall improvement for women, while men's scores declined. Thus, context is important and can make a significant difference to student performance. Our CCI includes items with cooking (boiling water, burning match) as well as more abstract (balancing equation) contexts, and it is not clear to what extent the contexts affect student performance. Sharma and Bewes (2011) have observed poorer performance by female undergraduate students across all ability levels in a physics quiz based on the FCI, but both genders were equally poor at self-monitoring. Madsen, McKagan, and Sayre (2013) reviewed the literature relating to the persistent gender gap in the FCI and explored its possible origins. They concluded that there are many factors leading to the observed gap and that "the gender gap on concept inventories should not be considered a well-understood or solved problem" (p. 020121-14).

In chemistry, Yezierski and Birk (2006) used the Particulate Nature of Matter Assessment (ParNoMA), from which items for the Foundation Chemistry Instrument were drawn. They found significant differences in performance by male and female students ranging from 8th grade to university, with males performing better. Kendhammer, Holme and Murphy (2013) have recently reported differential item functioning in multiple-choice chemistry items according to gender, with some questions favouring women and other questions favouring men. Each of the skills required to answer questions (reasoning, specific chemical knowledge, computation, visual-spatial) sometimes favoured women and other times, men.

To date, no gender-based analysis of the Chemical Concepts Inventory (Mulford & Robinson, 2002) has been reported. Thus, our data represents the first major analysis of chemical concept question responses according to gender. Further work is required to elucidate whether context or content are primarily responsible for the observed differences in outcomes, and whether interventions can be effective in reducing the difference in performance, as reported by Yezierski and Birk (2006).

5.3 Student motivation and self-regulation

One of the key challenges that arose during this project was the engagement of students with formative feedback and in self-directed learning activities. It quickly became evident that when students perceived there was no direct benefit in terms of summative marks to be gained that they did not access or persist with optional learning activities. The project team decided after the first iteration of diagnostic and intervention strategies that student perceptions of feedback and their motivation to learn in self-directed environments should be explored further. The following evaluation to collect data was completed:

- The Motivated Strategies for Learning Questionnaire (MSLQ) was embedded in the end-of-semester evaluation in semester 2, 2013 and semester 1, 2014 at UQ.
- Student engagement in face-to-face and online activities was evaluated at UoW in semester 1, 2014.
- End-of-semester focus groups were conducted at all participating institutions in June 2014.

5.3.1 Motivated Strategies for Learning Questionnaire

The MSLQ (Duncan & McKeachie, 2005) is a well-characterised literature instrument based on a seven-point likert scale that demonstrates reliable internal consistency for each of its clusters across multiple contexts. This enables researchers to select and combine the specific clusters that support their evaluation questions. The implementation of this questionnaire at UQ included five motivation scales and six learning strategies scales that resulted in Cronbach alpha values equivalent to, or having higher internal consistency than, published data (Table 5.4).

Table 5.3: Internal consistency (Cronbach alpha values) for selected clusters of core concept items from the MSLQ in S2, 2013 and S1, 2014, compared to literature data.

	Scale	**2013 S2 Mean**	**2013 S2 α**	**2014 S1 Mean**	**2014 S1 α**	**α Duncan & McKeachie, 2005**
Motivation Scales	Intrinsic motivation	**5.41**	0.74	**4.89**	0.74	0.74
	Extrinsic motivation	**4.94**	0.76	**5.37**	0.71	0.62
	Self-efficacy for learning and performance	5.08	0.94	4.94	0.95	0.93
	Task value	5.23	0.89	5.20	0.89	0.90
	Control of learning beliefs	5.70	0.83	5.60	0.76	0.68
Learning Strategies Scales	Organization	4.79	0.69	4.91	0.66	0.64
	Metacognitive self-regulation	4.51	0.75	4.49	0.76	0.79
	Time/study environmental management	4.45	0.71	4.66	0.72	0.76
	Effort regulation	4.68	0.65	4.83	0.68	0.69
	Peer learning	**4.18**	0.71	**4.03**	0.70	0.76
	Help seeking	**4.09**	0.43	**4.16**	0.35	0.52

An interesting outcome of this analysis was that students identified more with extrinsic motivation than intrinsic motivation at the end of their first semester. However, at the end of the second semester they identified more strongly with intrinsic motivation than extrinsic motivation. This aligns with other quantitative (LMS and web analytics) evidence and qualitative evidence (open-response questions and focus-group interviews) gathered throughout this project in regard to how summative assessment correlates with self-directed learning.

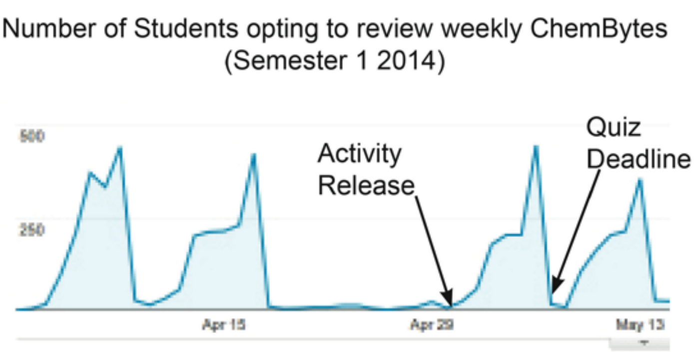


Figure 5.4: Google Analytics mapping student access to the ChemBytes websites during

semester 1, 2014 (the vertical axis indicates the number of views).

In semester 1, 2014, UQ students were provided with formative feedback from the concept cluster items that directed them to the online modules on the ChemBytes website. Google Analytics enabled collection of data in regard to how students accessed the web pages. These activities were structured in three stages: a concept cluster (formative), reviewing online modules (formative) and a short test related to lecture concepts (summative. The two formative elements were voluntary but all students were required to attempt the summative test. Analytics revealed that students began to access the resources immediately after their release on a Friday afternoon but the highest number of page views was in the 24-hour period before the summative test submission deadline. Students were advised to review the online modules prior to commencing the summative test (once accessed, the summative test had a time limit of two hours to complete four questions). When the task deadline had passed, analytics revealed that students stopped accessing the related resources – as the students' engagement with the formative resources was indirectly linked to the summative test, this represented a form of extrinsic motivation.

6. Timing and feedback

The core thread throughout this project has been the provision of formative feedback to increase students' awareness of their mental models in chemistry. Formative feedback is will only be perceived as useful by students and acted on when the timing and format of the feedback is optimal (Carless et al., 2010; Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). The original aim was to provide tailored feedback that provided students with an indication of the specific nature of their alternate conceptions if they existed. The strategy that was adopted by the project team was to assign the alternative answers in the concept items to specific alternate conceptions then combine these across the five items in each cluster. Several approaches were trialled including automated mail-merged pdf documents emailed once students had completed the diagnostic instrument (2013) and embedding the feedback in LMS quizzes so students received feedback as soon as they submitted their formative test. Evaluation of the semester 1, 2013 implementation through focus groups provided several insights into students' perceptions of the feedback that they received; these were categorised in three ways (Table 6.1).

Table 6.1 Student perceptions of feedback they received based on their outcomes in the diagnostic concept instrument (S1, 2013).

Utility of formative feedback	Example
A quantitative indicator of performance which students use to decide whether to act on it.	*'Like if I got seven or above I probably wouldn't bother to get feedback in that situation. But if I got less than that then I would seek it, but always on a quiz when they give it to you and you know where you've gone wrong, I'd always go and check the right answer, and yeah, usually I'd find out why you got it wrong and stuff like that'*
A reminder of the concepts students had studied at high school.	*'I was just like, "Yeah, I should really go back and remember all my stuff from high school before I go into uni"'*
A trigger for seeking an explanation or advice in regard to incorrect answers	*'an explanation of what you may have done wrong would be good, or like an explanation of how you can get it right, if you did get it wrong'*

Students made no reference to self-directed learning using resources supplied by the instructor, or otherwise, in this early evaluation. In the next semester, student perceptions of feedback were explored through an open question in the end-of-semester questionnaire.

6.1 Engagement with formative feedback

The evidence across the three semesters of implementation in this project indicates that the majority of students will act on formative feedback when it is linked directly or indirectly to summative assessment. In 2013, the feedback was delivered at the beginning of semester 1 but it was found that students did not link back to this feedback later in the semester. At a workshop involving academics at the International Conference on the First Year in Higher Education, strategies were sought from participants after presentation of our data (Lawrie et al., 2013). This informed the strategy adopted by several instructors (UQ, UWS and Monash case studies) where the diagnostic instrument was delivered as individual clusters close in timing to the teaching of related concepts. This resulted in greater engagement of students with the feedback and linked learning activities.

6.2 Evidence of change in conceptual understanding

The question that remains partially answered at this stage in this project is whether the students who acted on formative feedback and completed the tailored learning activities subsequently improved their understanding of related concepts. A number of strategies were applied by the project team in their own contexts to attempt to measure gains including post-testing of concept items and embedded exam questions in the end of semester assessment. While individual gains in conceptual understanding were widely evident, these shifts in outcomes were not universal across every concept. This aligns with literature in that shows that many alternate conceptions are very difficult to adjust in a single intervention hence are recognised as intransigent (Taber & Tan, 2011). Student feedback indicated that the activities of their instructors and resources related to this project raised their awareness of their own thinking (metacognition) and this impacted on their approaches to study. The substantial data collected across semester 1, 2014 will continue to be analysed over the next few months and disseminated in relation to this claim. However, it should be noted that a single recommendation for practice, based solely on the formative feedback and interventions alone, is not possible because instructors changed their practice as a result of participating in this project. It is clear that this has resulted in more constructivist environments in these large first-year chemistry classes and is an outcome for the project.

7. Project dissemination

The dissemination activities in this project were informed by the recommendations of the D-cubed framework (Hinton et al., 2011). Engaged dissemination was achieved through the development and extension of networks of academics and conference workshops. Dissemination for information will be achieved through journal publications (ongoing) and distributed instructor resources. A website (ChemBytes) as the platform for the online modules has been created and another, which will be linked, will provide downloadable resources is in preparation. These URLs will be shared on the chemistry discipline teaching and learning network, Chemnet and other online teaching and learning community hubs.

7.1 Networks of academics

This project has received widespread attention from academics in multiple countries through conference presentations by the project team. In particular, the diagnostic tool has attracted significant interest evidenced by:

- Individual academics requesting access to the core concept diagnostic instrument for use in their own contexts.
- An online chemistry teaching forum in the UK adopted the diagnostic instrument to

catalyse a discussion in regard to how concept inventories may be applied. The core concept instrument was then implemented at The University of Southampton.

7.2 Conference presentations and workshops

Table 7.1 Oral & poster presentations

	Date	Forum
International Meetings	October 2014	Conference for the International Society for Scholarship of Teaching and Learning (Quebec, Canada).
	October 2014	Transforming Institutions: 21st Century Undergraduate STEM Education Conference (Indianapolis, US)
	July 2014 (oral)	International Conference on Chemical Education (Toronto, Canada)
	July 2013 (oral)	International Conference on Chemical Education, (University of Rome, Rome, Italy)
	June 2013 (poster)	Gordon Research Conference on Chemical Education Research & Practice (Salve Regina University, Rhode Island, US)
	July 2013 (oral)	5th Eurovariety in Chemistry Education Conference (Limerick, Ireland)
National Meetings	September 2014 (oral)	Australian Conference on Science and Mathematics Education (University of Sydney, Sydney)

Table 7.2 Interactive conference workshops

	Date	Forum	Number of Participants
International Meetings	April 2014	HEA STEM Annual Learning and Teaching Conference 2014: Enhancing the STEM Student Journey. (Edinburgh, UK)	9
	July 2013	International Conference on the First Year in Higher Education (Wellington, NZ)	26
National Meetings	September 2013	Australian Conference on Science and Mathematics Education (Canberra)	30

Abstracts have also been accepted for oral presentations at the following meetings:

- ‘Do you see what I see? Translating between representations in chemistry’. RACI Congress. December 7–12, 2014, Adelaide.

7.3 Publications

A publication framework was established early in this project to distribute the leadership in realising publications across the project team. Five journal articles have been formulated and the process of collaborative authorship is to be led by the designated project team member (Table 7.3).

Table 7.3 Project publication strategy

Project team coordinator(s)	Publication Aim (Target Journal)	Data
Gwen Lawrie (all project team members are weighted equally as authors)	Dissemination of the validated concept diagnostic instruments.	Statistical and psychometric data.
Madeleine Schultz	The design of the online intervention modules.	Iterative action cycle of formulation and response to student activity and feedback.
Glennys O'Brien & Simon Bedford	Factors affecting student engagement in activities to address alternate conceptions.	Comparison between face-to-face and online activities.
Roy Tasker and Mark Williams	Graphical versus textual concept items.	Concept-inventory outcomes and student interviews
Gwen Lawrie	Academic engagement in the process of designing items for diagnostics.	Ethnographic processes over the two years of the project focussing on single items
Anthony Wright (Md Abdullah Al Mamun, PhD research)	Student engagement in online inquiry-based simulations.	Stimulated response interviews

Publication of outcomes from the project to date:

Lawrie, G., Wright, A, Schultz, M., Dargaville, T., O.Brien, G., Bedford, S., Williams, M., Tasker, R., Dickson, H., & Thompson, C. (2013). Using formative feedback to identify and support first-year chemistry students with missing or misconceptions. *International Journal of the First Year in Higher Education*, *4*(2), 111–6.

