be remembered in similar problem-solving situations (*e.g.*, a mastery test) but not available for use in new venues such as equilibrium applications (Whitehead, 1929; Brown *et al*., 1989).

The current means of instruction, in which abstract stoichiometry procedures are taught detached from their use in real-world chemistry and held in abeyance for future use, is not working. Even after repeated instruction (high school courses, review in college courses), many chemistry students do not exhibit the fluid and flexible use of stoichiometry in their subsequent coursework. Clearly, instruction that leads to proficiency with solving algorithmic exercises does not translate to the learning of chemical concepts (Nurrenbern and Pickering, 1987; Sawrey, 1990; Nakhleh, 1993; Nakhleh and Mitchell, 1993). An inability to recognize relationships among chemical concepts prevents students from applying their knowledge in new problem situations (Sumfleth, 1988). In order for learners to develop the highly interconnected knowledge frameworks necessary for complex equilibrium and acid-base problem solving, the topic of stoichiometry must move from being simply a collection of tools to being those tools in use (Evans *et al*., 2006).

Findings from research on cognition can inform an instructional design that makes stoichiometry both easier and more interesting to learn: Instruction should promote a learner's active construction of knowledge (Piaget, 1954; Wheatley, 1991; Simon, 2000.), should be mindful of the learner's limited processing capacity (Miller, 1956; Simon, 1974; Sweller, 1994; Johnstone, 1997), and should provide multiple opportunities for encoding information into the learner's long-term memory (Tulving and Thomson, 1973; Gick and Holyak, 1983; Bjork and Richardson-Klavhen, 1989, Ericsson and Charness, 1994). In addition, the level of intellectual sophistication needed to both understand and work on complex equilibrium and acid-base problems as well as on ill-defined real-world problems may best be realized through *authentic inquiry* (Chinn and Malhotra, 2002) in which tools such as stoichiometry are developed as needed in the planning and executing of experiments and in the interpretation of data (Evans *et al*., 2006). But authentic activity in chemistry is often too dangerous, too time consuming, or too obscured by the interaction of multiple variables to be of cognitive value to learners. Furthermore, without the experimental skills required for successful laboratory work, the quality of data from which inferences are made is questionable. The instructional and learning challenges of the stoichiometry toolbox are ones that a *technology-rich learning environment* (Lajoie and Azevedo, 2006) may be equipped to address.

Distinctive features of online technology include abilities to dynamically explicate abstract information, to provide timely feedback for practice, and to scaffold the execution of complex tasks so that learners focus on knowledge relationships rather than individual bits of information. More than a decade ago Osin and Lesgold (1996) proposed that intelligent computer systems coupled with domain simulations might facilitate a cognitive apprenticeship model of learning by which novices (the students) could be supported by experts (in this case, the computer) as they solve authentic, albeit difficult, tasks in the process of developing competency in a domain such as chemistry. These interactive opportunities may promote the construction of a fluid and flexible knowledge framework by actively engaging students in revision of and building on their current understandings through exploration and reflection.

Interactive learning opportunities facilitated by online homework activities have been used to support stoichiometry instruction. Arasasingham *et al*. (2005) has compared the learning of stoichiometry by introductory college students assigned traditional text-based homework with that of students working with a Web-based homework program that incorporates molecular-level visualizations and timely formative feedback. Overall, students working with the Web-based program were better at conceptual problem solving than those assigned the traditional text-based homework. Furthermore, students in the Web-based group were better at conceptual problem solving than with numerical or algorithmic procedures, a contradiction to earlier findings of chemistry problem-solving development (Nurrenbern and Pickering, 1987; Sawrey, 1990; Nakhleh, 1993; Nakhleh and Mitchell, 1993).

An online course may be an opportune way to provide the individual review needed to facilitate the development of the stoichiometric competencies necessary for future chemistry coursework. For example, during the first semester of the introductory chemistry course at Carnegie Mellon University (CMU), students are expected to self-review stoichiometry content in preparation for a mastery examination. In order to help students with their preparation, the instructor posts problems to be learned and provides multiple testing opportunities. In the past, students have been found to need up to six tries on the mastery test to pass it; the majority of students fail the first attempt. Furthermore, many of the students who pass the mastery test still have difficulty intuiting stoichiometry's application during the second-semester study of equilibrium and acid-base chemistry. To aid students in their preparation, an online stoichiometry review course was designed and developed collaboratively by experts in chemistry content, educational psychology, instructional design, and multimedia technology in order to produce a product that optimally integrated chemistry knowledge with its methods and medium for Instruction.

This paper reports the results of an investigation conducted to determine the efficacy of this online course in promoting stoichiometry learning. The purpose of the investigation was to compare performance of students on a test of stoichiometry concepts and procedures after they had studied one of two cognitively informed sets of instructional materials. Post-test scores were analyzed to determine if dynamic expositions, immediate supportive feedback, and an overarching cover story facilitated through online technologies promoted greater learning outcomes than studying from text-based materials alone.

