

A Multi-year Evaluation of the Flipped Format in a General Chemistry Course

Colleen Taylor¹

Abstract

A flipped classroom model was implemented for three general chemistry courses between the years 2011 and 2013 and final grades were compared to those earned in the equivalent standard lecture course with the same instructor taught in the semesters of 2009, 2010 (two sections), and 2013. A comparison was also made using a control course, a traditional class with a different instructor with the same course materials as the flipped classroom, for the fall semester of 2012. Retention in the flipped classroom was 24% higher relative to the traditional format for the same instructor. Comparison of student grades in the flipped class with those in the control group having a different instructor provided compelling albeit less conclusive data. Survey data administered for the flipped and the control class of 2012 suggests that one of the major contributing factors to student success was the ability to affect the time students spend outside of the classroom.

Background

General Chemistry I is an entry-level course for science-based curricula which is often termed a gate-keeper course for science, technology, engineering, and math students. There is the common perception that the major barrier for student success in chemistry is the requirement for higher-order thinking skills in the cognitive domain such as modeling and graphing to be applied during the course (Apple, Beyerlein, Leise, & Baehr, 2007)—skills that are often lacking in students entering college and not nurtured in the “chilly climate” of the traditional lecture class (Daempfle, 2002). A report to the Department of Education in early 2000 suggested that the traditional class paradigm does not address students’ growth needs as learners and is the major contributing factor to student dissatisfaction and eventual withdrawal or failure in chemistry courses (Daempfle, 2002). A small but significant movement in the field of chemical education began about the same time and is still emerging providing transformative change in the classroom. Educators involved in this growing movement include early POGIL initiators (Farrell, Moog, & Spencer 1999), practitioners focusing on student metacognition (Cook, Kennedy, & McGuire, 2013) and entire countries dedicated to overhauling the way that the science is taught (Avargil, Herscovitz, & Dori, 2013). A tipping point has yet to be reached to change the culture of teaching and learning in higher education and there is much resistance to deep pedagogical change (Nelson, 2014).

Media tools have also aided in innovative pedagogical approaches. References as early as the 1970s show the use of film in tutorial guidance for problem solving (Garland, 1971) to more modern methods employing mixed media (Amaral, Shank, Shibley, & Shibley, 2013). There are now an immense number of resources

such as smart homework and free lecture materials from respected professors at major universities and through publisher content. Some improvement has been seen in the retention of students as a result of more and varied ways to deliver content and tools outside of the classroom (Chase, 2015). When the use of e-materials and methods began to merge with more traditional approaches, the term *blended learning* was coined (Arnaud, 2013). The so-called flipped classroom has grown from the blended learning approach with more emphasis on the instructor as a guide rather than a performer. The flipped classroom can be described as blended learning where the emphasis is on pre-class materials, video, or reading, replacing traditional classroom content delivery (Chase, 2015; Arnaud, 2013; Graham & Burke, 2014). Activities usually described as homework are transferred to the face-to-face setting and the learning is shifted to the learner.

Including a large active component in the classroom has its challenges, especially in the lack of training that the typical instructor has for the teaching profession itself rather than just the subject matter. Physical constraints exist such as large class sizes and characteristic stadium seating or individual desk classroom layouts that do not encourage collaboration. Even with knowledge that the lecture format is the least effective method of teaching, institutions of higher learning continue to promote and sustain this model out of convenience to handle larger and larger class sections and to appease students’ expectations of what a *teacher’s* job entails (Graham & Burke, 2014). The transition from accepted methods of instruction to active learning in a blended format is not an easy change and one that is often met with resistance from administrators, students, and other instructors hesitant to adopt new methods (Avargil, Herscovitz, & Dori, 2013).