Lawrie, G., Wright, A, Schultz, M., Dargaville, T., O.Brien, G., Bedford, S., Williams, M., Tasker, R., & Thompson, C. (2015). Closing the loop: A model for inter-institutional collaboration through delivering formative assessment in large, first-year STEM classes. *Invited book chapter in 'Transforming Institutions: 21st Century undergraduate STEM education' (Ed. Weaver, G), Purdue University Press (Indiana), submitted Jan 21 2015.*

7.4 Links between this project and other OLT projects

The activities within this project have informed four current OLT funded initiatives:

- Professor Manjula Sharma (2013 OLT National Teaching Fellowship) has incorporated the Chemical Concept Inventory items in her project activities and engaged in discussions with project team members.
- The 2014 OLT Extension Grant (Leader: Michael Jennings) investigating translation of diagnostic testing from engineering to science disciplines will be informed by our concept instruments.
- Professor Roy Tasker (2014 OLT National Senior Teaching Fellowship), one of our project team members, is extending and disseminating his VisChem animation resources.
- The 2014 OLT seed project (Schultz & Lawrie) that aims to develop resources to support academics in the development of their pedagogical content knowledge has been catalysed through the activities in this project.

7.5 Project Impacts

- An innovative new approach to the application of diagnostic concept testing has been demonstrated and linked to provision of formative feedback and online activities that address students' alternate conceptions in a range of constructive learning environments.
- The activities in this project have irrevocably altered the practices of 10 academics involved in leading instruction in first-year chemistry programs across five universities and three states in Australia. This shift in practice has impacted on the student learning processes for thousands of students since 2012 and will continue through the sustained dissemination and transfer of practices.
- This project has enabled two undergraduate research projects (Hassan and Henders) and one PhD project (Al Mamun), the aim of the latter is to investigate how students construct and adjust their conceptual understanding and their motivation to engage with online activities.
- This project has also catalysed widespread interest and discussions regarding instructor awareness of student thinking, strategies for addressing alternate conceptions and the role of concept inventories: research versus practitioner tool. One of the direct impacts of the project is a successfully funded OLT seed project (Schultz and Lawrie, 2014) that aims to support early-career academics in their development of pedagogical content knowledge.

7.6 Project Team Reflections

This project has progressed successfully due to the commitment of the project team members who are all instructors invested in improving student learning. Through the collaborative processes and experiences in the project, several important factors were highlighted which the team feels are important for any academics who may wish to build on this project's outcomes. These include:

- Projects that involve activities implemented in courses (units) with students across multiple institutions require a minimum of three iterations to collect sufficient evidence to support conclusions and recommendations. This could take longer than originally anticipated, particularly when ethical approval from multiple institutions is involved.
- Academic-based teams should secure institutional support in the form of personnel

who have expertise in LMS to enable delivery of technology-enhanced learning initiatives across multiple institutions. There are unavoidable variations in the offerings and versions of LMS that require resources to resolve.

- In projects that involve evaluation of student learning outcomes from innovations, across multiple institutions, there needs to be at least a six-month period included after completion of data collection to allow comprehensive analysis to be completed prior to submission of the final project report. Project team members should expect to continue in both data analysis and dissemination of outcomes beyond the term of the project to maximise the impact of the project.

References

Arjoon, J. A., Xu, X., & Lewis, J. E. (2013). Understanding the state of the art for measurement in chemistry education research: Examining the psychometric evidence. *Journal of Chemical Education, 90*, 536–545.

Barbera, J., Adams, W. K., Wieman, C. E., & Perkins, K. K. (2008). Modifying and validating the colorado learning attitudes about science survey for use in chemistry. *Journal of Chemical Education, 85*, 1435–1439.

Barbera, J. (2013). A psychometric analysis of the chemical concepts inventory. *Journal of Chemical Education, 90,* 546–553.

Bell, P. & Volckmann, D. (2011). Knowledge Surveys in General Chemistry: Confidence, Overconfidence, and Performance. *Journal of Chemical Education,* 88, 1469−76.

Bodner, G. M. (1991). I have found you an argument: The conceptual knowledge of beginning chemistry graduate students. *Journal of Chemical Education, 68*, 385–388.

Carless, D., Salter, D., Yang, M., & Lam, J. (2010) Developing sustainable feedback practices. *Studies in Higher Education*, 36, 395–407.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts, *Learning and Instruction, 4*, 27–43.

Duncan, T.G., & McKeachie, W.J., (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, 40, 117−128.

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching*, *29*(6) 611–28.

Hadenfeldt, J. C., Bernholt, S., Liu, X., Neumann, K., & Parchmann, I., (2014). Using ordered multiple choice items to assess students' understanding of the structure and composition of matter. *Journal of Chemical Education, 90*(12) 1602−8.

Hattie, J., & Timperley, H. (2007) The power of feedback. *Review of Educational Research*, 77, 81−112.

Heredia, K., Xu, X. & Lewis, J. E. (2012). The application and evaluation of a two-concept diagnostic instrument with students entering college general chemistry. *Chemistry Education Research and Practice*, 13, 30−38.

Hinton, T., Gannaway, D, Berry, B., & Moore, K. (2011). The D-cubed guide: Planning for effective dissemination. Australian Learning and Teaching Council, Strawberry Hills, NSW. ISSN: 9781921856655

Johnson, P. (1998). Children's understanding of changes of state involving the gas state, part 1: Boiling water and the particle theory. *International Journal of Science Education*, 20, 567−583.

Kendhammer, L., Holme, T., & Murphy, K. (2013). Identifying differential performance in general chemistry: Differential item functioning analysis of ACS General Chemistry trial tests. *Journal of Chemical Education*, *90*, 846–853.

Kift, S. (2009) Articulating a transition pedagogy to scaffold and to enhance the first year student learning experience in Australian higher education. ALTC Fellowship Report. Accessed March 11, 2011 from: http://www.altc.edu.au/resource-first-year-learning-experience-kift-2009.

Kift, S., Nelson, K., & Clarke, J. (2010) Transition pedagogy: A third generation approach to FYE – A case study of policy and practice for the higher education sector. *International Journal of the First Year in Higher Education*, 1(1) 1–20.

Krause, S., Birk, J., Bauer, R., Jenkins, B., & Pavelich, M. J. (2004). Development, testing and application of a chemistry concept inventory. Paper presented at the 34th ASEE/IEEE Frontiers in Education Conference, Savannah, 20–23 October, 2004.

Krause, S., Kelly, J., Corkins, J., & Tasooji, A. (2009). The role of prior knowledge on the origin and repair of misconceptions in an introductory class on materials science and engineering. *Research in Engineering Education Symposium*. Palm Cove, Qld, Australia.

Lawrie, G., Wright, A, Schultz, M., Dargaville, T., O.Brien, G., Bedford, S., Williams, M., Tasker, R., Dickson, H., & Thompson, C. (2013). Using formative feedback to identify and support first-year chemistry students with missing or misconceptions.

International Journal of the First Year in Higher Education, 4(2) 111–6.
Linenberger, K. J., & Lowery Bretz, S. (2012). Generating cognitive dissonance in student interviews through multiple representations. *Chemistry education research and practice, 13*, 172–8.
Madsen, A., McKagan, S. B., & Sayre, E. C. (2013). Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? *Physical Review Special Topics - Physics Education Research, 9*, 020121.
Marbach-Ad, G., Mcadams, K. C., Benson, S., Briken, V., Cathcart, L., Chase, M., El-Sayed, N. M., Frauwirth, K., Fredericksen, B., Joseph, S. W., Lee, V., Mciver, K. S., Mosser, D., Quimby, B. B., Shields, P., Song, W., Stein, D. C., Stewart, R., Thompson, K. V., & Smith, A. C. (2010). A model for using a concept inventory as a tool for students' assessment and faculty professional development. *CBE - Life Sciences Education*, *9*, 408–416.
McCullough, L. (2004). Gender, context, and physics assessment. *Journal of International Women's Studies, 5*, 20−30.
Molecular Workbench website: http://mw.concord.org/modeler/
Mulford, D. R. & Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education, 79*, 739–744.
Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31, 199–218.
Osborne, R. J. & Cosgrove, M. M. (1983). Children's conceptions of the changes of state of water. *Journal of Research in Science Teaching, 20*, 825−838.
Othman, J., Treagust, D. F. & Chandrasegaran, A. L. (2008). An investigation into the relationship between students' conceptions of the particulate nature of matter and their understanding of chemical bonding. *International Journal of Science Education, 30*, 1531–1550.
Pavelich, M. J., Jenkins, B., Birk, J., Bauer, R., & Krause, S. (2004). Development of a chemistry concept inventory for use in chemistry, materials and other engineering courses, paper presented at the American Society for Engineering Education Annual Conference and Exposition, City.
Pentecost, T.C., & Barbera, J., (2013). Measuring learning gains in chemical education: a comparison of two methods. *Journal of Chemical Education*, *90*, 839–45.
PhET website: http://phet.colorado.edu/
Phillips, R., McNaught, C., & Kennedy, G. (2012). *Evaluating e-learning: Guiding research and Practice*. New York: Routledge.
Pitkethly, A. & Prosser, M. (2001) The first year experience project: A model for university-wide change. *Higher Education Research & Development*. 20(2), 185–198.
Potgieter, M., Davidowitz, B., and Venter, E. (2008). Assessment of preparedness of first-year chemistry students: development and application of an instrument for diagnostic and placement purposes. *African Journal of Research in Mathematics, Science and Technology Education, Special Edition*, 12, 1−18.
Potgieter, M., Ackerman, M., & Fletcher, L. (2010). Inaccuracy of self-evaluation as additional variable for prediction of students at risk of failing first-year chemistry. *Chemistry Education Research and Practice*, 11, 17−24.
Potgieter, M., & Davidowitz, B. (2011). Preparedness for tertiary chemistry: Multiple applications of the chemistry competence test for diagnostic and prediction purposes. *Chemistry Education Research and Practice, 12*, 193–204.
Prince, M., Vigeant, M., & Nottis, K. (2012). Development of the heat and energy concept inventory: Preliminary results on the prevalence and persistence of engineering students' misconceptions. *Journal of Engineering Education*, *101*(3) 412–38.
Sandi-Urena, S., Cooper, M., & Stevens, R. (2010). Enhancement of metacognition use and awareness by means of a collaborative intervention. *International Journal of Science Education,* 33, 323−40.
Schraw, G, Crippen, K., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader persepective on learning. *Research in Science Education*, 36, 111−139.
Schwartz, P., & Barbera, J. (2014). Evaluating the content and response process validity of data from the chemical concepts inventory. *Journal of Chemical Education, 91*, 630–

640.
Sharma, M. D., & Bewes, J. (2011). Self-monitoring: Confidence, academic achievement and gender differences in physics. *Journal of Learning Design, 4*(3), 1–13.
Stains, M. (2007). Classification of chemical substances, reactions and interactions: the effects of expertise. PhD Thesis, University of Arizona. Accessed October 27, 2014, from http://arizona.openrepository.com/arizona/bitstream/10150/194835/1/azu_etd_2054_sip1_m.pdf
Taber, K.S., & Tan, K.C.D. (2011). The insidious nature of 'hard-core' alternate conceptions: Implications for the constructivist research programme of patterns in high school students' and pre-service teachers' thinking about ionisation energy. *International Journal of Science Education*, 33, 259−97.
Treagust, D. (1986). Evaluating students' misconceptions by means of diagnostic multiple choice items. *Research in Science Education*, 16, 199−207.
Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education, 10* (2), 159−169.
Treagust, D., Chandrasegaran, A., Zain, A., Ong, E., Karpudewan, M., & Halim, L. (2011). Evaluation of an intervention instructional program to facilitate understanding of basic particle concepts among students enrolled in several levels of study. *Chemistry Education Research and Practice*, 12, 251−261.
Vischem website: http://www.vischem.com.au/
Wei, S., Liu, S., Wang, Z., & Wang, X. (2012). Using Rasch measurement to develop a computer modeling-based instrument to assess students' conceptual understanding of matter. *Journal of Chemical Education, 89*, 335–45.
Yezierski, E. J., & Birk, J. P. (2006). Misconceptions about the particulate nature of matter: using animations to close the gender gap. *Journal of Chemical Education, 83*, 954–960.

Appendix A

2013 Core Concept Diagnostic Instrument

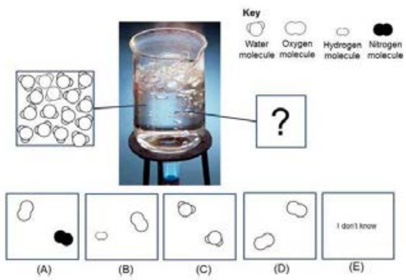
In 2013, a 25-item instrument was delivered across four institutions comprised of five clusters of five items exploring student conceptions in phase change, heat and energy, conservation of matter, aqueous solutions and equilibrium. A series of demographic questions were included:

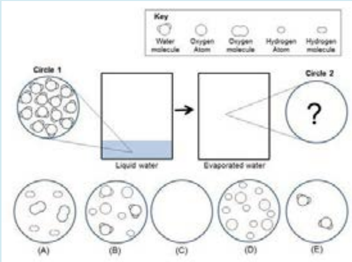
- Which program are you currently enrolled in at [institution]?
- Please indicate your age (not compulsory):
  Options: Under 17/17/18/19/20/21/22/23/Over 23
- Please indicate your gender (not compulsory):
  Options: Male/Female
- Please indicate your OP score or tertiary entrance rank:
  Options: OP/ATAR/Other
- When did you finish your most recent studies
  Options: In the last 6 months/Between 6-12 months ago/More than one year ago but less than two years ago/More than two years ago but less than 5 years ago/More than 5 years ago/I have not studied chemistry before.
- Where did you finish your most recent studies:
  Options: Queensland secondary school/New South Wales secondary school/Victoria secondary school/Other
- Is English your first language?
  Options: Yes/No. If No, what is your first language?