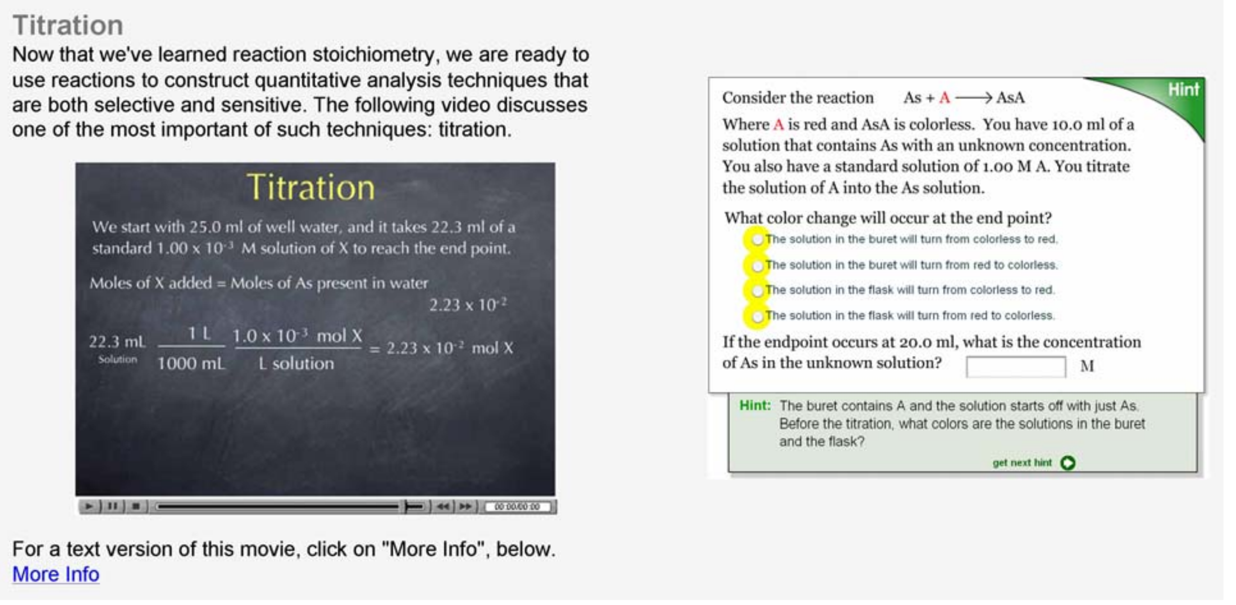

Fig. 1 Screenshot of Titration Module multimedia explanation with *More Info* link to text version (left) and practice exercise with hint (right).

Materials and methods

The online course

The online stoichiometry review course was developed with the support of the William and Flora Hewlett Foundation through the Open Learning Initiative (OLI) at CMU. The design principles included: a belief in the power of an overarching real-world story or context to both motivate students and integrate ideas; the use of an exploratory virtual laboratory in support of conceptualizing and practicing competencies; the need for a variety of practice contexts to support conceptual understanding; the importance of a variety of feedback experiences as students practice problems, from being able to track the effects of certain actions to getting responses to their submitted answers. In addition, the course works from a principle of explanation and example-based learning. The course can be accessed free of charge (http://www.cmu.edu/oli/courses/enter_chemistry.html) and is compatible with both PC and Mac OS10.1 platforms.

The course is divided into two units, each with multiple modules. A drop-down syllabus allows students to navigate to any module at will. The first unit develops the context of arsenic contamination of the drinking water in Bangladesh, an issue that suggests problems that get to the heart of stoichiometry. The contextual setting of groundwater contamination operates at the macroscopic level (*e.g.* grams, liters) from which interpretations can be made at the submicroscopic level (*e.g.* number of atoms, molecules, etc.) and then recorded using symbolic notation (e.g., AsO_2^-, g/mL, 0.05 M). By embedding stoichiometric knowledge in a real-world setting that highlights its utility, students learn and practice concepts in a context that may support their development of a coherent cognitive framework. The first unit also explains the use of the Virtual Lab (Yaron *et al.*, 2001) and other interactive features of the OLI course. Unlike a physical laboratory in which students can see *only* the macroscopic results (e.g., color change, gas production) of chemical interactions, the Virtual Lab also provides simultaneous quantitative representations of the abstract and invisible chemical species present. These responsive representations serve to link mathematical computations and actual chemical phenomena during problem solving, an action that may serve to promote development of a flexible knowledge base. Finally, the first unit reviews basic measurement skills and addresses basic compound and solution stoichiometry. This review is facilitated through the use of multimedia explanations of stoichiometric tasks and the procedures by which these tasks can be accomplished, as well as practice exercises with feedback and hints. The second unit develops the use of stoichiometric tools within the analytic (*e.g.* titration, percent yield, elemental assays) work that chemists do (Evans *et al.*, 2006). Each of the two units ends with a recap module for review of, reflection upon, and extension of, the concepts addressed. The course requires an estimated 20-25 hours to complete. During this investigation the course could be accessed only through a password protected website.

As an example of the content and context provided by the OLI course, Figures 1 and 2 illustrate by screenshots the multiple dynamic learning objects available to the student in the Titration Module located in the second unit of the course. It is this unit that develops the use of stoichiometric tools within the real-world activity of chemical reaction analysis. Titration is an example of a quantitative analysis technique that is explained in the context of efficiently, effectively, and inexpensively determining the amount of arsenic in the water supply. The left side of Figure 1 depicts a voiceover movie that includes a thorough explanation of the titration procedure as it used to accomplish a real-world task along with a link to a text version of the same content. The right side of Figure 1 shows one of the interactive questions that immediately follows the titration lesson. This type of question provides hints (from 3-6 per response) to guide a student through a calculation, with the last hint being a bottom-out hint that

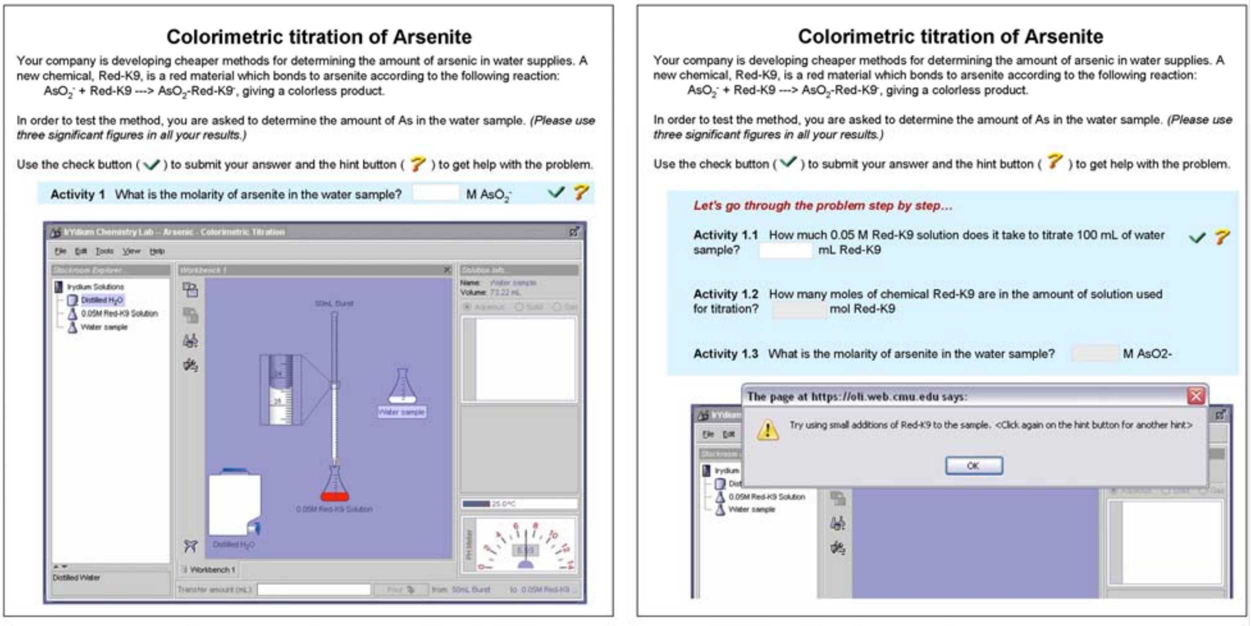

Fig. 2 Screenshot of Virtual Lab titration activity (left) with examples of requested hints and scaffolding (right).