¹Virginia State University

The changes implemented and described in this paper were designed to reach the students in the same way one may reach them during office-hour sessions—with direct intervention. The typical cycle followed by this instructor prior to changing the classroom format was similar to that of many chemistry instructors (Figure 1). The students were expected to prepare for class by reading the text and a lecture was delivered during class time with a mixture of PowerPoint®, videos, and chalk board/overhead work with specific problem-solving examples in the form of worksheets. If time allowed, some problems on the worksheet were completed in class, however, working the problems usually meant showing the students how to work the problems. The students were required to complete homework assignments related to those questions on the worksheets and in the notes using online homework associated with the text. Interim quizzes were given to assess the students' retention of this material in order to help them prepare for the exam. The reality was that students seldom read the book because all the information needed was delivered in the lecture or PowerPoint notes. Students copied the models presented and learned patterns to solving the problems without becoming problem solvers themselves, because no true self-assessment or problem solving was happening. The instructor intended for the quizzes to provide opportunities for self-assessment, however, quizzes were viewed by students as evaluations, or models to translate directly to the exam.

In order to affect how students approached learning directly, this instructor believed that a major change in classroom approach was necessary (Figure 2). Class preparation at home was to take the place of traditional in-class lectures by providing video and supplemental reading assignments. An expectation was set at the beginning of the course that worksheets were to be completed by students with help given as needed by the instructor who would

rotate between groups. Just-in-time lectures were provided only when the majority of the class needed further intervention to solve the problems for that day. In this new cycle, the need for students to perform during each class time was the major driving force for students to come to classes prepared. Integral to the changes implemented was an emphasis on the idea of self-assessment.

Student outcomes in the form of grades obtained by the students in the courses and on test materials based on their experience in a traditional or flipped format is described here. The intent is to add to the breadth of evidentiary approaches for those wishing to implement major pedagogical change.

Methodology

The Course and Students

The course under study at Virginia State University, VSU, is General Chemistry I, a three-hour introductory course in a two-semester series required of all science-track majors. These majors include, in order of usual population, biology (Biol), engineering (Engr), agriculture (Agri), math (Math), physical education (Hper), criminal justice (Cjus), and psychology (Psyc). The absence of chemistry majors in this science track class is due to the availability of a similar course offered for chemistry majors only; chemistry majors are therefore not included in the study. General Chemistry I has a listed pre-requisite of a math placement in college algebra and trigonometry or higher. This prerequisite was established in 2008 through the curriculum committee structure based on a 2007 internal study which equated success in the course with a minimal math placement score. The math department has undergone some reorganization and did not require the placement test in 2012. The chemistry faculty therefore substituted a short math

Figure 1 The traditional teaching cycle of in-class lecture and evaluation

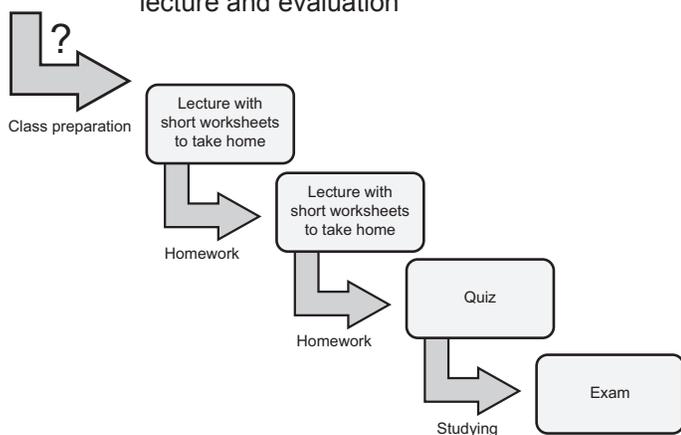
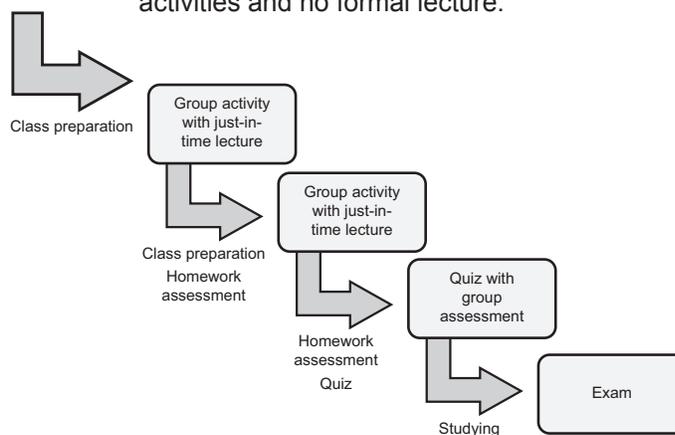