Phase Change Cluster

We often see several states of matter coexisting. Consider what is involved as substances changes between the solid, liquid and gaseous states.

Note: students in semester 1 2013 were randomly allocated 1A or 1B in their survey.

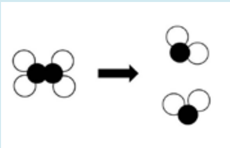
Item number	Stem	Options
1A	Assume a beaker of pure water has been boiling for 30 minutes. What is in the bubbles in the boiling water?	(A) Air. (B) Oxygen gas and hydrogen gas. (C) Oxygen gas. (D) Water vapour. (E) I don't know.
1B	Assume a beaker of pure water has been boiling for 30 minutes. What is in the bubbles in the boiling water?	

2	Circle 1 shows a magnified view of a very small portion of liquid water in a closed container.  	Which of the circles labelled (A) to (E) above would show the magnified view after the water evaporates, in Circle 2 on the right. (A) (B) (C) (D) (E)
3	When water changes from a liquid to a gas, heat is transferred to:	(A) form new bonds between atoms to form new molecules. (B) break an O-H bond within a water molecule. (C) break the hydrogen bonds between water molecules. (D) make the molecules larger so they take up more space. (E) I don't know.
4	A small dry glass is filled with ice and left to stand on a kitchen bench. Fifteen minutes later some of the ice has melted and the outside of the glass is covered in water droplets. Where has the water on the outside of the glass come from?	(A) Water from the melted ice evaporates and condenses on the outside. (B) The water from the melted ice passes through the glass. (C) Water vapour from the air condenses on the cold surface of the glass. (D) The coldness causes oxygen and hydrogen from the air to combine forming the water on the glass. (E) I don't know.
5	At room temperature and pressure, H_2O is a liquid and N_2 is a gas because:	(A) an N_2 molecule is heavier than an H_2O molecule. (B) the bonds between H_2O molecules are stronger than the bonds between N_2 molecules. (C) an N_2 molecule is lighter than an H_2O molecule. (D) an H_2O molecule contains more atoms than an N_2 molecule. (E) I don't know.

A confidence scale was included after each of the questions in this cluster only, where students were asked: ‘How confident are you about your response to the question above?’. Answer options were: Totally guessed answer/Almost a guess/Almost certain/Certain.

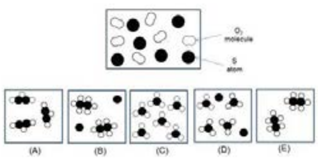
Heat & Energy Cluster

Consider what is involved as energy is transferred during chemical and physical processes.

Item number	Stem	Options
6	Heat can be best described as:	(A) energy flow between bodies at different temperatures. (B) friction from particles rubbing together. (C) a reading on a thermometer. (D) the absence of cold. (E) a substance that makes objects feel warm.
7	Heat is given off when hydrogen gas burns in air according to the equation: $2H_2 + O_2 \rightarrow 2H_2O$ Which of the following is responsible for the heat?	(A) Breaking hydrogen bonds gives off energy. (B) Breaking oxygen bonds gives off energy. (C) Forming oxygen-hydrogen bonds gives off energy. (D) Both (a) and (b) are responsible. (E) (a), (b) and (c) are responsible. (F) I don't know.
8	A molecule of dinitrogen tetroxide, N_2O_4, contains an N-N bond. The bond can break to form two molecules of nitrogen dioxide, NO_2.  Which of the following statements is true?	(A) When the bond breaks energy will be absorbed. (B) When the bond breaks energy will be released. (C) When the bond breaks no energy will be absorbed or released. (D) It is not possible to predict whether energy would be absorbed or released when the bond breaks. (E) I don't know.
9	A piece of iron railway track will get longer when it is in the hot sun which is why there are usually gaps between pieces of track. When the track expands the iron atoms:	(A) form bigger molecules so that they take up more space. (B) vibrate more so the atoms take up more space. (C) increase in size so they take up more space. (D) contain electrons that move faster so the atom takes up more space. (E) I don't know.
10	In a classroom containing metal and plastic chairs, people often say that the metal chairs feel colder than the plastic chairs. This is because:	(A) metal naturally attracts and holds cold. (B) plastic is an insulator so it absorbs and retains heat. (C) the metal chairs are at a lower temperature than the plastic chairs. (D) metal quickly conducts heat away from your skin. (E) I don't know.

Conservation of Matter Cluster

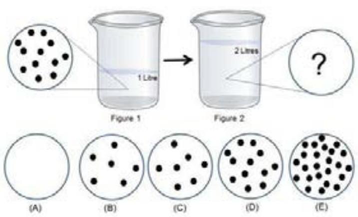
Chemical reactions occur all around us. Accounting for atoms is very important.

Item number	Stem	Options
11	Iron reacts with oxygen and water from the air to form rust. If an iron nail was allowed to rust completely, the rust would weigh:	(A) less than the nail it came from. (B) the same as the nail it came from. (C) more than the nail it came from. (D) it is impossible to predict. (E) I don't know.
12	The large box below represents a mixture of S atoms and O_2 molecules in a closed container:  Which of the smaller diagrams labelled (A) to (E) above represents the contents of the container after the original mixture reacts as completely as possible according to the equation: $2S + 3O_2 \rightarrow 2SO_3$	(A) (B) (C) (D) (E)
13	Which of the following must be the same before and after a chemical reaction?	(A) The sum of the masses of all the substances involved. (B) The number of molecules of all substances involved. (C) The number of atoms of each type involved. (D) The sum of the masses and the number of atoms of each type. (E) The sum of the masses, the number of molecules of all substance and the number of atoms of each type.
14	Nitroglycerine is used as an explosive. The equation for the chemical reaction is: $4C_3H_5(NO_3)_3(l) \rightarrow 12CO_2(g) + 10H_2O(g) + 6N_2(g) + O_2(g)$ When 24 molecules of	(A) 32 molecules of carbon dioxide and 20 molecules of oxygen. (B) 72 molecules of carbon dioxide and 20 molecules of oxygen. (C) 72 molecules of carbon dioxide and 24 molecules of oxygen. (D) 32 molecules of carbon dioxide and 6 molecules of oxygen.

	nitroglycerine react, how many molecules of carbon dioxide and how many molecules of oxygen are produced?	(E) 72 molecules of carbon dioxide and 6 molecules of oxygen.
15	When a match burns, what is the fate of most of the mass of the match?	(A) It converted into heat. (B) The mass of the ash is the same as the mass of the match. (C) It is converted into gaseous carbon dioxide, water and heat. (D) It is converted into ash and heat. (E) It is converted into gaseous carbon dioxide and water.

Aqueous Solutions Cluster

Consider substances dissolved in water and what chemical species are present.

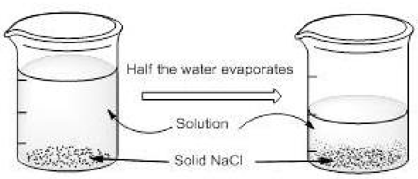
Item number	Stem	Options
16	When a teaspoon of calcium chloride ($CaCl_2$) dissolves in a glass of water, a clear liquid is formed. What happens to the calcium chloride?	(A) The calcium chloride and water form a new substance. (B) The calcium chloride separates into calcium atoms and chlorine molecules. (C) The calcium chloride molecules become separated and surrounded by water. (D) The calcium chloride forms calcium ions and chloride ions which are surrounded by water molecules. (E) I don't know.
17	The picture below represents a 1.0 L solution of sugar dissolved in water. The dots in the magnification circle represent the sugar molecules. In order to simplify the diagram, the water molecules have not been shown.  Which of the circles labelled (A) to (E) above represents the view after 1.0 L of water is added to the beaker in Figure 1?	(A) (B) (C) (D) (E)

18	What happens to the concentration of salt in a dilute solution if half the volume of water evaporates? The concentration will:	(A) increase. (B) decrease. (C) stay the same. (D) need to be measured because it is not possible to say. (E) I don't know.
19	What is the reason for your answer to the question above (Q18)?	(A) There is the same amount of salt in less water. (B) Solid salt will form. (C) The salt and water both evaporate. (D) Ions are constantly leaving and returning to the solid. (E) I don't know.
20	Alum, $KAl(SO_4)_2$, is often used in water treatment plants. What ions would you find in a dilute alum solution?	(A) K^+ and $Al(SO_4)^{2-}$ (B) KAl^{4+} and SO_4^{2-} (C) K^+ and Al^{3+} and SO_4^{2-} (D) Al^{3+} and KSO_4^{2-} and SO_4^{2-} (E) I don't know.

Chemical Equilibrium Cluster

Equilibrium is a complex concept. Consider what happens in the systems below.

Item number	Stem	Options
21	Which one of the following best describes a dynamic equilibrium?	(A) Reactions continue with no effect on the concentrations of reactants and products. (B) Concentrations of products are high. (C) Reactions have stopped. (D) There are equal concentrations of reactants and products (E) I don't know.
Q22&23	Consider the following diagram and answer the two questions below:	



In a beaker, sodium chloride salt (NaCl) is added to water and the mixture is stirred until no more salt dissolves. The salt that does not dissolve is allowed to settle out (above left beaker).

22	What happens to the concentration of salt in solution if water evaporates until the volume of the solution is half the	The concentration: (A) increases. (B) decreases.

	original volume? (above right beaker). You can assume that the temperature remains constant throughout the experiment.	(C) constantly changes. (D) stays the same. (E) I don't know.
23	What happens to the concentration of salt in a dilute solution if half the volume of water evaporates? The concentration will:	(A) There is the same amount of salt in less water. (B) More solid salt forms. (C) Salt does not evaporate and is left in solution. (D) There is less water. (E) I don't know.
24	The two linked equilibrium reactions shown below are important in your blood, in the ocean and also in fizzy drinks. $2H_2O(l) + CO_2(g) \rightleftarrows H_2CO_3(aq) + H_2O(l) \rightleftarrows H_3O^+(aq) + HCO_3^-$(aq  If you open a bottle of fizzy water (without spilling any), bubbles rapidly form. The amount of water:	(A) increases, because the gas pressure in the bottle decreases and H_2CO_3 reacts to produce more CO_2 (B) decreases, because water evaporates when the CO_2 gas leaves the bottle. (C) constantly changes, because the new equilibrium is constantly changing. (D) stays the same, because water is present on both sides of the left hand equilibrium equation. (E) I don't know.
25	Formic acid is the simplest carboxylic acid, HCOOH. It is the cause of pain in bee and ant stings and nearly a million tonnes are made each year because it is a good preservative. In water, formic acid behaves as a weak acid: $HCOOH(aq) + H_2O(l) \rightleftarrows HCOO^-(aq) + H_3O^+(aq\ )$ Which of the statements below is the best explanation of the process of returning to equilibrium after some sodium formate, HCOONa, has been added to a formic acid solution?	(A) The HCOOH concentration will decrease, because it reacts with the added HCOONa. (B) The H_3O^+ concentration will increase, because the ratio of H_3O^+ to $HCOO^-$ remains constant at equilibrium. (C) The $HCOO^-$ concentration will increase and concentration of the other species will decrease to balance the reaction. (D) The H_3O^+ concentration will decrease, because some will react with $HCOO^-$ to return to the equilibrium product/reactant ratio. (E) I don't know.

Appendix B

Impact of completing senior high-school chemistry

The difference in achievement for students was compared based on whether they had completed senior high-school chemistry or not (semester 1, 2013). This data is for the combined set across four institutions and three states.

* indicates an item where there was no significant difference in achievement between students who had completed high-school chemistry and those who had not.

Table B1: Statistics for the pilot diagnostic concept items comparing students' outcomes with and without high-school chemistry. For items 2 to 25, N=1100 for students who had completed high-school chemistry (HS Chem) and N=155 for students had not studied chemistry previously (No HS Chem).