provides the answer. Figure 2 includes screen shots of the parameterized tutor within the Virtual Lab. The student is given the opportunity first to solve the problem, for which hints and feedback that check for common errors are provided (left). Students may request the tutor mode, which assists them by providing sub-goals to be solved in a step-by-step fashion. Hints and feedback are available for each sub-goal if requested by the student (right).

The comparison course

Since the goal of the investigation was to ascertain whether the features of dynamic and interactive online technology along with an overarching real-world story would promote stoichiometry learning to a greater degree than a static text-only format, a text-based study guide consisting of sixteen lessons that mirrored the topics presented by the modules in the OLI course was developed, but without the dynamic interface, timely and informative feedback, or overarching storyline. Each lesson of this study guide included a brief explanation in the direct service of a specific problem type, a worked example problem with all moves explained as to purpose, a worked example problem with no explanation, and three practice problems for which no solutions (feedback) were available. The format of these materials was similar to that found in a textbook, except that in a textbook not all the topics addressed would be found as a cohesive unit. By developing text-only contrast materials, identical content would be accessed by studying from either format. If performance differences were found between the two formats, they could be attributed to the design principles and their execution (dynamics, feedback, context), not to the content. Because there are no video explanations or virtual lab exploratory opportunities, the text-only contrast materials are estimated to require 12-15 hours to complete (as opposed to the 20-25 hours for the OLI course). Nevertheless, this text-only study guide was an improvement over the traditional practice at CMU that consisted of posting problems to be learned and providing testing situations. The complete study guide was posted as a PDF file on a secure website.

The titration lesson in the text-only study guide, for example, addresses the same content as the Titration Module found in the OLI course and illustrated in Figures 1 and 2. First the student reads a brief description of the topic. Then a titration problem is presented in which both condition and action descriptions are shown. The most difficult problem from the OLI Titration Module was chosen as the example problem. This example problem is followed by another worked titration problem in which only the actions are shown. Finally, three practice problems are posed for the students to solve. Whenever possible, example and practice problems for the sixteen lessons in the text-only study guide were drawn from the OLI course.

The student population

Volunteers, who were at least 18 years of age and whose residences were identified as being in the United States (to avoid possible computer access problems), were solicited by email during the summer from the incoming (but not yet enrolled) CMU freshman class. Email solicitations were limited to students majoring in the sciences, engineering or mathematics since they comprise the vast majority of students who study introductory chemistry. A total of 426 students were contacted; 45 (27 males and 18 females) completed either the OLI course (21 students) or the text-only comparison course (24 students).

Experimental procedure

Communication with participants was through email and secure websites. Each respondent was directed to a website to complete an electronic background survey. The survey requested educational data regarding SAT (formerly known as the Scholastic Aptitude Test) scores and completed math and science courses, as well as general background information so that possible relationships of prior knowledge and/or demographics to stoichiometry learning could be analyzed. Upon receipt of completed surveys participants were randomly assigned to either the OLI course or the comparison text-only course. Depending upon the date of enrollment, participants had from 10 to 24 days to complete the study materials. During that time participants were free to choose how much they interacted with either the OLI or text-only study materials. Five days after the end date of the study, participants completed a proctored post-test of stoichiometry concepts and procedures on campus.

Data sources

Proctored post-test

A table of specifications was constructed relating content to desired learning objectives (see Appendix 2). Test items were developed from this blueprint in several ways. Whenever possible, standardized items from the American Chemical Society (ACS) or Advanced Placement (AP) Chemistry exams were used and converted into open-ended response items. For example, the following ACS item tests procedural knowledge about the qualitative analysis of pure substances (see Figure 4).

> A compound of sodium, sulfur, and oxygen contains 29.08% Na, 40.56% S, and 30.36% O. Which formula is correct?
>
> (A) Na_2SO_3, (B) Na_2SO_4, (C) $Na_2S_2O_3$, (D) $Na_2S_2O_8$, (E) $Na_2S_4O_6$

The first sentence of this item was used exactly as written, without the responses provided but with a request for students to show all work. If an item for a desired objective was not available from a standardized source, then a task for that objective was developed. For example, no items from either the ACS or AP Chemistry exams test the procedural knowledge needed for determining the molarity of a mixture of solutions (see Appendix 2) so the following item was created:

> The contents (A, B, C) of three different bottles of fructose solutions were combined in a 1000-mL volumetric flask. The flask was then filled to capacity with distilled water. Solution A was 74 mL of 0.527 M fructose, solution B was 632 mL of 0.872 M fructose, and solution C was 139 mL of 1.166 M fructose. What was the final concentration of the fructose in the volumetric flask? Please show all your work.

Scoring criteria for the post-test items were developed jointly by two researchers. Each item was allocated five points. In scoring procedural items, researchers deducted one point for a significant-figures or arithmetical/units error. Scoring for conceptual items was not standardized but was dependent upon the nature of the question. The first author scored all of the items for the post-tests. Five items from each of the post-tests (approximately 15% of the tests) were scored for reliability which was 100%. Post-test scores were converted into percentages.

Background information

Participants completed a background survey before being assigned to one of the two treatments. Since prior knowledge and gender have both been shown to affect chemistry performance (Ozsogomonyan and Loftus, 1979; Boli *et al.*, 1985; Spencer, 1996; Shibley *et al.*, 2003), information regarding participants' SAT scores, previous chemistry coursework, and gender was requested. These data were analyzed to determine if the backgrounds of the sample of participants were representative of the population from which they came, and if prior knowledge and/or gender contributed to learning outcomes.