Figure 2 The revised cycle with emphasis on in-class activities and no formal lecture.



assessment given in the first few weeks of the chemistry course. Students were vetted for their readiness at the advisor level and not prevented from taking the course if they chose to ignore the recommendation of the instructor. The major classification/major population in the course is the sophomore biology student.

The course includes a one-hour recitation section which meets once per week in the evening. Recitation is with the same instructor and same students as the three-hour course. VSU does not have a graduate program in chemistry; therefore, there are no teaching assistants or graders. Recitation would ideally be used as the traditional problem-solving session with small groups. Instead, without graduate students to facilitate this offering, recitation is often used to administer and review quizzes and tests and to go over practice problems. Class sizes average around 50 students with two to three sections offered in the fall semester and one in the spring. The sections are loosely organized by a course coordinator, using the same text book, a similar syllabus, and schedule of course coverage. Otherwise, faculty members are allowed academic freedom in the classroom approach. Lecture halls of the stadium seating type or single level desk-type rooms are the most common type of classroom where these classes are scheduled. One collaborative room has been used in the past few years including the 2012 study. The lab portion of the course is a separate credit and grade. Students may, but do not generally, have the same instructor for lab and class.

Most students enrolled at VSU are in-state first-generation low income African American students who live on campus. The class profile in general chemistry was very much like that of the overall university.

The Instructors

The primary instructor and course coordinator was the instructor of focus for this study. The instructor has taught at VSU since 2003 and has taught General Chemistry I or Chemistry I for majors every fall semester during her tenure at VSU. She is well-versed in developing online courses and content, having experienced the Quality Matters Rubric (Hays, 2010) serving as the Blackboard CMS trainer and leader in helping develop an online presence at VSU. The instructor for the control group, Tongwen Wang, was in his second year at VSU and was comfortable with the lecture style PowerPoint-type classroom pedagogy and accepting of the use of materials prepared by the lead instructor.

The Course E-Content, Resources, Tools, and Platform

The course content and tools for the general chemistry course were developed using the Blackboard platform. Videos of course lectures were initially recorded over a period of years for students to use as a resource, so they were mostly developed prior to making the flipped classroom switch. Active learning worksheets were already in existence and were edited as needed to fit the format of the course. Prior to making the conversion, the developed online course materials were used to teach in the hybrid environment.

The course platform provided a student repository of PowerPoint® lecture notes, video tutorials, demonstration videos, online interactive homework, practice quizzes, and other publisher content. Students were allowed to use the group function in Blackboard to collaborate, and to use the journal function to reflect, but they generally did not utilize these two functions. Online interactive homework varied with the publisher but included the required e-text with *OWL* (Cengage), *Connect* (McGraw-Hill) or *Mastering Chemistry* (Pearson). In more than one year of the study, the text was provided in e-format free of charge.

Grading

Students in the traditional class and flipped classroom had the same grade percent distribution: 60% for four one-hour exams, 20% for the final exam, 10% for quizzes, and 10% for assignments, which included in-class and/or out-of-class work. To encourage participation, most courses provided up to 3% in bonus work for groups or other special projects such as keeping an electronic study journal. The laboratory component was graded separately.

Tests generally included 20 – 30 multiple-choice test bank questions with around three long-answer word problem/ application types. Some short answer or fill-in-the-blank type questions were included on occasion. Multiple-choice question types and levels on the exams varied from basic recall to application but were generally not numerical in nature unless they were one-step formula use. The final exam was standard for all sections with all multiple-choice questions. Assignments in the form of group work were graded as joint assignments for the students in the group in the flipped setting. The assignments were not always collected but a detailed key was provided. Students in the traditional class had some opportunities to work on problems in class but problem solving was usually instructor-led. Tests and the majority of quizzes were

administered in class. Some quizzes were moved online if time constraints became an issue.