Item	Topic	Prior chemistry	Mean	Standard Deviation	t statistic	Significance
1A	Water boiling, bubbles (text)	HS Chem (N=556)	0.51	0.500	7.179	<0.001
		No HS Chem (N=89)	0.18	0.386		
1B	Water boiling, bubbles (graphic)	HS Chem (N=545)	0.53	0.500	3.050	0.003
		No HS Chem (N=65)	0.34	0.477		
2	Water evaporating	HS Chem	0.52	0.500	7.280	<0.001
		No HS Chem	0.25	0.433		
3	Heat transferred in changing water to a gas	HS Chem	0.62	0.484	4.057	<0.001
		No HS Chem	0.45	0.499		
4	Water condensing on a glass	HS Chem	0.82	0.388	4.371	<0.001
		No HS Chem	0.64	0.482		
5	Nitrogen gas and water intermolecular forces	HS Chem	0.75	0.433	8.977	<0.001
		No HS Chem	0.38	0.487		
6	Heat can be described as	HS Chem	0.62	0.486	4.616	<0.001
		No HS Chem	0.43	0.49		
7	Hydrogen gas burning in air	HS Chem	0.33	0.471	0.885	0.376*
		No HS Chem	0.30	0.458		
8	Dinitrogen tetroxide	HS Chem	0.25	0.431	9.488	<0.001
		No HS	0.05	0.208		

	bond breaking	Chem				
9	Iron railway expanding	HS Chem	0.70	0.459	6.316	<0.001
		No HS Chem	0.43	0.497		
10	Plastic and metal temperature	HS Chem	0.49	0.500	4.027	<0.001
		No HS Chem	0.33	0.471		
11	Rusting nail mass change	HS Chem	0.57	0.496	5.985	<0.001
		No HS Chem	0.32	0.469		
12	Sulfur dioxide stoichiometry	HS Chem	0.35	0.477	7.566	<0.001
		No HS Chem	0.12	0.329		
13	Conservation of matter	HS Chem	0.34	0.475	2.412	0.017
		No HS Chem	0.25	0.435		
14	Nitroglycerine stoichiometry	HS Chem	0.75	0.432	7.805	<0.001
		No HS Chem	0.42	0.495		
15	Burning match fate of matter	HS Chem	0.22	0.414	4.142	<0.001
		No HS Chem	0.10	0.307		
16	Calcium chloride dissolving	HS Chem	0.67	0.471	7.272	<0.001
		No HS Chem	0.37	0.485		
17	Sugar solution dilution	HS Chem	0.74	0.440	2.187	0.030
		No HS Chem	0.65	0.479		
18	Dilute salt solution evaporation	HS Chem	0.86	0.349	3.452	0.001
		No HS Chem	0.73	0.446		
19	Reason for Q19	HS Chem	0.88	0.329	3.540	0.001
		No HS Chem	0.75	0.435		
20	Ion species in solution	HS Chem	0.54	0.498	4.704	<0.001
		No HS Chem	0.35	0.478		
21	Dynamic equilibrium description	HS Chem	0.54	0.499	8.364	<0.001
		No HS Chem	0.23	0.420		
22	Saturated salt solution evaporation	HS Chem	0.33	0.471	1.689	0.093*
		No HS Chem	0.27	0.443		
23	Reason for Q22	HS Chem	0.30	0.460	5.934	<0.001
		No HS Chem	0.12	0.331		
24	Fizzy water equilibrium	HS Chem	0.34	0.475	7.630	<0.001
		No HS Chem	0.12	0.322		
25	Formic acid Le Chatelier	HS Chem	0.41	0.492	10.166	<0.001
		No HS Chem	0.11	0.314		

Impact of gender on outcomes in concept diagnostic

Students were asked to indicate their gender; this question was not required.

Table B2: Analysis of gender as a variable in student outcomes (S1, 2013). N=628 females and N=439 males)

Item	Topic	Gender	Mean (Std Dev)	t statistic (Significance)	□^{2} (Significance)
1A	Water boiling, bubbles (text)	Female (N=311)	0.43 (0.496)	-4.276 (<0.001)	17.660 (<0.001)
		Male (N=228)	0.61 (0.488)		
1B	Water boiling, bubbles (graphic)	Female (N=317)	0.48 (0.500)	-2.897 (0.004)	8.221 (0.004)
		Male Chem (N=211)	0.61 (0.490)		
2	Water evaporating	Female	0.47 (0.499)	-4.576 (<0.001)	20.417 (<0.001)
		Male	0.61 (0.489)		
3	Heat transferred in changing water to a gas	Female	0.57 (0.496)	-4.943 (<0.001)	23.198 (<0.001)
		Male	0.71 (0.453)		
4	Water condensing on a glass	Female	0.80 (0.403)	-2.367 (0.018)	5.345 (0.021)
		Male	0.85 (0.356)		
5	Nitrogen gas and water intermolecular forces	Female	0.73 (0.443)	-1.822 (0.069)*	3.236 (0.072)*
		Male	0.78 (0.414)		
6	Heat can be described as	Female	0.58 (0.494)	-3.433 (0.001)	11.441 (0.001)
		Male	0.68 (0.466)		
7	Hydrogen gas burning in air	Female	0.33 (0.472)	0.008 (0.994)*	<0.001 (0.994)*
		Male	0.33 (0.472)		
8	Dinitrogen tetroxide bond breaking	Female	0.22 (0.415)	-1.846 (0.065)*	3.486 (0.062)*
		Male	0.27 (0.445)		
9	Iron railway expanding	Female	0.65 (0.477)	-4.446 (<0.001)	18.592 (<0.001)
		Male	0.77 (0.418)		
10	Plastic and metal temperature	Female	0.42 (0.495)	-5.335 (<0.001)	27.771 (<0.001)
		Male	0.59 (0.493)		
11	Rusting nail	Female	0.48	-7.243	48.801

	mass change		(0.500)	(<0.001)	(<0.001)
		Male	0.69 (0.462)		
12	Sulfur dioxide stoichiometry	Female	0.30 (0.460)	-4.186 (<0.001)	17.722 (<0.001)
		Male	0.43 (0.495)		
13	Conservation of matter	Female	0.28 (0.448)	-5.306 (<0.001)	28.456 (<0.001)
		Male	0.44 (0.496)		
14	Nitroglycerine stoichiometry	Female	0.72 (0.451)	-3.313 (<0.001)	10.425 (0.001)
		Male	0.80 (0.398)		
15	Burning match fate of matter	Female	0.16 (0.364)	-5.868 (<0.001)	36.222 (<0.001)
		Male	0.31 (0.464)		
16	Calcium chloride dissolving	Female	0.65 (0.477)	-2.066 (0.039)	4.184 (0.041)
		Male	0.71 (0.455)		
17	Sugar solution dilution	Female	0.71 (0.453)	-2.602 (0.09)*	6.529 (0.011)*
		Male	0.78 (0.413)		
18	Dilute salt solution evaporation	Female	0.84 (0.368)	-2.240 (0.025)	4.757 (0.029)
		Male	0.89 (0.318)		
19	Reason for Q19	Female	0.86 (0.345)	-2.244 (0.025)	4.738 (0.030)
		Male	0.91 (0.292)		
20	Ion species in solution	Female	0.55 (0.498)	-0.002 (0.998)*	<0.001 (0.998)*
		Male	0.55 (0.498)		
21	Dynamic equilibrium description	Female	0.49 (0.500)	-3.902 (<0.001)	14.911 (<0.001)
		Male	0.61 (0.488)		
22	Saturated salt solution evaporation	Female	0.26 (0.439)	-6.061 (<0.001)	37.073 (<0.001)
		Male	0.44 (0.497)		
23	Reason for Q22	Female	0.24 (0.430)	-5.061 (<0.001)	26.148 (<0.001)
		Male	0.39 (0.488)		
24	Fizzy water equilibrium	Female	0.30 (0.459)	-3.105 (<0.001)	9.782 (0.002)
		Male	0.39 (0.489)		
25	Formic acid Le	Female	0.38	-2.024	4.121

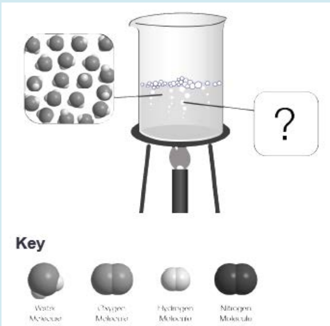
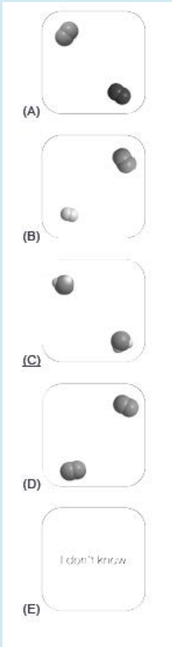
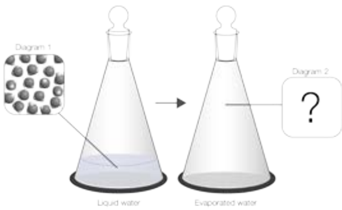
	Chatelier		(0.487)	(0.043)	(0.042)
		Male	0.45 (0.498)		

Appendix C

2014 Core Concept Diagnostic Instrument

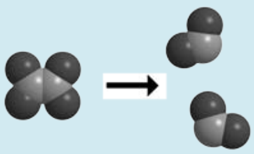
In 2014, the revised 25-item instrument was delivered either as the complete instrument or as individual concept clusters at four institutions.

Phase Change Cluster

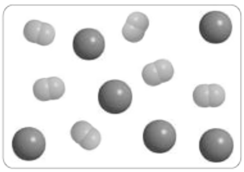
Item number	Stem	Options
1	Assume a beaker of pure water has been boiling for 30 minutes. What is in the bubbles in the boiling water? 	
2	Diagram 1 shows a magnified view of a very small portion of liquid water in a closed container.  Which option represents the magnified view after the water evaporates, in diagram 2 on the right? Key 	

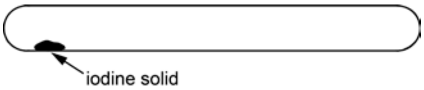
3	When water changes from a liquid to a gas, heat is transferred to:	(A) form new bonds between atoms to form new molecules. (B) break an O-H bond within a water molecule. (C) break the hydrogen bonds between water molecules. (D) make the molecules larger so they take up more space. (E) I don't know.
4	A small dry glass is filled with ice and left to stand on a kitchen bench. Fifteen minutes later some of the ice has melted and the outside of the glass is covered in water droplets. Where has the water on the outside of the glass come from?	(A) Water from the melted ice evaporates and condenses on the outside. (B) The water from the melted ice passes through the glass. (C) Water vapour from the air condenses on the cold surface of the glass. (D) The coldness causes oxygen and hydrogen from the air to combine forming the water on the glass. (E) I don't know.
5	At room temperature and pressure, water is a liquid and nitrogen is a gas because:	(A) a nitrogen molecule is heavier than a water molecule. (B) the bonds between water molecules are stronger than the bonds between nitrogen molecules. (C) a nitrogen molecule is lighter than a water molecule. (D) the bonds within water molecules are stronger that the bonds within nitrogen molecules. (E) I don't know.

Heat & Energy Cluster

Item number	Stem	Options
6	Heat can be best described as:	(A) energy transfer between bodies at different temperatures. (B) friction from particles rubbing together. (C) a reading on a thermometer. (D) the absence of cold. (E) a substance that makes objects feel warm. (F) I don't know.
7	Heat is given off when hydrogen gas burns in air according to the equation: $2H_2 + O_2 \rightarrow 2H_2O$ Which of the following is responsible for the heat?	(A) Breaking hydrogen bonds gives off energy. (B) Breaking oxygen bonds gives off energy. (C) Forming oxygen-hydrogen bonds gives off energy. (D) Both (a) and (b) are responsible. (E) (a), (b) and (c) are responsible. (F) I don't know.
8	A molecule of dinitrogen tetroxide, N_2O_4, contains an N-N bond. The bond can break to form two molecules of nitrogen dioxide, NO_2.  Which of the following statements is true?	(A) When the bond breaks, energy is absorbed (B) Energy is released when the bond breaks (C) When the bond breaks no energy will be absorbed or released. (D) It is not possible to predict whether energy would be absorbed or released when the bond breaks. (E) I don't know
9	A piece of iron railway track will get longer when it is in the hot sun, and that is why there are usually gaps between pieces of track. When the track expands the iron atoms:	(A) form bigger molecules so that they take up more space. (B) vibrate more so the atoms take up more space. (C) increase in size so they take up more space. (D) contain electrons that move faster so the atom takes up more space. (E) I don't know.
10	In a classroom containing metal and plastic chairs, people often say that the metal chairs feel colder than the plastic chairs. This is because:	(A) metal naturally attracts and holds cold. (B) plastic absorbs and retains heat. (C) the metal chairs are at a lower temperature than the plastic chairs. (D) metal quickly conducts heat away from your skin. (E) I don't know.

Conservation of Matter Cluster

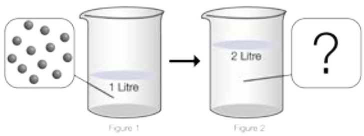
Item number	Stem	Options
11	Iron reacts with oxygen and water from the air to form rust. If an iron nail was allowed to rust completely, the rust would weigh:	(A) less than the nail it came from. (B) the same as the nail it came from. (C) more than the nail it came from. (D) it is impossible to predict. (E) I don't know.
12	The diagram below represents a mixture of S atoms and O_2 molecules in a closed container:  Which option represents the contents of the container after the original mixture reacts as completely as possible according to the equation: $2S + 3O_2 \rightarrow 2SO_3$ Key:  	 (A)  (B)  (C)  (D)  (E)
13	Which of the following must be the same before and after a chemical reaction?	(A) The sum of the masses of all the substances involved. (B) The number of molecules of all substances involved. (C) The number of atoms of each type involved. (D) The sum of the masses and the number of atoms of each type. (E) The sum of the masses, the number of molecules of all substance and the number of atoms of each type. (F) I don’t know.
14	A 1.0 g sample of solid iodine is placed in a tube and the tube is sealed after all of the air is	(A) less than 27.0 g because iodine gas weighs less than solid iodine. (B) more than 27.0 g because iodine

	removed. The total mass of the tube and the solid iodine is 27.0 g?  The tube is then heated until all of the solid iodine evaporates and the tube is filled with iodine gas. The total weight after heating will be:	molecules expand on heating. (C) less than 27.0 g because iodine gas is lighter than air. (D) 27.0 g because mass is conserved. (E) I don't know.
15	When a match burns, what happens to most of the mass of the match?	(A) It converted into heat. (B) The mass of the ash is the same as the mass of the match. (C) It is converted into gaseous carbon dioxide, water and heat. (D) It is converted into ash and heat. (E) It is converted into gaseous carbon dioxide and water. (F) I don't know.