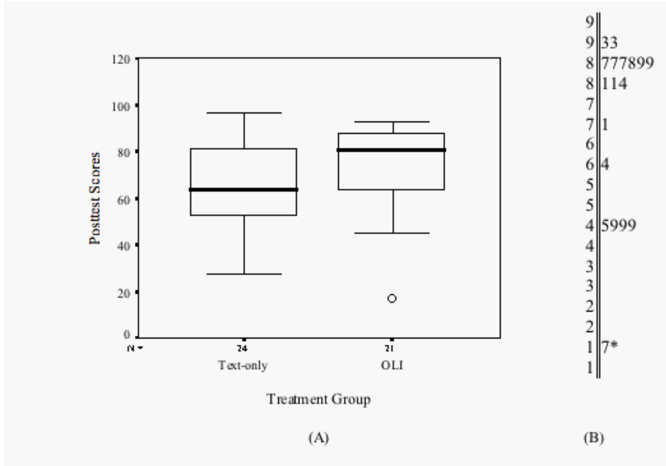

Fig. 3 (A) Box-and-whisker plots showing the distribution of post-test scores by treatment group. (B) Stem-and-leaf display of post-test scores from the OLI treatment group. The outlier score (17*) is nearly three standard deviations below the mean score (mean=74, *SD*=21).

Log files

The OLI's delivery system records (logs) every student move in the online stoichiometry review course. These log files can give an estimation of the time users spend with, as well as the frequency of their interaction with, particular features of the course such as the Virtual Lab. An analysis of these log files was completed for the OLI participants to ascertain whether there was any relationship to post-test performance of either their time spent with the course or their numbers of interactions with specific course features.

Results

Effect of treatment

Exploratory data analysis of the relationship of post-test scores to treatment suggested that there might be an advantage to studying with the OLI course (mean=74, SD=21) over the text-only materials (mean=65, SD=21). However, a single outlier score of 17 from the OLI treatment group is nearly three standard deviations below the group mean (Figure 3). In order to reduce the sensitivity of the data to the presence of this outlier, it was removed from subsequent analyses.

An independent samples t-test (t[42] = -1.989, p = .05) suggests that studying from the OLI course (mean=77, SD=16, without outlier) is indeed advantageous compared to studying from the text-only materials (mean=65, SD=21). When post-test scores are regressed on treatment, however, only six percent (adj. R^2 = .06, p = .05) of the variability in performance is explained by treatment. Since participation in

Table 1 Background attributes of treatment groups.

Background attributes	Treatment group: Text-only (n=24)	OLI (n=21)
AP completion	16 (67%)	9 (43%)
SAT scores (mean ± *SD*)	1369 ± 101	1389 ± 104
Number (proportion) of males	13 (54%)	14 (67%)

the technology-rich OLI course appears to explain so little of the variation in learning as measured by the post-test scores, a finer-grained analysis of the scores was undertaken. The results of additional analyses of performance and error types on specific post-test items (procedural, or conceptual, or both) favored studying with the OLI course but were statistically inconclusive.

Effect of background characteristics

The findings to this point suggest only a small advantage for the carefully designed OLI stoichiometry course. Certain background experiences and characteristics may be more closely related to stoichiometry post-test performance than participation in one or another form of a brief review course. Before investigating the role of specific background characteristics, it is appropriate to validate the random assignment process with regard to the distribution of participants' background attributes. Each treatment group was examined in terms of the participants' AP completion, SAT scores, and gender. Although there were small differences between the two groups in terms of mean SAT scores and in the percentage of males, with both favoring the OLI treatment group, there were no statistically significant differences between the two treatment groups with regard to any of these criteria. Table 1 summarizes the background composition of the two treatment groups.

Although the two treatment groups were statistically equivalent with regard to criteria that may be related to chemistry achievement, these student background experiences and characteristics themselves may be more closely related to stoichiometry post-test performance than participation in a brief review course. In order to address the question of the relationship between participant background experiences and characteristics and the learning of stoichiometry, the distribution of the post-test scores with regard to prior knowledge (AP chemistry, SAT scores) and gender was explored. Exploratory data analysis suggested that participation in an advanced high school course such as AP chemistry was not related to post-test performance. Previous analyses of the effect of AP on college performance have found that the level of student performance level on AP examinations (grade of 3 or higher) is strongly related to college performance but merely participating in AP or other honors-level courses in high school is not a valid predictor of superior performance in college (Geiser & Santelices, 2004). Therefore the lack of a relationship between AP chemistry experience and performance on the post-test is not surprising. Exploratory data analysis did suggest that SAT score and gender were related to post-test performance. Table 2 presents a correlation matrix of background variables and post-test scores. The significant positive correlations between SAT and post-test scores, gender and post-test scores, and gender and SAT scores suggest the need for additional analyses of these variables' influences on the learning of stoichiometry.

Table 2 Bivariate correlations for post-test scores and background characteristics

	Post-test Score	AP completion	SAT score
AP completion	.18		
SAT score	.51*	-.02	
Gender	.49*	-.18	.47*

*p=.01.

Table 3 Results of stepwise multiple regression of post-test scores on treatment, SAT score, gender, and all possible interaction variables.

Model	adj. R^2	*SE*	β	*p*
SAT	.25	17.1	.51	.001
SAT + Gender	.31	16.4	.36, .32	.02, .03

Effect of SAT

When post-test scores are regressed on SAT scores, β=.51 (p=.001) and 25% of the variability (adj. R^2) in the post-test scores is explained by the SAT scores. These results support the previous findings of a positive relationship between SAT scores and chemistry performance (Ozsogomonyan and Loftus, 1979; Spencer, 1996). It is important to note that the effect of prior knowledge as indicated by SAT performance explains more than four times the variability in post-test scores than does studying from either set of review materials (25% vs. 6%).

Effect of gender

When post-test scores are regressed on gender, β=.49 (p=.001), 22% of the variability in scores (adj. R^2) is explained by gender with males having the advantage. The concerns of documented gender differences in science (Grigg *et al.*, 2006) are borne out in this situation even though the female participants were enrolled in science and math departments of an upper level institution. Gender explains nearly four times more of the variability in post-test scores than does studying from either set of review materials (22% vs. 6%). Furthermore, the significant correlation of SAT score with gender (r=.47, p=.01) suggests an interaction of these variables and the possibility that the study's results may be confounded.

Effect of multiple factors

Treatment, SAT scores, gender, and all possible interaction variables (gender-SAT, gender-treatment, SAT-treatment, gender-SAT-treatment) were systematically added (stepwise) to the regression equation to determine a model that best explains the participants' post-test performance (see Table 3). Two models resulted, with no explanation of variability due to treatment or any interaction variable. SAT scores explain 25% of the variability in post-test scores in the first model. SAT scores and gender together explain 31% of the variability in post-test scores, with high scorers on the SAT and males having the advantage, in the second model.