The Video Platform

Videos in several formats were made available to the students. Those recorded by the instructor were from 30 to 50 minutes in length and mimicked what a normal lecture would look like. (1) Mediasite – instructor lecturing to a camera with streaming videos housed on a VSU server (2) MP4 format filmed during previous classes (3) Videos provided by the publisher. On occasion there were issues with the availability of the university server, so having downloadable or alternative choices became a necessity. Major issues with the availability of the server did occur during the 2012 study.

Groups

Students in the flipped classroom in 2011 were assigned teams of 4-5 individuals. These teams were balanced by GPA and major. The rigid conformity to the same team members was overly cumbersome and sometimes problematic due to personal factors. Later in 2012 and 2013 better interactions were found by allowing groups to self-assign and by providing flexibility in allowing groups members to shift, or by adjusting group members if the students were non-productive or not balanced in terms of student preparation.

Results and Discussion

Two separate analyses were conducted to view the differences between the flipped and traditional formats. The first compares the retention between the two approaches when the instructor is the same. The second compares the students’ exam scores between the different formats and instructors using math placement as a norming factor.

Semester & Year	Instructor	Enrollment	Format
Fall 09	Taylor	44	Traditional
Fall 10	Taylor	57	Traditional
Fall 10	Taylor	60	Traditional
Fall 11	Taylor	44	Flipped
Fall 12	Taylor	43	Flipped
Fall 13	Taylor	44	Flipped
Fall 13	Taylor	36	Traditional

Table 1 Courses used in the quantitative data analysis for the same instructor using comparison of final grades

Comparing Retention with the Same Instructor and Different Format

Final grades in the form of percent A, B, and C were compared for all traditional and flipped classrooms taught by Taylor from 2009 – 2013 (Table 1). Individual exam scores were not compared for reporting.

The average retention rate in the three flipped classes was 55% with a standard deviation of 3% whereas that for the traditional classes was 31% with a standard deviation of 7% (Figure 5). The improvement is 24% higher than the traditionally taught classes by the same instructor. Even looking at the lower end of the flipped classroom, 52%, with the higher end of the traditional, 38%, this is a dramatic improvement in one measure of outcomes, especially when measured over several classes over a five-year period.

Comparing Outcomes by Student Exam Grades with Different Formats and Instructors

In the 2012 comparison study, the flipped-class Blackboard course was copied over to the traditionally taught control class and the course syllabus was shared with only minor changes. All content was the same and the publisher content was used whenever possible. The quizzes and tests were the same, and efforts were made to keep this fact from the students. Only the objective multiple-choice questions were used in the analysis. The homework in the control class was paper-based and only collected from some students randomly to encourage completion and to provide a reasonable work load for the instructor. All students had access to the course text and online homework without charge. The paper text was made available in addition to the free e-text. The control class fell behind

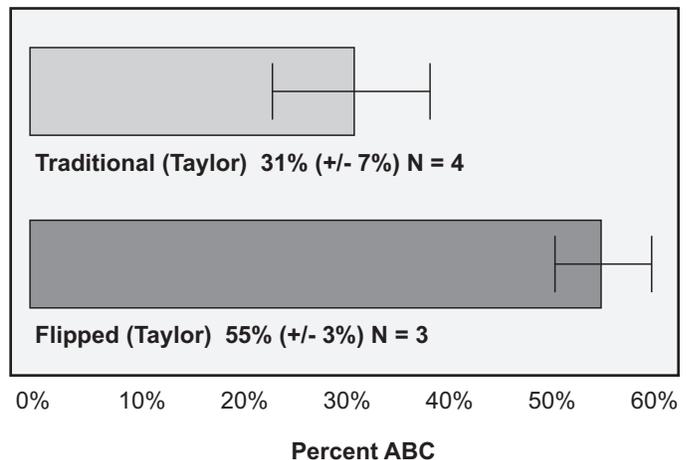


Figure 5 Retention of students (%ABC) in traditional versus flipped courses with the same instructor over a five-year period.