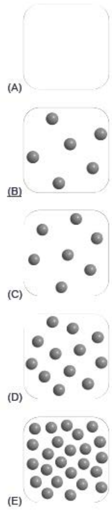
Aqueous Solutions Cluster

Item number	Stem	Options
16	When a teaspoon of calcium chloride ($CaCl_2$) dissolves in a glass of water, a clear liquid is formed. What happens to the calcium chloride?	(A) The calcium chloride and water form a new substance. (B) The calcium chloride separates into calcium atoms and chlorine molecules. (C) The calcium chloride molecules become separated and surrounded by water. (D) The calcium chloride forms calcium ions and chloride ions which are surrounded by water molecules. (E) I don't know.

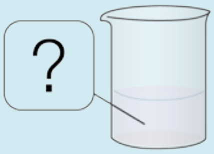
17 The diagram below represents a 1.0 L solution of sugar dissolved in water. The dots in the magnification diagram represent the sugar molecules. In order to simplify the diagram, the water molecules have not been shown.



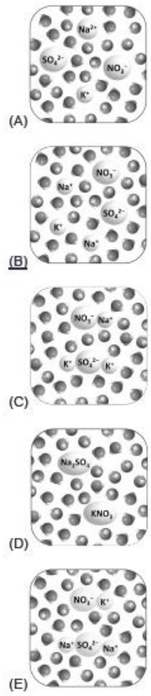
Which option represents the view after 1.0 L of water is added to the beaker in Figure 1?



18 Dilute solutions of potassium nitrate (KNO_3) and sodium sulfate (Na_2SO_4) are mixed together. There is no precipitate. The magnification diagram shows a view of a small portion of this solution:



Key



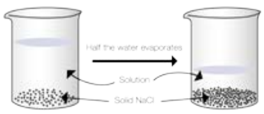
19 Equal volumes of solutions of sodium chloride (NaCl) and sodium bromide (NaBr) are mixed. The initial concentrations of the solutions are identical. The final concentration of sodium ions is?

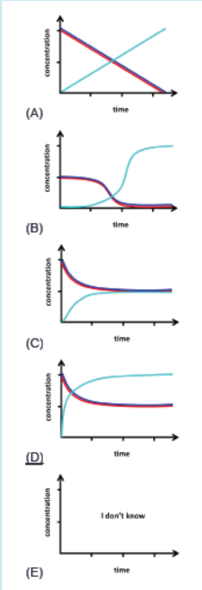
(A) double the initial concentration.
(B) the same as the initial concentration.
(C) half the initial concentration.
(D) not possible to determine without more information.
(E) I don't know.

20 Alum, $KAl(SO_4)_2$, is often used in water treatment plants. What ions would you find in a dilute alum solution?

(A) K^+ and $Al(SO_4)^{2-}$
(B) KAl^{4+} and SO_4^{2-}
(C) K^+ and Al^{3+} and SO_4^{2-}
(D) Al^{3+} and KSO_4^{2-} and SO_4^{2-}
(E) I don't know.

Chemical Equilibrium Cluster

Item number	Stem	Options
21	Which one of the following best describes a chemical equilibrium?	(A) Reactions continue with no effect on the concentrations of reactants and products. (B) Forward and backward reactions alternate. (C) Reactions have stopped. (D) There are equal concentrations of reactants and products. (E) I don't know.
22	In the beaker on the left, sodium chloride salt (NaCl) is added to water and the mixture is stirred until no more salt dissolves. The salt that does not dissolve is allowed to settle out:  What happens to the concentration of salt in solution if water evaporates until the volume of the solution is half the original volume? (Shown in the beaker on the right). You can assume that the temperature remains constant throughout the experiment. The concentration:	(A) increases because the salt does not evaporate and is left in solution. (B) decreases because more solid salt forms. (C) constantly changes depending on the rate of the reaction. (D) stays the same because salt crystalises as the water evaporates. (E) I don't know.
23	Hydrogen gas and iodine gas are mixed together in a container at constant temperature. Hydrogen iodide gas is formed according to the following chemical equation: $H_2(g) + I_2(g) \rightleftharpoons 2HI(g)$ Which graph shows the progress of the reaction to equilibrium?	

		
24	The two linked equilibrium reactions shown below are important in your blood, in the ocean and also in fizzy drinks. $2H_2O(l) + CO_2(g) \rightleftarrows H_2CO_3(aq) + H_2O(l) \rightleftarrows H_3O^+(aq) + HCO_3^-(aq$  If you open a bottle of fizzy water (without spilling any), bubbles rapidly form. The amount of water:	(A) increases, because the gas pressure in the bottle decreases and H_2CO_3 reacts to produce more CO_2 (B) decreases, because water evaporates when the CO_2 gas leaves the bottle. (C) constantly changes, because the new equilibrium is constantly changing. (D) stays the same, because water is present on both sides of the left hand equilibrium equation. (E) I don't know.
25	Formic acid is the simplest carboxylic acid, HCOOH. It is the cause of pain in bee and ant stings and nearly a million tonnes are made each year because it is a good preservative. In water, formic acid behaves as a weak acid: $HCOOH(aq) + H_2O(l) \rightleftarrows HCOO^-(aq) + H_3O^+(aq)$ The concentration of formate ions ($HCOO^-$) is increased by adding sodium formate (HCOONa). Which of the statements below is the best explanation of the process of returning to equilibrium after the sodium formate has been added?	(A) The HCOOH concentration will decrease, because it reacts with the added HCOONa. (B) The H_3O^+ concentration will increase, because the ratio of H_3O^+ to $HCOO^-$ remains constant at equilibrium. (C) The $HCOO^-$ concentration will increase and concentration of the other species will decrease to balance the reaction. (D) The H_3O^+ concentration will decrease, because some will react with $HCOO^-$ to return to the equilibrium product/reactant ratio. (E) I don't know.

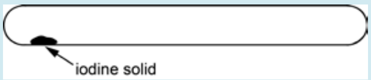
Appendix D

2014 Foundation Concept Diagnostic Instrument

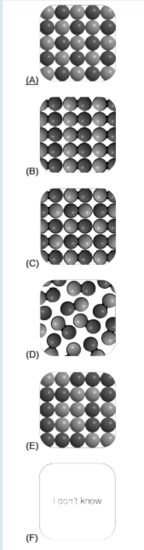
This instrument was delivered as an alternative option to the core chemistry concept instrument and is intended for use with students who have not completed high-school chemistry.

Behaviour of Matter Cluster

Item number	Stem	Options
1	In a pure sample of oxygen gas, in between the oxygen molecules there is:	(A) Matter. (B) Air. (C) Water vapour. (D) Nothing. (E) Atmosphere. (F) I don’t know.
2	If you could remove all the atoms of a chair, what would happen to the chair?	(A) Nothing would change. (B) The chair looks the same but weighs less. (C) A viscous liquid would remain. (D) Nothing of the chair would be left over. (E) I don’t know.
3	Iron can be obtained in the form of blocks as well as in the form of filings:  If you had an instrument with which you could visualize the atoms in the iron block and in the iron filings, what would you see?	(A) The iron atoms in the block and the filings are the same. (B) The iron atoms in the filings are darker coloured than the iron atoms in the block. (C) The iron atoms in the filings are rough while the iron atoms in the block have a smooth surface. (D) The iron atoms in the block are completely different to those in the filings. (E) The iron atoms in the filings are smaller than those in the block. (F) I don’t know.
4	A water molecule in the gas phase is ________ a water molecule in the solid phase.	(A) smaller than (B) lighter than (C) heavier than (D) larger than (E) the same mass as (F) I don’t know.
5	A 1.0 g sample of solid iodine is placed in a tube and the tube is sealed after all of the air is removed. The total mass of the tube and the solid iodine is 27.0	(A) less than 27.0 g because iodine gas weighs less than solid iodine. (B) more than 27.0 g because iodine molecules expand on heating. (C) less than 27.0 g because iodine gas is

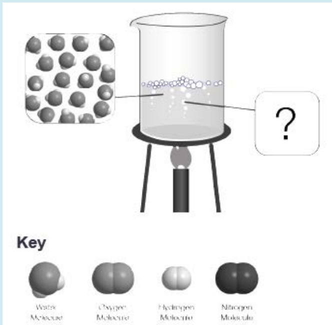
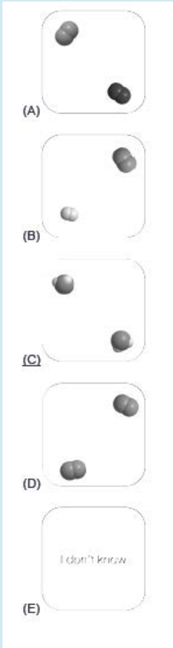
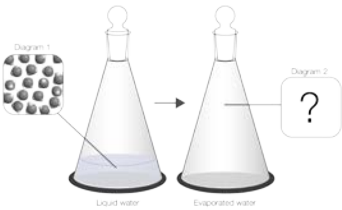
	g.  The tube is then heated until all of the solid iodine evaporates and the tube is filled with iodine gas. The total weight after heating will be:	lighter than air. (D) 27.0 g because mass is conserved. (E) I don't know.

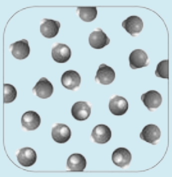
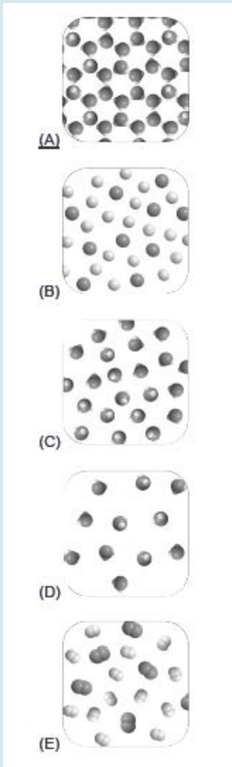
Nature of Matter Cluster

Item number	Stem	Options
6	The particles of a pure substance may be atoms, ions or molecules. Which of the following best represents the ionic solid, NaCl?	
7	The atoms of two different elements are represented as  and  . Which of the following best represents a possible molecular compound.	
8	Which of the following is characteristic of atoms?	(A) Dense. (B) Stationary. (C) Fixed size. (D) Mostly empty space. (E) Always spherical. (F) I don't know any characteristics of atoms.
9	Solid sodium chloride is composed of:	(A) A crystal lattice of sodium atoms sharing electrons with chlorine atoms. (B) A crystal lattice of sodium atoms and

		chlorine atoms. (C) A crystal lattice of sodium ions and chloride ions. (D) Molecules, each composed of a sodium atom and a chlorine atom bonded together. (E) Molecules, each composed of a sodium ion and a chloride ion attracted together. (F) I don’t know.
10	Which of the following must be the same before and after a chemical reaction?	(A) The sum of the masses of all the substances involved. (B) The number of molecules of all substances involved. (C) The number of atoms of each type involved. (D) The sum of the masses and the number of atoms of each type. (E) The sum of the masses, the number of molecules of all substance and the number of atoms of each type. (F) I don’t know.

States of Matter Cluster

Item number	Stem	Options
11	Assume a beaker of pure water has been boiling for 30 minutes. What is in the bubbles in the boiling water? 	
12	Diagram 1 shows a magnified view of a very small portion of liquid water in a closed container.  Which option represents the magnified view after the water evaporates, in diagram 2 on the right? Key 	
13	What happens when ice melts?	(A) The molecules in ice get smaller because liquid water takes up less volume than ice. (B) The molecules in ice get hotter because

		the liquid water is hotter than ice. (C) The molecules move faster in the liquid water than they did in ice. (D) Ice molecules change into water molecules to melt. (E) Heat causes molecules to expand leading to a separation of molecules during melting. (F) I don’t know.
14	A small dry glass is filled with ice and left to stand on a kitchen bench. Fifteen minutes later some of the ice has melted and the outside of the glass is covered in water droplets. Where has the water on the outside of the glass come from?	(A) Water from the melted ice evaporates and condenses on the outside. (B) The water from the melted ice passes through the glass. (C) Water vapour from the air condenses on the cold surface of the glass. (D) The coldness causes oxygen and hydrogen from the air to combine forming the water on the glass. (E) I don't know.
15	A diagram representing water molecules in the liquid phase is shown below.  Which of the diagrams below best shows what water would look like after it freezes (changes to a solid)?	