Table 4 Comparisons of mean SAT scores of entering freshmen

	SAT Section Scores	
Group	Verbal Mean(*SD*)	Math Mean(*SD*)
Participants		
Females (*n*=18)	636(77)[1]	682(43)[2,3]
Males (*n*=27)	683(65)[1]	735(39)[2]
Non-participants		
Females (*n*=182)	653(76)	715(53)[3]
Males (*n*=401)	660(72)	740(48)
Overall		
Females (*n*=200)	651(76)	712(53)[4]
Males (*n*=428)	662(72)	739(47)[4]

[1, 2, 3, 4] Each pair of numerals indicates a significant difference of at least p=.05.

The gender effect—looking for an explanation

The similarity of SAT performance and gender individually to post-test scores prompted a deeper investigation of the relationship between SAT scores and gender in the sample of participants and the population from which they came. To ascertain whether male SAT scores differed from female SAT scores in the general population from which the sample in the present study originated, a comparison of SAT scores by gender was made. Table 4 is a summary of the findings. In each group, males outperform females on both the verbal and math sections of the SAT. The differences between the genders are significant for the overall population on the math section (p=.0001). It is unclear what this suggests; it could be that students are offered places with similar scores, but perhaps a higher percentage of highly-qualified males accept a place at CMU than do females, or that the combination of admission criteria tends to have females with higher high school GPA's and lower SAT scores or some other factor. Since freshman success, including science performance, has been linked to SAT performance, females may be at a disadvantage during the introductory chemistry or other quantitative science courses. The difference in SAT scores between males and females is even more pronounced within the sample of students who participated in the study. There is a .66 SD difference in SAT verbal scores and a 1.3 SD difference in SAT math scores. Furthermore, there are no significant differences in verbal or math scores between male participants and male non-participants but there are significant differences (.69 SD) in math scores between female participants and the female non-participants. Female participants' math scores are significantly lower (p=.01), a finding that may suggest an underlying difference in motivation between the males and females who agreed to participate in the study. Female participants may have felt underprepared for their chosen courses, and perhaps participated in this study (which occurred before classes began for the fall term) as a way to strengthen their skills.

Learning practices

Although background experiences and gender explain more of the variability in post-test scores than either treatment condition, the OLI course more effectively supports engagement with the study materials than does the text-only resource. Participants in the OLI group self-reported spending significantly (p=.01) more minutes working (mean=625; SD=297) with their study materials than the participants in the text-only treatment group (mean=381; SD=295). Log files generated by the actions of the OLI participants both supported their self-report times and enabled a closer examination of their specific study practices.

There is no significant correlation between time engaged with the materials and post-test performance (r=-.01, p=.98) for participants in the OLI course. Although log files accurately report when the OLI program is on, participants may not be engaged during that entire period. They may step away from the computer to answer the phone or grab a snack. In addition, study time itself may not be an optimal indicator of learning since some learners require more time on task than others to achieve the same level of understanding (Carroll, 1963; Bloom, 1974; Cooley and Leinhardt, 1980; Gettinger, 1984). Although the chemistry content of the text-only and OLI treatments was designed to be comparable, the instructional delivery of the OLI course also included multiple opportunities for interactive problem solving and exploration with feedback. Therefore an examination of the relationship of post-test performance to the level of engagement with these interactive opportunities within the OLI treatment group was undertaken.

The OLI course's Virtual Lab provides students with both support and feedback when solving stoichiometry problems. Visualizations of both submicroscopic and macroscopic interactions of chemical entities can aid students in their connection and integration of the multiple levels of chemical knowledge. In addition, hints for and feedback to proposed solutions may be requested by the users when solving problems in the Virtual Lab interface. Yet measuring time spent in Virtual Lab activities is subject to the same questions of engagement as arose from the measure of time spent overall in the course. An exploration of the participants' actual number of interactions with the Virtual Lab may be a more accurate measure of their engagement with the study materials. Each time a participant's mouse clicks in the Virtual Lab interface it is recorded as an event in that user's log file. A scatter plot of the distribution of the post-test scores and the total number of Virtual Lab events for each participant reveals a positive correlation (r=.43, p=.06); but the wide range (0-5000) of the number of Virtual Lab events suggests a scale issue (see Figure 4, left). When the number of Virtual Lab events is transformed to a logarithmic (base 10) scale, however, a strong positive and statistically significant correlation (r=.65, p=.02) is revealed between post-test score and the level of interaction (events) with the Virtual Lab (see Figure 4, right). When post-test scores are regressed on the $\log_{10}$ of the numbers of Virtual Lab events, 39% of the variability in performance (β=.65, p=.002) is explained by the level of participant interaction with Virtual Lab learning activities. These results suggest that the degree to which students take advantage of the instructional interaction afforded by the OLI course, not just being assigned to the technology-rich treatment, is highly related to learning as measured by post-test scores.