Figure 3 Comparison of Classification.

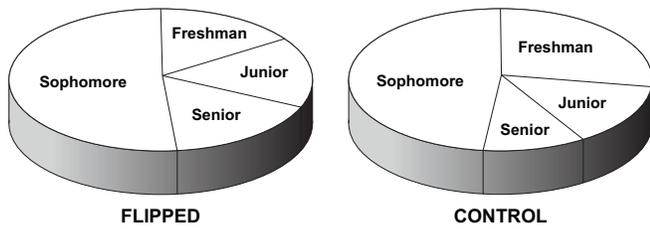
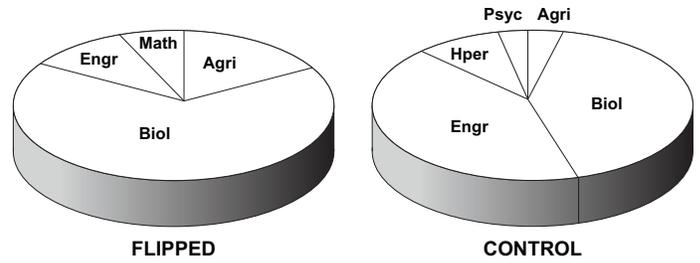


Figure 4 Comparison of Major. A time conflict affected engineering student enrollment in the flipped class.



by one week and we were unable to keep the students from obtaining information from the other section at that point. Therefore, only the first three tests were used to compare outcomes in the course. Thus, Taylor and Wang’s sections for fall 2012 were compared quantitatively by looking at the performance on exams 1, 2 and 3 and comparing average student scores on these three exams against their expected aptitude based on their math placement scores. The breakdown of the two courses is depicted in Table 2. The size of the course is comparable as is the class year of the students (Figure 4). The major breakdown of the students is different between the two classes as shown in Figure 3.

Table 2 Courses used in the quantitative and qualitative data analysis for different instructors using average test grades 1-3 and survey data

Semester & Year	Instructor	Enrollment	Format
Fall 12	Want	47	Traditional
Fall 12	Taylor	43	Flipped

The success of the students was normalized between the two classes by using a math assessment to categorize the initial level of the students as *low*, scoring from 1-4 points, *medium*, scoring 5 to 9 points, and *high*, scoring 10 to 16 points on 16 questions. The assessment questions tested knowledge of scientific unit analysis, exponents, fractions, algebra, and scientific notation. There was previous precedent for using the math placement as a reliable

predictor of expected performance in the course based on an internal study conducted in the fall of 2007 taken on behalf of the curriculum committee of the university to add a math placement minimum as a prerequisite for the course. The sample size of 31 for the traditional and 29 for the flipped courses was limited by those who took the assessment test, completed at least two of the first three tests, and signed a release form.

The math assessment administered in the first two weeks of the course was used to normalize the data across the two sections: the flipped class and the traditional class with different instructors. This approach was taken for two reasons: (1) the students in each class varied greatly in their classifications and majors, and (2) a previous study used internally showed that math assessment data was a reliable predictor of success in general chemistry. The data were scattered and hard to interpret when using assessment data versus the individual multiple choice average exam scores, so the assessment scores were grouped by low, medium, and high (Table 3). It is important to note that the low scorers were sent to their advisors to drop the course but chose to stay enrolled. With different instructors, if error bars are included, there is no statistical difference in the data. However, there is an increase in chemistry test scores with an increase in math placement in the flipped course that is absent in the traditional course. These differences were consistent regardless of the perimeters examined, such as using only sophomore or biology scores or when only plotting each individual exam average Scantron score.