Temperature & Heat Cluster

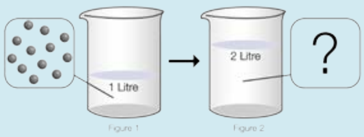
Item number	Stem	Options
16	Heat can be best described as:	(A) energy transfer between bodies at different temperatures. (B) friction from particles rubbing together. (C) a reading on a thermometer. (D) the absence of cold. (E) a substance that makes objects feel warm. (F) I don’t know.
17	In a classroom containing metal and plastic chairs, people often say that the metal chairs feel colder than the plastic chairs. This is because:	(A) metal naturally attracts and holds cold. (B) plastic absorbs and retains heat. (C) the metal chairs are at a lower temperature than the plastic chairs. (D) metal quickly conducts heat away from your skin. (E) I don't know.
18	A flat sheet of aluminium foil and a thicker aluminium baking tray are both placed in an oven at a temperature of 200 ^{o}C for a long period of time. If you touched these items, which would be more likely to burn you?	(A) (A) Both objects are equally likely to burn you because they are at the same temperature. (B) The baking tray is more likely to burn you because it is hotter. (C) The aluminium foil is more likely to burn you because it is hotter. (D) The baking tray is more likely to burn you because it has a higher mass. (E) I don’t know.
19	A piece of iron railway track will get longer when it is in the hot sun, and that is why there are usually gaps between pieces of track. When the track expands the iron atoms:	(A) form bigger molecules so that they take up more space. (B) vibrate more so the atoms take up more space. (C) increase in size so they take up more space. (D) contain electrons that move faster so the atom takes up more space. (E) I don't know.
20	When water molecules in the gas phase are heated, the molecules themselves:	(A) expand. (B) move faster. (C) become lighter. (D) change to a liquid. (E) release air. (F) I don’t know.

Solutions Cluster

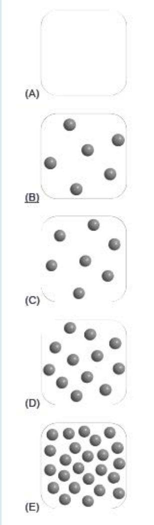
Item number	Stem	Options
21	Crystals of sugar are placed in a beaker of water. If the mixture is left to stand long enough without stirring, the sugar crystals eventually can no longer be seen, and the water will taste sweet?	(A) The sugar molecules gain heat energy from the surroundings and melt, forming a liquid. This liquid then mixes with the water. (B) Water molecules surround sugar molecules on the surfaces of the crystals and pull them away into the solution. (C) The sugar fills the air spaces in the water and 'disappears'. (D) The sugar crystals will only dissolve when stirred. Stirring breaks up the crystals into smaller particles that spread throughout the water and so can no longer be seen. (E) I don't know.
22	What happens to the concentration of sodium chloride in a dilute solution if half the volume of water evaporates? You can assume the temperature is constant throughout. The concentration will:	(A) increase. There is the same amount of salt in less water. (B) decrease. Solid salt will form. (C) stay the same. The salt and water both evaporate. (D) stay the same. Ions are constantly leaving and returning to the solid. (E) I don't know.
23	Sodium chloride, NaCl, is used in daily life for many food, health and industrial purposes. What would you find when you dissolve sodium chloride in water?	(A) Na^+ and Cl^- ions. (B) NaCl molecules. (C) Very small NaCl crystals. (D) Na^+Cl^- particles. (E) I don't know.
24	When a teaspoon of calcium chloride ($CaCl_2$) dissolves in a glass of water, a clear liquid is formed. What happens to the calcium chloride?	(A) The calcium chloride and water form a new substance. (B) The calcium chloride separates into calcium atoms and chlorine molecules. (C) The calcium chloride molecules become separated and surrounded by water. (D) The calcium chloride forms calcium ions and chloride ions which are surrounded by water molecules. (E) I don't know.

25

The diagram below represents a 1.0 L solution of sugar dissolved in water. The dots in the magnification diagram represent the sugar molecules. In order to simplify the diagram, the water molecules have not been shown.



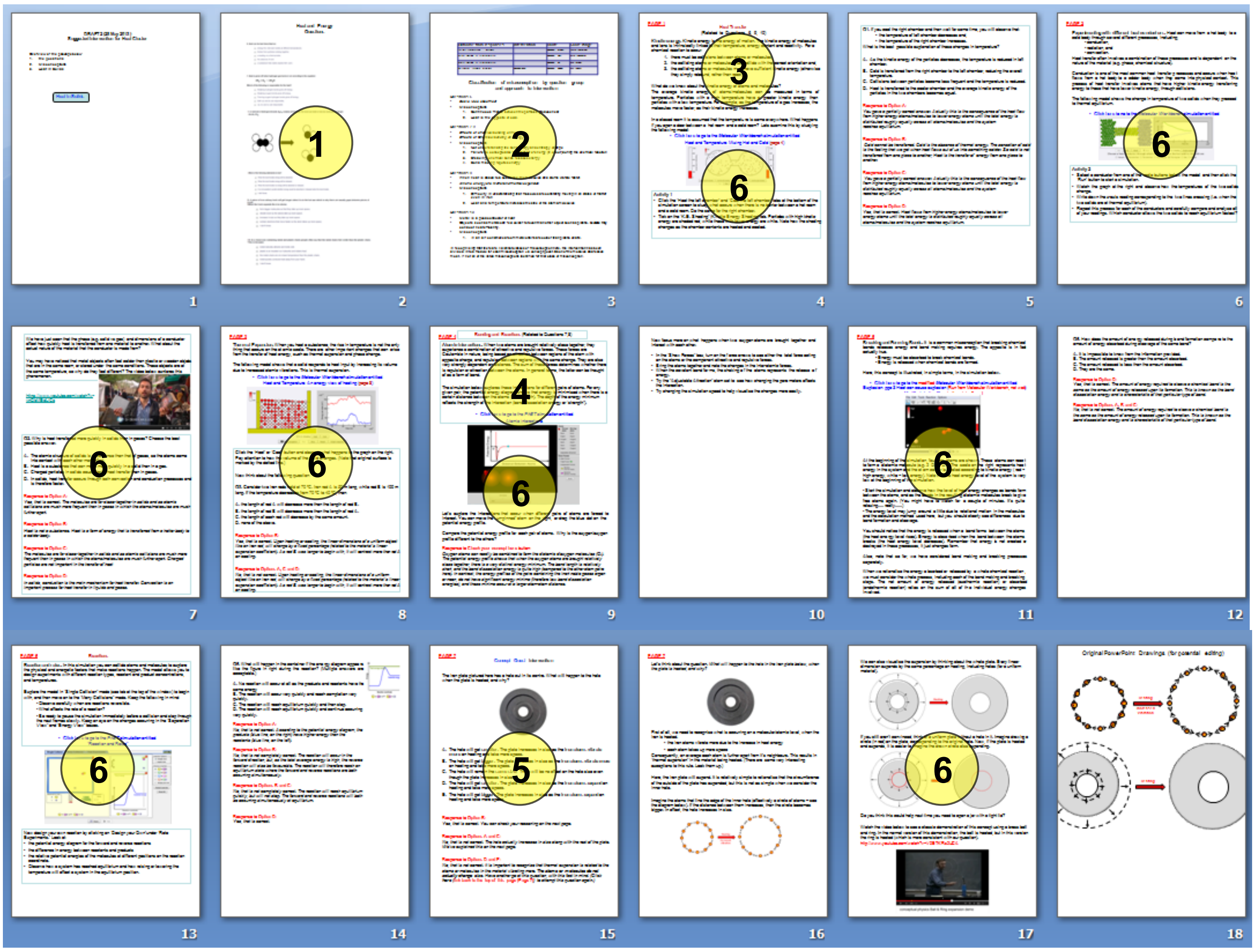
Which option represents the view after 1.0 L of water is added to the beaker in Figure 1?



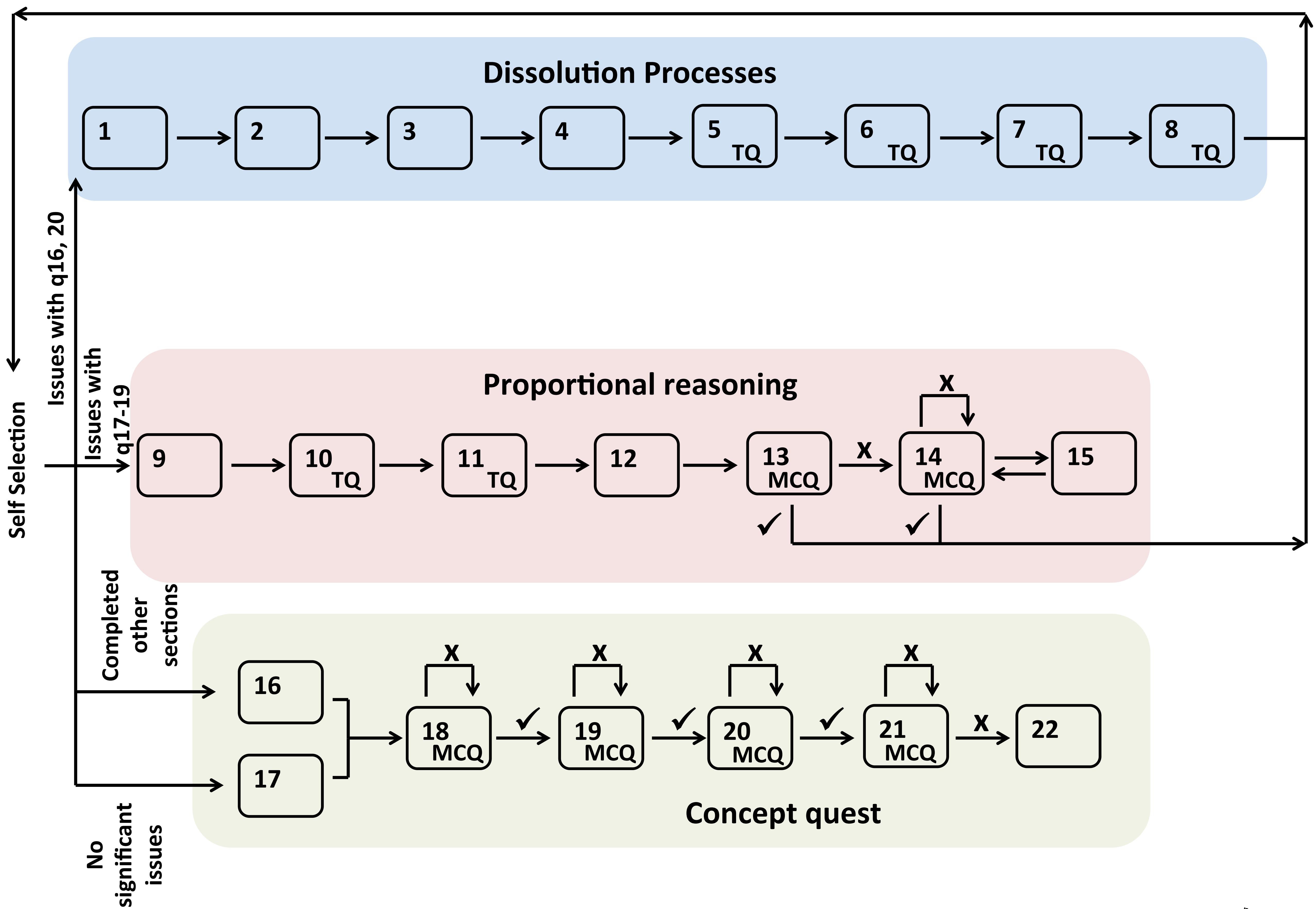
APPENDIX E
Storyboard for online module for Heat and Energy Cluster

Overview of the package below:

1. the questions
2. a description of the apparent misconceptions revealed by the questions
3. first intervention screen for misconceptions based around heat transfer processes
4. first intervention screen for misconceptions based around bonding and reactions
5. first screen for Concept Quest
6. Interactive module: either a url link to an interactive online visualisation (Molecular Workbench or PhET; two VisChem animations or a Youtube video (may contain short answer or multiple choice questions).



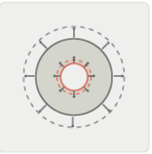
ReSOLv Website Module Workflow



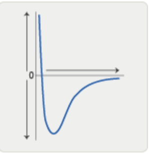
Heat and Energy

Heat Transfer

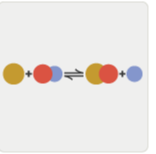
Heat Transfer



Thermal Expansion



Making Bonds



Energy and Reactions

Concept Quest

After completing this module, you should be able to:

- Grasp the relationships between heat/thermal energy, particle kinetic energy and temperature
- Understand the principles of heat transfer
- Predict, at a basic level, the efficiency of on object to conduct heat, based on the nature of the material and its shape

Overview

Heat and temperature are concepts that we are all familiar with, on the surface. Yet, if we look closely at our understanding of these concepts we sometimes find that they are poorly defined in our own minds, and can be difficult to explain clearly in plain language. In this module, you will be introduced to the relationships between heat/thermal energy, particle kinetic energy and temperature, and explore simulations to clarify your understanding of heat transfer.

Heat and Kinetic energy

Kinetic energy is the energy of motion. The temperature, energy content and reactivity of a material is intrinsically linked to the kinetic energy of the molecules and/or ions that make up that material. For example, for a chemical reaction to occur between two particles (atoms, ions or molecules):

1. there must be collisions between the particles,
2. the colliding particles must collide with the correct orientation and,
3. the colliding particles must have sufficient kinetic energy (otherwise they simply rebound, rather than react).

What do we know about the kinetic energy of atoms, ions and molecules?

The average kinetic energy of a group of particles (atoms, ion or molecules) can be measured in terms of temperature. In other words, temperature is a measure of the average kinetic energy of a group of particles. We can think about this in a number of ways.