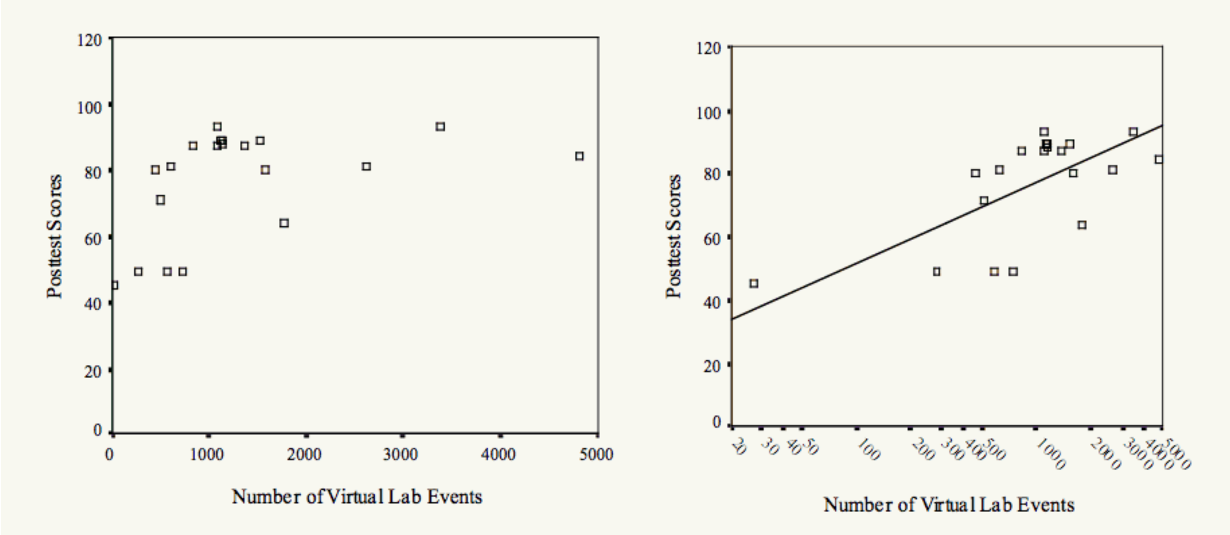

Fig. 4 Scatter plots of level of engagement with the Virtual Lab as measured by events and post-test scores for the OLI group. Transforming the scale of the number of Virtual Lab events reveals a strong positive relationship with post-test scores (right).

Stepwise regression of post-test scores on the $\log_{10}$ of the numbers of Virtual Lab events, gender, SAT scores, and all possible interaction variables yields two explanatory models for performance in the OLI treatment group. The first is that the greater the number of Virtual Lab events, the higher the post-test score; the second model includes a gender factor as well as the number of Virtual Lab events. What is important is that SAT, or aptitude, is not a factor. Both attribute a high proportion of the variability in performance to the level of engagement with the Virtual Lab. These models suggest that engagement with an interactive resource may overcome deficiencies in prior knowledge and gender differences (see Table 5), since SAT scores are not included in either model. Care should be taken in drawing conclusions from the second model (Log_{10} Virtual Lab events + gender). The small size of the OLI group (n=20) together with the low proportion of female participants (six out of twenty) calls into question the conclusiveness of this particular model with regard to gender. It should be noted, however, that the female participant having the greatest level of interaction with the Virtual Lab also scored highest (81) on the post-test among the female participants.

Table 5 Regression models for post-test performance in the OLI treatment group. Log_{10} of the number of Virtual Lab events, gender, SAT scores, and all possible interaction variables were added stepwise to the regression equation.

Model	adj. R^2	*SE*	β	*p*
Log_{10} Virtual Lab events	.39	12.8	.65	.002
Log_{10} Virtual Lab events + gender	.58	10.6	.49 + .47	.01+ .01

Discussion

The development of the OLI course was motivated by the desire to provide incoming freshman college students with a learning experience that would result in their fluid and flexible use of the stoichiometric competencies needed for the complex demands of solution chemistry problem solving. Although stoichiometry is addressed in most high school courses, college instructors have noticed that students do not have command of this central tool for chemistry work even if the content is reviewed early during a freshman chemistry course by direct instruction or self-study. To test the efficacy of this course's online features (dynamic expositions, immediate supportive feedback, overarching cover story) for promoting the learning of stoichiometry, a randomized investigation was conducted for the purpose of comparing post-test performance of those participants studying with the OLI materials to those studying from a text-only control set of materials. The results encouraged a cautious optimism with regard to the efficacy of the carefully designed online stoichiometry review course but also generated concerns with regard to the mastery of stoichiometric competencies overall and to the revelation of a significant gender gap in performance among science and engineering majors at an upper level university. Carefully designed modifications to the OLI's instructional locations and to the course's implementation may both enhance learners' overall stoichiometry mastery levels as well as work to minimize the gender gap in performance.

Modification of instructional locations

Although the performance of the OLI group (mean=76, SD=16) statistically significantly exceeded that of the text-only group (mean=65, SD=21), the mean post-test score overall was only 69%. These results suggest that there is still room for improvement in the instructional design. The online learning experiences could be enhanced by providing more examples and tasks (practice opportunities) for each topic, and/or by revising the example and task formats to encourage greater engagement of the participants.

Increasing the number of examples and tasks

When compared to four other online chemistry courses, the OLI course used as the technology-rich treatment in this study provided the highest degree of cognitive complexity among, but the fewest number of, examples and tasks for each stoichiometric topic (Evans and Leinhardt, 2008). Perhaps the developers of the OLI course felt that, since this course served as a review, the content was already familiar to the users and therefore the need for numerous examples and tasks was not as important as it would have been if the content were new. Yet the OLI course only provided an average of two examples and less than five tasks per topic. Considering that the quantity and variety of examples are important for students to be able to compare and to distinguish relevant from incidental features for specific problem types (Quilici and Mayer, 1996), the explication of only two examples may not have been sufficient for meaningful learning. Likewise, the practice opportunities provided by less than five tasks per topic may not be sufficient for development of any degree of fluidity or accuracy with stoichiometric procedures.

Revising the structure of examples and tasks

The relationship between conceptual understanding of, and procedural fluency with, stoichiometry competencies is not a simple one, due at least in part to the tripartite nature of chemistry knowledge. The limiting reagents and dilution conceptual items from the study's post-test required that participants work with only one (submicroscopic) of the three levels of chemistry knowledge whereas the corresponding limiting reagents and dilution procedural items required participants to integrate all three levels (macroscopic, submicroscopic, and symbolic). This act of integrating the three levels of chemistry knowledge is a conceptual task itself, albeit embedded within a stoichiometric procedural task. In an effort to aid participants in working with proportional reasoning across these three levels, the procedure of dimensional analysis was reviewed. Dimensional analysis supports proportional reasoning skills by showing how the units of measure are assigned and transformed during the arithmetic computation of ratios and proportions. But this numerical manipulation is simply a routine to be memorized rather than a way of reasoning through the multiple knowledge levels required by a stoichiometric task. Mechanistic learning of this type has been shown to block reflective competence on the part of the students, leaving them unable to learn from the problems they have done (Hiebert and Wearne, 1985; Hiebert, 1992). Stoichiometry is taught at the pre-college level as a set of tools divorced from use and through the procedure of dimensional analysis. As evidenced by the post-test results, simply re-teaching the dimensional analysis procedure through direct instruction was not effective in promoting stoichiometric procedural competence. Students need to be cognitively engaged with the solution process rationale in order for any possibility of transfer to new situations.