Table 3 Quantitative data for comparisons of exam scores for a traditional and flipped classroom in the fall of 2012

Grouping	Traditional				Flipped			
	Average Math Score	(+/-)	Scantron Score Exam 1-3 Average (%)	(+/-)	Average Math Score	(+/-)	Scantron Score Exam 1-3 Average (%)	(+/-)
Low	2.6	0.5	59.2	14	3.1	1.1	44.9	15
Medium	6.6	1.3	56.9	17	6.3	1.3	50.4	14
High	11.1	1.8	56.9	14	11.6	2.6	77.1	14

The End-of-Course Survey

A survey was launched for the flipped classes in 2012 and 2013, with several questions probing a variety of student impressions of the course, with 34 and 33 valid respondents respectively. Similar questions were framed to try to obtain parallel information where applicable from the control class in 2012 with 37 valid respondents. The surveys were anonymous, available to all enrolled and limited to one attempt per IP address. Extra credit was provided for completion by having the students provide a screen shot of the completion page.

One of the contributing factors to the improved outcomes for the flipped versus traditional courses with the same instructor over five years may have come from the amount of time students spent on work outside the classroom. The students in the flipped classroom were more likely to have spent the university-suggested six to eight hours on their chemistry work relative to the students in the traditional class. The students in the traditional class spent less time outside the classroom but were more likely to say that they *should* have spent more time outside the classroom, although the questions were worded slightly differently.

- Flipped: “How much time did you spend on work outside of class including reading the book, watching the lectures, and on OWL?” and “How much time did you think should be spent outside of class including reading the book, watching the lectures, and on OWL to master the course material?”
- Control: “How much time did you spend on work outside of class including studying, reading the book, completing homework, and organizing your notes?” and “How much time did you think should be spent outside of class to master the course material in a science course?”

A second contributing factor to the success was the overwhelmingly positive attitudes about the group work in the course. 76% of 37 respondents in 2012 reported positive attitudes, with 22% neutral and 3% negative. 88% of 33 respondents in 2013 reported positive attitudes, with 9% neutral and 3% negative.

Conclusion and Insights

As presented here, over a number of years the success of the flipped classroom is convincing evidence of improvement in student outcomes due to the pedagogical change of the course, judging from the 24% increase in the students' final grades when comparing only one instructor's results. A similar percent improvement was reported by Paulson in an organic chemistry course (1999)

and more recently in another general chemistry course in Southern Maine (Ford, 2014). The result in both cases can be attributed to transferring the active component of the class to the face-to-face environment using a blended learning flipped approach. Having a successful experience inside the classroom requires that the students complete the preparatory work and that they have a positive experience with their peers. In the case of this instructor's experiences, the students' self-reported time spent outside the classroom and perceptions on peer interactions are contributions to the improvement in final grades. A great deal of time was spent when switching to the method, especially in dealing with group dynamics. Training in effective facilitation rather than on-the-job training may be a useful intervention in promoting widespread adoption of the flipped classroom method.

Comparison of the flipped course to the control group in terms of outcomes of the students with different instructors highlights the difficulty in trying to normalize data with so many different parameters to control (Lewis, 2014). The only significant outcome is the better performance of the more prepared students in the flipped class relative to those in the traditional course. This result is not surprising, as with a peer-to-peer led exercise the prepared student often explains to others and thus learns better (Lewis 2011). The preference would be to expand the data set over a number of instructors which can prove daunting in a small school but perhaps possible in a collaborative study (Zare, 2008). The premise of using a math assessment in normalizing the class is brought into question when the control group shows almost no advantage to the more prepared student for the standard format. Upon reflection, the use of only multiple-choice questions of the non-numerical multistep type could explain the flat trend in the data. A more effective method would be to select performance on key questions sprinkled throughout the course, perhaps with organization by learning level (Apple, Beyerlein, Leise, & Baehr, 2007). Using this approach may also cut down on students in one section giving away the test contents to another. It would be useful to follow those students through to their second-half of the course experience to see the effect of the flipped experience on matriculation through their higher-level chemistry courses. In this case the data set is further reduced given that VSU's engineers are not required to take the second half of the general chemistry course. The limited size of the data set would make such a study difficult and again would require several more years or a collaborative project between schools to complete. Interestingly, there is no evidence that this researcher could find that support a longitudinal improvement as a result of the student's flipped class experience and Lewis notes caution in this area (Lewis, 2014). Networks of educators completing and measuring similar and long-term improvement in student

success as resulting from flipped classroom experiences should provide support and legitimacy and arm the educator to make effective changes in the larger community towards large-scale reform in education (Avargil, Herscovitz, & Dori, 2013; Zare, 2008).