- A group of particles with high average kinetic energy will have a higher temperature than a group of particles with lower average kinetic energy.
- If two bodies are at the same temperature, the particles that make up Body 1 will have the same average kinetic energy as those that make up Body 2.
- For example, as the temperature of a gas increases, the kinetic energy of the gas molecules increases, and the molecules move faster (obviously, their mass remains constant).

Heat, or thermal energy, is related to the ***total*** amount of kinetic energy possessed by the particles that make up the object of interest. We can conclude from this, that if two objects are made of the same material, and are at the same temperature, the one with the largest mass will contain the most thermal energy. If the objects are made of different materials however, the amount of thermal energy possessed by each will depend on the nature of each material.

Heat Transfer

If temperature is related to the average kinetic energy of the particles involved, how does heat move from one material or place to another?

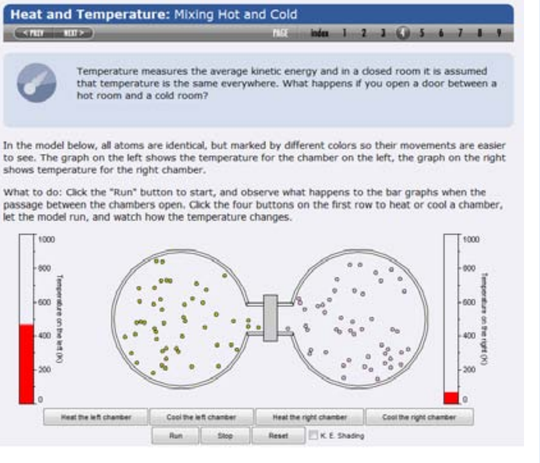
Heat can move from a hot body to a cold body through several different processes, including;

1. **conduction,**
2. **radiation, and**
3. **convection.**

Heat transfer often involves a combination of these processes and is dependent on the nature of the material (e.g. phase, chemical structure).

Molecular Workbench "Heat and Temperature" Activity

In the *"Heat and Temperature: Mixing Hot and Cold"* activity, you can explore the transfer of heat from a closed hot room to a cold room, by opening the door that separates them. In a closed room it is assumed that the temperature is same everywhere.



- Click the "Heat the left chamber" and "Cool the left chamber" buttons at the bottom of the simulation screen to study what occurs when there is no barrier between a hot room and a cold room. Do the same for the right chamber.
- Tick on the "K.E. Shading" (Kinetic Energy Shading) tab. Particles with high kinetic energy are shaded red, while those with lower energy are white. Note how the shading changes as the chamber contents are heated and cooled.

If you cool the right chamber and then wait for some time, you will observe that:

- the temperature of left chamber decreases and,
- the temperature of the right chamber increases.

What is the best possible explanation of these changes in temperature?

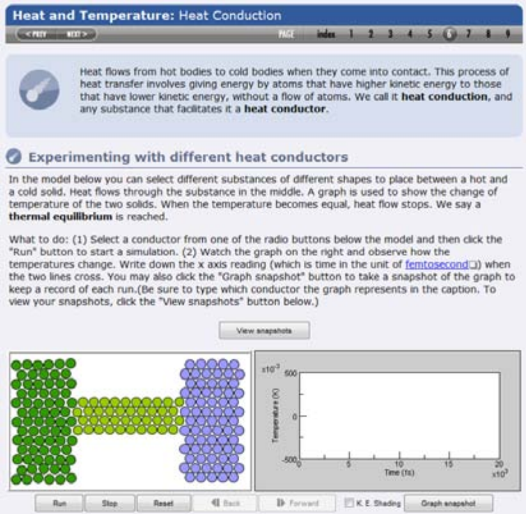
a) As the kinetic energy of the particles decreases, the temperature is reduced in left chamber.
b) Cold is transferred from the right chamber to the left chamber, reducing the overall temperature.
c) Collisions between particles become less frequent and the temperature is reduced.
d) Heat is transferred to the cooler chamber and the average kinetic energy of the particles in the two chambers becomes equal.

Conduction

Conduction is one of the most common heat transfer processes and occurs when heat flows from a hot body to a colder body when they come into physical contact. This process of heat transfer involves atoms that have higher kinetic energy transferring energy, through collisions, to those that have lower kinetic energy. The efficiency of conduction depends on the nature and shape of the materials involved.

Molecular Workbench "Heat and Temperature" Activity

This Molecular Workbench "Heat and Temperate: Heat Conduction" model allows you to explore the change in temperature of two solids as they proceed to thermal equilibrium.



- Select a conductor from one of the radio buttons below the model and then click the 'Run' button to start a simulation.
- Watch the graph at the right and observe how the temperatures of the two solids change.
- Write down the x-axis reading corresponding to the two lines crossing (i.e. when the two solids are at thermal equilibrium).
- Repeat this process for each of the conductors and carefully compare and analyse all of your readings. Which conductor allows the two solids to reach equilibrium fastest?

Why is heat transferred more quickly in solids than in gases? Choose the best possible answer.

a) The atomic structure of solids is more dense than that of gases, so the atoms come into contact with each other more often.
b) Heat is a substance that can move more quickly in a solid than in a gas.
c) Charged particles in solids cause faster heat transfer than in gases.
d) In solids, heat transfer occurs through both convection and conduction processes and is therefore faster.

The Nature of the Conductor

We have just seen that the phase (e.g. solid *vs* gas) and dimensions of a conductor affect how quickly heat is transferred from one material to another. What about the actual nature of the material that the conductor is made from?

You may have noticed that metal objects often feel colder than plastic or wooden objects that are in the same room, or stored under the same conditions. These objects are at the same temperature, so why do they feel different? This video explores this phenomenon.

Appendix F
Evaluation Framework & External Evaluator's Report

ID12-2277: Enhancing the Secondary–Tertiary Transition in Chemistry through Formative Assessment and Self-regulated Learning Environments

Authentic nature of this project

This project, funded by the Office for Learning and Teaching (OLT), addresses a complex issue at the centre of challenges to science education in universities world-wide. There is an extensive literature – going back several decades – on the deeply entrenched unhelpful conceptions many students have about many fundamental scientific concepts. There is also a great deal of evidence that many interventions to teach these concepts 'better' have not been successful, and that students revert to their earlier beliefs or live in a 'two worlds' culture of reproducing what is expected for assessment while, at the same time, retaining what is actually believed (Solomon, 1983).

It has been a privilege to work with a project team determined to break through this seeming impasse of assisting students to negotiate their unhelpful conceptions in chemistry. The team comes from five universities of different character, all having diverse first-year chemistry cohorts, and all having some differences from each other in the overall patterns of diversity. This microcosm of the Australian higher-education sector has brought challenges to the project but also a much-needed element of authenticity. One consequence of this diversity was the decision to tailor the intervention design to the needs of each institution.

Conceptual-change theory

The details of the project are in the main report and I will not repeat them here. The point I would like to emphasize is how difficult it is to change mental models about *abstract* concepts. The topics chosen for this project are phase change, heat and energy, chemical equilibria, aqueous solutions, and conservation of matter. All of these topics can only be understood by forming appropriate mental models of matter and of explanations to explain observed properties of matter. This is hard stuff!

Like most previous interventions, the project has broadly adopted conceptual-change theory as an overall framework for its endeavours. A quick summary is as follows: The pioneering work of Lewin (1952) on bringing about social change through group decision-making focused on considering how to change deep-seated beliefs. He suggested a three-step procedure: unfreezing, moving and then freezing at the new position. Others have subsequently utilized this work in educational contexts, mostly in science education. Nussbaum and Novick (1982) and West (1985) described a similar three-stage process for bringing about conceptual change in these terms (McNaught & Curtis, 2009).

1. a process for diagnosing existing conceptual frameworks and revealing them to those involved;
2. a period of disequilibrium and conceptual conflict which makes the subject dissatisfied with existing conceptions; and
3. a reforming or reconstruction phase in which a new conceptual framework is formed.

1. Stage 1 of this project required the development of a set of validated concept inventory items. I am very satisfied that this has been done in a thorough and rigorous fashion.

2. The second stage requires students to consciously let go of prior views before accepting new and more helpful explanations. Students do this by working through optional media-enriched online activities. The optional nature of the interventions deserves a comment. The lack of compulsory ("is it for marks?") activity has inevitably reduced student uptake but, arguably, those students who do the interventions are motivated by wanting to learn. Being willing to engage in the hard work of understanding abstract concepts is a pre-requisite for learning. One design feature that is worth highlighting is that, at The University of Queensland (UQ) in semester 1 2014, the modules were embedded in a sequence that was linked to summative assessment; good take-up with student self-directed access to the interventions was achieved with 800 out of 1500 still opting to use the materials right until the end of semester.

   The diversity across the institutions resulted, as mentioned above, in tailored approaches for interventions. These approaches were *context-specific* (depending on student prior learning experiences and learning environment); *flexible* (designed to be adapted to fit most first-year chemistry curricula); and *modular* (enabling teachers to 'pick & mix'). More will be noted about these three design principles below.

3. The third stage is supported by a number of test items that students can use to test their new ideas.

Stages 1 and 3 still need further work in order for the system to give automated individualised rich formative feedback to students from the diagnostic items. The survey tools and learning-management system (LMS) quiz tools all deliver scores which need further development to link to multiple pathways in the interventions. This is ongoing work that is clearly in a direction that is likely to result in students being well-supported in untangling and personalizing the learning of these complex concepts in chemistry.

Enacting the evaluation plan

The first evaluation plan was produced in February 2013 after a project meeting and a number of email iterations. The complexity of the task was recognized at the outset and the evaluation plan was detailed and ambitious right from the outset (an annotated version is in the Appendix to this evaluator's report). Given the resistance with which unhelpful conceptions are held, this detailed project planning was entirely appropriate.

The evaluation data set is complex and has increased in complexity as the interventions were (very sensibly) tailored to each institution in the project. This has meant that analysis is still ongoing. However, I am satisfied with the nature of the analysis, and completely agree with the need to gain authentic and relevant evidence about how well the interventions supported student learning, and how lasting this learning remains.

The 'LEPO' (Learning, Environment, Processes, Outcomes) conceptual framework for curriculum design (Phillips, McNaught, & Kennedy, 2010, 2011) has informed thinking about the context and interactions involved in curriculum design in this project. Within a *learning environment*, students attain *learning outcomes* by going through *learning processes.*

The project has been well-managed. For example, the team maintained all records on project-management software ('Basecamp'). Weekly Skype meetings of the project team were essential to the success of the work.

The major challenge has been time – see my comments below. The time-lines for analysis activities have slipped. This is understandable with such an ambitious project but it is an issue that OLT needs to recognize and consider.

Completing the final project report in a timely fashion was a challenge and I would like to acknowledge the tenacity of the Project Leader, Dr Gwen Lawrie, in her leadership of the process of collation of a vast amount of data and extraction of the key themes and outcomes.

I have maintained contact with the group as a ‘critical friend’ at regular intervals. I am privileged to be a participant observer in what I consider to be an excellent example of design-based research (DBR). The outcomes of the project are contributing to theoretical understandings as well as practical resources and guidelines. It is this dual nature of theoretically informed planning and evidence-based outcomes that makes DBR so useful for projects such as OLT projects.

Reeves (2006) described the iterative nature of DBR as involving:

- analysis of practical problems;
- development of solutions based on existing knowledge;
- evaluation research of the solution in practice; and
- reflection to produce *design principles*.

It is not the purpose of this brief report to examine the intricacies of DBR; additional useful references are Barab and Squire (2004); Herrington, Reeves, and Oliver (2009); Phillips, McNaught, and Kennedy (2011); Van den Akker et al. (2006); and Wang and Hannafin (2005). Rather it is to record that this project is a good example in the Australian context of DBR in university science education, and has *both* a good theoretical basis *and* the potential for good evaluation evidence. This was the same comment that I made about an earlier OLT chemistry project (CG9-1112: IS-IT learning? Online interdisciplinary scenario-inquiry tasks for active learning in large, first-year STEM courses), with UQ as the lead institution.

Design principles for interventions in teaching fundamental abstract concepts in chemistry

As noted above, the three design principles are (1) *context-specific* (depending on student prior learning experiences and learning environment); (2) *flexible* (designed to be adapted to fit most first-year chemistry curricula); and (3) *modular* (enabling teachers to ‘pick & mix’).

This project illustrates what I consider to be a healthy trend away from a one-size-fits-all solution to educational issues and, instead, towards a focus on accommodating diversity of contexts, as well as academics’ diversity in appetite and capacity to bring about change in their own contexts. As this project has developed, it has become clear that academics teachers of chemistry in Australia are not seeking neatly packaged complete resources. They appear to want access to ideas and resources that can be adapted to their own contexts and student needs.

From an evaluation point of view, this is much messier, but then authentic environments usually are. For example, the semester 1 2014 delivery data is diffuse because all the institutions adapted the intervention to their own contexts: timing, level, learning-management system, student prior experiences and learning environments. However, this appears to be necessary to achieve the outcome of flexibility in the diagnostic/ intervention options for adaptation so that scalability becomes viable.

Overall comments on issues highlighted by the project

It is my considered opinion that the project has been conducted well and is successful in its outcomes at this stage of analysis. However, it is important to examine a bit further the unfinished state of the data analysis.