The work of Chi *et al.* (1989) demonstrated that students who self-explain worked examples learn more than those who tend to them in a more cursory manner. Since most students do not spontaneously provide effective self-explanations when studying worked examples, instruction needs to provide prompts for eliciting them (Renkl, 1997; Renkl *et al.*, 1998). The explanations should focus on the conceptual understanding of the tripartite nature of chemistry knowledge, not just the process of dimensional analysis. Practice with actual tasks would still be necessary to develop fluidity and accuracy. Therefore, after cognitively engaging students with their self-explanations of worked examples, support could be faded gradually from the use of incompletely worked examples to independent problem solving. Such a backward fading process (starting with last step of a worked example's solution) along with prompts for self-explanations have been shown to foster both near and far transfer performance (Renkl and Atkinson, 2003).

The potential of online interactivity

Although overall time spent with either treatment's materials was not related to learning outcomes, analysis of the log file data from the OLI participants revealed a clue for optimizing the use of allocated study time in the technology-rich OLI course. Nearly 40% of the variability in post-test scores from the OLI treatment group was related to the degree of participant interaction with the Virtual Lab. Furthermore, the relationship between SAT scores and post-test scores was eclipsed by participants' interaction with the Virtual Lab. These findings suggest an opportunity for mathematically less-advantaged chemistry students, in terms of their SAT scores, to build the critical quantitative and conceptual competencies needed for success in introductory chemistry. What remains unanswered by this study, however, is how these interactions with the Virtual Lab work to increase learning. Are increased interactions a sign of more practice with solving stoichiometric problems? Or do increased interactions indicate a deeper engagement within a problem, such as looking back and forth between a macroscopic flask on the Workbench and the submicroscopic entities and symbolic notations of the Solution Information Table? Or do the interactions reflect exploratory actions by learners as they generate and test self-developed hypotheses?

Although most students come to college with remarkable skills for accessing and downloading music from the Internet, few have had experience with using online technologies to optimize their learning in academic areas such as chemistry. To facilitate this optimization, there are two recommendations based on the findings from this study: (1) Students should be explicitly advised that interacting with the Virtual Lab might lead to increased learning—this could be done by including in the OLI course's instructions what this study has discovered. (2) All students should be encouraged to use the Virtual Lab—This could be done by setting default navigation options (*e.g.* continue buttons) within the course so that they lead to Virtual Lab practice. If users choose to bypass this default option, a pop-up reminder message of the benefits of interactive engagement with the Virtual Lab simulation should be displayed.

Final thoughts

The large expansion of tertiary education during the latter half of the past century has resulted in an increased need to support a diverse population of students. Many students need to be supported vigorously in order to succeed with their early quantitative coursework such as chemistry. The dynamic features of Internet technology can facilitate a pedagogical paradigm shift from the passive dissemination of content (e.g., through textbooks, videos, lectures) to the active support of these learners (e.g. through immediate informative feedback and exploratory environments) as they transform information into meaningful knowledge. Instructors may choose to implement online courseware along with–or even instead of–traditional textbooks as a way of fulfilling this need. But just as textbook review is based on criteria such as content and organization, online courses will need to be systematically evaluated to determine if the full potential of the individualized learning resources (e.g. interactive and exploratory environments) has been exploited. Nachmias and associates (Mioduser et al., 2000; Nachmias and Tuvi-Arad, 2001; Tuvi-Arad and Nachmias, 2001; Tuvi-Arad and Nachmias, 2003) have created a descriptive taxonomy of pedagogical and technological characteristics for online chemistry learning materials. What is missing, however, is an analytic tool that measures the cognitive quality of instructional materials—the features of course design that promote meaningful learning. An effective analytical framework to assess the quality of online chemistry instruction would ascertain whether a given course integrates the distinctive features of modern technology with instructional strategies informed by research on the constructivist nature of cognition. Such an analytical framework with which to compare the cognitive quality of online chemistry courses has recently been proposed (Evans and Leinhardt, 2008).

Acknowledgments

The authors want to acknowledge Michael Karabinos and Jordi Cuadros for their help in collecting the data for this study. This work was funded in part by the CCLI program of the National Science Foundation (NSF 0127455) and by the William and Flora Hewlett Foundation through the Open Learning Initiative Project.

References

Arasasingham R., Taagepera M., Potter F., Martorell I. and Longers S., (2005), Assessing the effect of Web-based learning tools on student understanding of stoichiometry using knowledge space theory, *J. Chem. Educ.,* **82**, 1251-1262.

Bjork R. and Richardson-Klavhen A., (1989), On the puzzling relationship between environment context and human memory, In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 313-344), Hillsdale, NJ: Erlbaum.

Bloom B., (1974), Time and learning, *Am. Psychol.,* **29**, 682-688.

Boli J., Allen M. and Payne A., (1985), High-ability women and men in undergraduate mathematics and chemistry courses, *Am. Educ. Res. J.,* **22**, 605-626.

Brown J., Collins A. and Duguid P., (1989), Situated cognition and the culture of learning, *Educ. Res.,* **18,** 32-42.

Carroll J., (1963), A model of school learning, *Teach. Coll. Rec.,* **64**, 723-733.

Chi M., Bassok M., Lewis M., Reimann P. and Glaser R., (1989), Self-explanations: how students study and use examples in learning to solve problems, *Cognitive Sci.,* **13**, 145-182.

Chinn C. and Malhotra B., (2002), Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks, *Sci. Educ.,* **86**, 175–218.

Cooley W. and Leinhardt G., (1980), The instructional dimensions study, *Educ. Eval. Policy An.,* **2**, 7-25.

Ericsson K. and Charness N., (1994), Expert performance: its structure and acquisition, *Am. Psychol.,* **49**, 725-747.

Evans K. and Leinhardt G., (2008), A cognitive framework for the analysis of online chemistry courses, *J. Sci. Educ. Technol.,* **17**, 100-120.

Evans K., Leinhardt G., Karabinos M. and Yaron D., (2006), Chemistry in the field and chemistry in the classroom: a cognitive disconnect? *J. Chem. Educ.,* **83,** 665-661.

Geiser S. and Santelices V., (2004), The role of Advanced Placement and honors courses in college admissions. *Center for Studies in Higher Education*, Paper CSHE-4-04, published online July 1, 2004, retrieved January 18, 2008, from http://repositories.cdlib.org/cshe/CSHE-4-04

Gettinger M., (1984), Achievement as a function of time spent in learning and time needed for learning, *Am. Educ. Res. J.,* **21**, 617-628.