References

- Amaral, K. E., Shank, J. D., Shibley, I. A. Jr., & Shibley, L. R. (2013). Web-enhanced general chemistry increases student completion rates, success, and satisfaction. *Journal of Chemical Education*, 90(3) 296.
- Apple, D. K., Beyerlein, S. W., Leise, C., & Baehr, M. (2007) Classification of learning skills. In D. K. Apple, S. W. Beyerlein, & C. Holmes (Eds.). *Faculty guidebook: A comprehensive tool for improving faculty performance* (4th ed.) (p. 201). Lisle, IL: Pacific Crest.
- Arnaud, C. H. (2013). Flipping chemistry classrooms. *Chemical and Engineering News*, 91(12) 41.
- Avargil, S., Herscovitz, O., & Dori, Y. J. (2013). Challenges in the transition to large-scale reform in chemical education. *Thinking Skills and Creativity*, 10 189. Web. 12 Feb. 2015.
- Chase, C. (2015). Blended learning: combining online technology with classroom instruction: 1 of 3. *Make EdTech Happen*. Resources for Practically Implementing Tech Trends in Schools, Mar. 15. Web. 12 Feb. 2015. <http://chip-chase.com/2012/03/21/blended-learning-combining-online-technology-with-classroom-instruction-1-of-3/>
- Cook, E., Kennedy, E., & McGuire, S. Y. (2013). Effect of teaching metacognitive learning strategies on performance in general chemistry courses. *Journal of Chemical Education*, 90(8) 961.
- Daempfle, P. A. (2002). *An analysis of the high attrition rates among first-year college science, math and engineering majors*. Report to the Department of Education. ERIC. Web. 11 Feb. 2015
- Farrell, J. J., Moog, R. S., & Spencer, J. N. (1999) A guided-inquiry General Chemistry. *Journal of Chemical Education*, 76(4) 570.
- Ford, J. R. (2014). Flipping crazy: The large lecture flipped classroom model at the University of Southern Maine: *Addressing the Millennial Student in Undergraduate Chemistry*. By Lucille A. Benedict. Vol. 1180. Washington DC: American Chemical Society, 2014. 59-70. Print. ACS Symposium Ser.
- Garland, J. (1971). A TV based guide for freshman chemistry. *Journal of Chemical Education*, 48(3), 207.
- Graham, K., & Burke, K. (2014). Perception on a blended and flipped classroom. *International Journal of Process Education*, 6(1), 21.
- Hays, M. (2010). SLOAN-C: Quality matters: An overview. *EdTech Connect*. N.p., 23 Apr. Web. 11 Feb. 2015.
- Lewis, J. E. (2014). Doing chemistry education research in the real world: Challenges of multi-classroom collaborations. In D. M. Bunce, & R. S. Cole (Eds.). *Tools of chemistry education research*. Vol. 1166. Washington DC: American Chemical Society, 267-78.
- Lewis, S. E. (2011). Retention and reform: An evaluation of peer-led team learning. *Journal of Chemical Education*, 88(6), 703-7.
- Lewis, S. E. (2014). Investigating the longitudinal impact of a successful reform in general chemistry on student enrollment and academic performance. *Journal of Chemical Education*, 91(12), 2037-44.
- Nelson, C. (2014). *Overcoming personal incredulity: Reducing our resistance to deep pedagogical change*. Process Education Conference 2014: Plenary: Fri June 20.
- Paulson, D. R. (1999). Active learning and cooperative learning in the organic chemistry lecture class. *Journal of Chemical Education*, 76(8), 1136.
- Zare, R. N. (2008). Questions to chemical educators from the chemistry community. In D. M. Bunce, & R. S. Cole (Eds.), *Nuts and bolts of chemical education research*. Vol. 976. (11-18). Washington DC: American Chemical Society. ACS Symposium Ser.