The project is necessarily complex: genuinely tackling how to support the learning of individual students in large classes in diverse institutional settings across Australia cannot be done in a one-size-fits-all approach. As a result, in similar projects, I have three suggestions for OLT to consider:

1. *More careful scrutiny of the suggested time-lines*. In any project, it takes about a semester to get agreements in place and the project team gelling into a working unit. Also, once the teaching innovation has been designed and implemented, there needs to be more time for embedding and evidence of sustainability to emerge. In my view, this project needed to be, say, a semester longer.

2. *Careful scrutiny of any technical requirements and clear allocation of appropriate resources in the budget*. The project team are 'experts' in content and teaching but not in websites and course management platforms. More dedicated technical assistance would have been beneficial.

3. *Ensuring that data analysis is well-supported.* This project has amassed a great deal of data – both qualitative and quantitative. Analysing educational data and writing educational reports and papers is quite different to doing research in chemistry. OLT may wish to consider developing guidelines to assist projects in resourcing this vital aspect of the work. I think this project has the potential to make a significant contribution to the international chemistry education community but there is still a great deal of analytic work that needs to be completed. Team members in this project proactively discussed a publication strategy throughout the project, with apparent consensus which involved individual teams at each institution leading a paper. However, in reality, this has not happened due to time pressures.

   There is a great deal of talk about the scholarship of teaching and learning (SoTL) in the higher-education sector. OLT may wish to consider how best to ensure that maximum benefit is derived from OLT-funded projects through complete evidence of project outcomes being produced and disseminated, not only at conferences and discipline meetings but also in scholarly publication venues.

References

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14.

Herrington, J., Reeves, T. C., & Oliver, R. (2009). *A guide to authentic e-learning*. New York and London: Routledge.

Lewin, K. (1952). Group decision and social change. In G. E. Swanson, T. M. Newcomb & F. E. Hartley (Eds.). *Readings in social psychology* (pp. 459–473). New York: Holt.

McNaught, C., & Curtis, A. (2009). Using policy initiatives to support both learning enhancement and language enhancement at a Hong Kong university. In P. Coverdale-Jones & P. Rastall (Eds.). *Internationalizing the university: The Chinese context* (pp. 85–104). Basingstoke: Palgrave MacMillan.

Nussbaum, J., & Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy. *Instructional Science*, *11*, 183–200.

Phillips, R. A., McNaught, C., & Kennedy, G. (2010). Towards a generalised conceptual framework for learning: the Learning Environment, Learning Processes and Learning Outcomes (LEPO) framework. In J. Herrington & W. Hunter (Eds.), *ED-MEDIA 2010* (pp. 2495–2504). Proceedings of the 22nd annual World Conference on Educational Multimedia, Hypermedia & Telecommunications, Toronto, Canada, 28 June–2 July. Chesapeake VA: Association for the Advancement of Computers in Education.

Phillips, R. A., McNaught, C., & Kennedy, G. E. (2011). *Evaluating e-learning: Guiding research and practice.* Abingdon, Oxon.: Routledge.

Reeves, T. C. (2006). Design research from a technology perspective. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 52–66). London: Routledge.

Solomon, J. 1983. Learning about energy: How pupils think in two domains. *European Journal of Science Education*, *5*(1), 49–59.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). *Educational design research*. Abingdon, U.K.: Routledge.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), 5–23.

West, L. H. T., & Pines, A. L. (Eds.). (1985). *Cognitive structure and conceptual change*. New York: Academic Press.

Carmel McNaught
Emeritus Professor of Learning Enhancement
The Chinese University of Hong Kong

October 2014

Appendix to External Evaluator's Report
Annotated Evaluation Plan

The first evaluation plan was produced in February 2013 in a form very similar to this version. Regular and systematic updates on the progress of data collection have taken place and been reported to the whole project team through Basecamp. The evaluation plan has thus been a working document that has enabled the team (especially at UQ) to consider the project in a holistic fashion.

The evaluation data set is complex and has increased in complexity as the interventions were (very sensibly) tailored to each institution in the project. This has meant that analysis is still ongoing. However, I am satisfied with the nature of the analysis, and completely agree with the need to gain authentic and relevant evidence about how well the interventions supported student learning, and how lasting this learning remains.

This is a complex evaluation plan. Many of my comments are repeated in my main report so that they are not 'lost' in the myriad of detail.

Aspect & possible evaluation questions	***Data sources***	***Data collected as of July 14 plus comments by External Evaluator in September 14***
1. Learning environment		
A. Design of diagnostic		
1.1 What are the five core concepts to address? 1.2 What are good items to probe alternate conceptions? 1.3 Which items will provide answers that can act as triggers for the intervention?	• Project team item evaluation processes seen by logs and minutes on BaseCamp	• Minutes (not detailed) and Basecamp logs. • Evolution of question development is clear on Basecamp.
	• Student focus group (FG)	• Nov 12 & Nov 13. These interviews have been coded to identify themes related to support students seek.
	• Comparison to literature instruments	• Ongoing comparison to literature items. • Validation of 2013 questionnaire shown by expert-review process (iterative refinement), and psychometric testing against multiple contexts. This has been completed in a rigorous fashion. • Analysis of outcomes for 2014 clusters and Rasch analysis to assess viability of using clusters independently as a research instrument or only for practitioner feedback. This has also been completed and is important for achieving flexibility in the diagnostic/ intervention options.
B. Delivery of formative feedback		
1.4 How do we format and deliver effective feedback? 1.5 How do we know if students have reviewed	• Design of feedback in terms of hints not answers: degree of coherence of the feedback with literature	• Participant feedback from the FYHE Nuts & Bolts Workshop Dissemination (Jul 13) and the two ACSME ideas exchange sessions (Sept 13)

feedback?	findings on students' alternative conceptions	
	• Can we track views or have a receipt on opening feedback?	• A 'persistence' rate has been tracked through how many students accessed the resources. • Stimulated-response interviews completed with 10 students in Sept-Oct 13. Also see the comment under 3.1.
	• Student FG	• Interviews held at QUT, UQ and Wollongong in Jul 13 • At UQ and Wollongong in Nov 13 • These interviews have been coded to identify themes related to support students seek.
	• Questions in end-of-semester evaluation	• Online questionnaire completed at end of semester, Oct 13 & Jun 14 • Have students who have consented for us to look at their Chemistry Review Online Module (CROM) concept questions and this can be correlated with overall access to online modules.
C. Design of interventions		
1.6 What type of activity works really well? 1.7 Which aspects of the activities appear to be critical?	• Evolution: rationale for versions over course of project – project team reflections and coherence with literature findings on good learning design	• There is substantial project data showing the different mechanisms that have been trialled and evaluated including: UWS website, on paper, face to face and also via Blackboard (Bb). • Discussions have occurred about fitting the analysis to a design-based research framework. The project overall is an excellent example of using the principles of design-based research.
	• Web logs	• Collected but low student uptake in some cases. • Emerging patterns in students accessing online modules identified through exploration of data logged using Bb and Google analytics.
	• Student FG using critical incident	• Addressed in part by stimulated-response interviews.
1.8 Is there enough scaffolding?	• Student FG using critical incident	• Addressed in part by stimulated-response interviews. • Open question in semester 2, 2013 end-of-semester online questionnaire and focus groups
1.9 Was the web interface easy to use?	• Student FG using critical incident	
	• Web logs	• Have this data through the UWS website, though UWS use the same platform for other initiatives and so it is difficult to assign uptake directly to project activities.
D. Sustainability		

1.10 How can these diagnostic/ feedback/ interventions be translated between contexts?	• Evidence from pilot processes implementing for each type of LMS and remediation required	• This might be presented nicely using a form of narrative research based on reflections from the range of folk involved at each institution. 1. Team member reflections through audio recording of Jun 13 & Jan 14 day project meetings. 2. Skype minutes and semester 2 question development process. • Online questionnaire completed at end of semester 1 2014. • Overall, the case studies presented have captured the essence of how diversity across institutions has been accommodated.
1.11 Why would academics bother with an intervention?	• Interviews with academics who express interest in implementing the diagnostic/ therapy	• Expressions of interest from 13 academics in becoming partners in 2014. • Feedback from workshop collected. • Invitations to participate in 2014 delivery resulted in one academic (WA) requesting resources.
1.14 What are the costs of implementation?	• Post-activity survey with open questions to evaluate the process and resourcing for delivering the diagnostic/ therapy • Design flexibility	• Costs to date available to make an estimate for each activity. This data has not yet been mined.

2. Learning process		
A. Student engagement		
2.1 What motivates students to do the interventions once they have received their feedback? 2.2 What motivates students to persist when their alternate conceptions are deeply embedded?	• Student feedback (FG/ end-of-semester survey)	• Have semester 1, 2013 data from FG interviews: UQ, QUT & Wollongong & semester 2 data from UQ and UoW. • Motivated Strategies Learning Questionnaire (MSLQ, 4 clusters) delivered at end of semester 2, 2013 (UQ). MSLQ analysis and initial triangulation to determine student strategies in seeking resources to support self-directed learning has been completed.
B. Formative feedback		
2.4 What is the best format for providing students with useful feedback without giving out the answers? 2.5 Do students know what to do with feedback of this type?	• Process of embedding hints into exercises (see 1 B) • Peer workshop activities – UoW & UWS	• We have semester 1, 2013 data from FG interviews: UQ, QUT & Wollongong and semester 2 data from UQ. • One design feature that I would like to comment on is the optional nature of the interventions. This inevitably has reduced student uptake but, arguably, those students who do the interventions (CROMs) are motivated by wanting to learn. One design feature that is worth highlighting is that, at UQ in

		semester 1 2014, the CROMs were embedded in a sequence that was linked to summative assessment; good take-up with student self-directed access to the interventions was achieved with 800 out of 1500 still opting to use the CROMs right until the end of semester.
C. Concept reconstruction		
2.6 How do we know if the activity has challenged or remodelled conceptions?	• Concept check questions embedded in activity • Final assessment results; this will involve focusing on performance on specific examination questions	• Student outcomes in pre/post (low numbers) • Feedback in FGs but perception data not a strong indicator. • Question embedded in final exam at UQ semester 1, 2014. • This has been a major discussion point; it is difficult to distinguish the effects of specific intervention activities from other course activities. • In semester 1 2014, four institutions planted a question in the exam to look at retention but this data has not yet been analysed.

3. Learning outcomes		
A. Metacognition		
3.1 Do students recognize any processes of cognitive conflict when they change their conceptions? 3.2 Is there evidence of student awareness of their own conceptions and ways of learning chemistry?	• FG Interviews • Open comment on post-survey "How has the diagnostic and intervention changed the way you think about learning chemistry" "On reflection, was this a useful activity for students entering university chemistry?" • Reflection at end of intervention? (is this possible) or run a focus group	• This has not been addressed yet – but will be in part by stimulated-response interviews. A series of these interviews is still in train as part of a PhD study; four have been completed and another six are planned. This should provide much more detailed information and insights than usually is obtained in focus groups.
B. Learning gains		
3.4 How do we know the apparent remodelling of a conception is retained?	• Student Assessment of their Learning Gains (SALG; http://www.salgsite.org/) instrument in end-of-semester evaluations	• UQ semester 1 2014 End of semester survey has MSLQ and SALG questions related to online learning activities (CROMs). • There seems to be a strong intention to get firmer learning-outcome data. With two extension proposals being prepared, and with my own knowledge of the rigour and commitment shown by the team, I am very hopeful that this project, over the next year or two, will make a significant contribution to the field of learning fundamental concepts in chemistry.

Evaluation of the project overall		
Questions	***Data for these five questions***	***Comments by External Evaluator in September 14***
1. Has the project resulted in a pragmatic and flexible set of learner resources? 2. Has the project resulted in a pragmatic and flexible set of teacher guidelines? 3. Is this project feasible within the context of Australian higher-education institutions at present? 4. To what extent does diffusion of the project require a local champion?	All of the above plus • Adherence to the project timelines. • Sense of time pressures – regular feedback at Skype meetings – minutes for data • Response to dissemination activities (e.g. concerned about pragmatics of academic workload and those interested seeing themselves as being mainstream and not off on a limb).	• Overall, as noted elsewhere, I am highly impressed by the project team and the way in which this project has been conducted. • The weekly Skype meetings have been an excellent way to keep a distributed project on track. • The use of Basecamp has also been excellent in keeping all the project team informed of developments. All documents can also be easily found and retrieved. • As this project has developed, it has become clear that academics teachers of chemistry in Australia are not seeking neatly packaged complete resources. They appear to want access to ideas and resources that can be adapted to their own contexts and student needs. The design of the ChemBytes website supports this notion and is gaining good feedback: https://shire.science.uq.edu.au/bb/CHEM1100/Index.html#/Home/Main • I have been concerned for some time in a number of projects about the need for a local champion who kick-starts and sustains innovation and change at an institution. This project illustrates what I consider to be a healthy trend away from a one-size-fits-all solution to educational issues and, instead, towards a focus on accommodating diversity of contexts, as well as academics' diversity in appetite and capacity to bring about change in their own contexts. • From an evaluation point of view, this is much messier, but then authentic environments usually are. For example, the semester 1 2014 delivery data is diffuse because all the institutions adapted the intervention to their own contexts: timing, level, learning-management system, student prior experiences and learning environments. However, this appears to be necessary to achieve the outcome of flexibility in the diagnostic/ intervention options for adaptation so that scalability becomes viable.