Gick M. and Holyoak K., (1983), Schema induction and analogical transfer, *Cognitive Psychol.*, **15**, 1-38.

Grigg W., Lauko M. and Brockway D., (2006), *The nation's report card: Science 2005* (NCES 2006-466), U. S. Department of Education, National Center for Education Statistics, U.S. Government Printing Office, Washington, DC.

Herron J., (1975), Piaget for chemists, *J. Chem. Educ.,* **52**, 146-150.

Hiebert J., (1992), Mathematical, cognitive, and instructional analyses of decimal fractions, In G. Leinhardt, R. Putnam and R. Hattrup, (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 283-322), Hillsdale, NJ: Erlbaum.

Hiebert J. and Wearne D., (1985), A model of students' decimal computation procedures, *Cognition Instruct.,* **2**, 175-205.

Johnstone A. H., (1997), Chemistry teaching—science or alchemy? *J. Chem. Educ.,* **74**, 262-268.

Johnstone A. H., (2000), Teaching of chemistry-logical or psychological? *Chem. Educ. Res. Pract.,* **1**, 9-15.

Kousathana M. and Tsaparlis G., (2002), Students' errors in solving numerical chemical equilibrium problems, *Chem. Educ. Res. Pract.,* **1**, 5-17. (http://www.uoi.gr/cerp)

Lajoie S. and Azevedo R., (2006), Teaching and learning in technology-rich environments, In P Alexander and P Winne, (Eds.), *Handbook of Educational Psychology* (2nd ed.) (pp. 803-821), Mahwah, NJ: Erlbaum.

Miller G., (1956), The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychol. Rev.,* **63**, 81-97.

Mioduser D., Nachmias R., Lahav O. and Oren A., (2000), Web-based learning environments: current pedagogical and technological state, *J. Res. Comput. Educ.,* **33**, 55-76.

Nachmias R. and Tuvi-Arad I., (2001), Taxonomy of scientifically oriented educational websites, *J. Sci. Educ. Technol.,* **10**, 93-104.

Nakhleh M., (1993), Are our students conceptual thinkers or algorithmic problem solvers? *J. Chem. Educ.,* **70**, 52-55.

Nakhleh M. and Mitchell R., (1993), Concept learning versus problem solving: there is a difference, *J. Chem. Educ.,* **70**, 190-192.

Nurrenbern S. and Pickering M., (1987), Concept learning versus problem solving: is there adifference? *J. Chem. Educ.,* **64**, 508-510.

Osin L. and Lesgold A., (1996), A proposal for the reengineering of the educational system, *Rev. Educ. Res.,* **66**, 621-656.

Ozsogomonyan A. and Loftus D., (1979), Predictors of general chemistry grades, *J. Chem. Educ.,* **56**, 173-175.

Quilici J. and Mayer R., (1996), Role of examples in how students learn to categorize statistics word problems, *.J. Educ. Psych.,* **88**, 144-161.

Piaget J., (1954), *The construction of reality in the child (M. Cook, Trans.)*, New York: Basic Books.
Renkl A., (1997), Learning from worked-out examples: a study on individual differences, *Cognitive Sci.*, **21**, 1-29.
Renkl A. and Atkinson R., (2003), Structuring the transition from example study to problem solving in cognitive skill acquisition: a cognitive load perspective, *Educ. Psychol*, **38**, 15-22.
Renkl A., Stark R., Gruber H. and Mandl H., (1998), Learning from worked-out examples: the effects of example variability and elicited self-explanations, *Contemp. Educ. Psychol.,* **23**, 90-108.
Robinson W., (2003), Chemistry problem solving: symbol, macro, micro and process aspects, *J. Chem. Educ.,* **80**, 978-982.
Sawrey B., (1990), Concept learning versus problem solving: revisited, *J. Chem. Educ.,* **67**, 253-254.
Shibley I., Milakofsky L., Bender D. and Patterson, H., (2003), College chemistry and Piaget: an analysis of gender difference, cognitive abilities, and achievement measures seventeen years apart, *J. Chem. Educ.,* **80**, 569-573.
Simon H., (1974), How big is a chunk? *Science*,**183**, 482-488.
Simon H., (2000), Observations on the sciences of science learning, *J. Appl. Dev. Psychol.,* **21**, 115-121.
Spencer H., (1996), Mathematical SAT test scores and college chemistry grades, *J. Chem. Educ.,* **73**, 1150-1153.
Sumfleth E., (1988), Knowledge of terms and problem solving in chemistry, *Int. J. Sci. Educ.,* **10**, 45-60.
Sweller J., (1994), Cognitive load theory, learning difficulty, and instructional design, *Learn. Instr.,* **4**, 294-312.
Sweller J., van Merrienboer J. and Paas F., (1998), Cognitive architecture and instructional design, *Educ. Psych. Rev.,* **10**, 251-296.
Tsaparlis G., Kousathana M., and Niaz M., (1998), Molecular-equilibrium problems: manipulation of the logical structure and M-demand, and their effect on student performance, *Sci. Educ.*, **82**, 437-454.
Tulving E. and Thomson D., (1973), Encoding specificity and retrieval processes in episodic memory, *Psychol. Rev.*, **80**, 352-373.
Tuvi-Arad I. and Nachmias R., (2001), Current state of web sites in science education—focus on atomic structure, *Journal J. Sci. Educ. Technol.,* **10**, 293-303.
Tuvi-Arad I. and Nachmias R., (2003), A study of web-based learning environments focusing on atomic structure, *J. Comput. Math. Sci. Teach.,* **22**, 225-240.
Wheatley G., (1991), Constructivist perspectives on science and mathematics learning. *Sci. Educ.,* **75**, 9-21.
Wheeler A. and Kass H., (1977 March), *Proportional reasoning in introductory high school chemistry*, Paper presented at the 50th Annual Meeting of the National Association for Research in Science Teaching, Cincinnati, OH.
Whitehead A. N., (1929), *The aims of education*, New York: New American Library.
Yaron D., Freeland R., Lange D., Karabinos M., Milton J. and Belford R., (2001), Uses of a flexible virtual laboratory simulation in introductory chemistry courses, Presented as part of a CONFCHEM online conference, *Online Teaching Methods*, American Chemical Society, Washington, DC.
Yaron D., Leinhardt G. and Karabinos M., (2004), *Shape, transformation, and energy: critical resources for thinking in chemistry*, Unpublished manuscript, University of Pittsburgh, Pittsburgh, PA.

