Spring 2011

Retention of Science Majors Through Different Avenues of General Chemistry Education

Benjamin P. Bilodeau
Longwood University

Melissa C. Rhoten
Longwood University

Retention of Science Majors Through Different Avenues of General Chemistry Education

Senior Honors Thesis

Spring 2011

Benjamin P. Bilodeau & Melissa C. Rhoten

Department of Chemistry & Physics
Longwood University
Farmville, Virginia 23909

Table of Contents

Abstract

For the past two academic years, the chemistry division at Longwood University has implemented POGIL methodology in some sections of the general chemistry sequence for science majors. Initially we were interested in determining if students exposed to POGIL were more successful in general chemistry than those students that were not exposed to this teaching pedagogy. Subsequent studies compare conceptual learning outcomes of POGIL versus non-POGIL students and also try to determine if SAT and GALT scores are a good predictors of student success in general chemistry. This project also addresses the following question: do science majors exposed to POGIL pedagogy in first semester general chemistry demonstrate increased success in subsequent chemistry courses compared to those students with no POGIL experience? Details about POGIL implementation and challenges as well as the methods of assessment used are discussed.

Introduction

Research on learning theory tells us that learning is a personal activity. Educators generally

agree that knowledge cannot be transmitted intact from the brain of the instructor to the brain of the

student (Bodner, 1986). However, the teaching model used historically in numerous disciplines,

including chemistry, has been one where students are lectured to in a passive environment. Proponents

of Constructivist theory assert that the assimilation of knowledge begins by the senses processing input

from the environment. This sensory data then interacts with existing knowledge in order for new

knowledge to be constructed by an individual. The process depends on a number of personal factors

including biases, misconceptions, beliefs, prejudices, likes, and dislikes of the learner (Butts, 1977).

Many new teaching approaches are learner-centered rather than instructor- or content-centered to

match the constructivist model of learning (Caprio, 1999; Johnstone, 1997).

Striving to understand education and its most effective methods is nothing new. Jean Piaget, a

Swiss developmental psychologist, worked throughout the twentieth century to understand how

children learn. He is credited with works like the theory of cognitive development, which assigns phases

of understanding in youths (Bunce, 2001). Although this model is that of a psychologist, not an

educator, there are many useful elements to Piaget's work that can be used to benefit students in

chemistry and other fields.

Piaget's first stage is known as the sensorimotor stage, from birth to age 2. In it, children

experience the world through movement and their five senses. At this age, children are egocentric; they

have an inability to perceive the world from others' points of view. Infants move through several sub-

stages in this phase, working on their simple reflexes, habits, cyclic actions, novelty, and familiarity

(Bunce, 2001). Stage 2 is the preoperational stage, representing ages 2-7. In this stage, a child acquires

motor skills and 'magical' thinking takes over as egocentrism fades. Logical thinking and reasoning skills

evade this age group. Next is the concrete operational stage from ages 7-12. Here, logical thinking begins but only in a concrete sense. Abstract thoughts and concepts are still not used in reasoning or even retained. In this stage, practical aids in thinking processes may still be used. Egocentrism has faded completely by this stage (Bunce, 2001). The final stage of child development according to Piaget is the formal operational stage, which spans from age 12 onward. Abstract reasoning is developed, and the retention of these logical and abstract elements allows these children to think in their own minds. As we will discuss later, it is believed that some incoming college freshman are not yet functional in this stage, leading to some of their academic issues (Bunce, 2001).

Piaget monitored children of all ages to construct a generalized timeline of human cognitive development through adulthood. Beginning at about age 12, he suggests that children are fully developed but are able to reason through experiences and improve their understanding independently. Although Piaget's work was not aimed at finding the best teaching method, he submits that children after this age learn well if encouraged to interact with the environment around them. After this work started to circulate in the 1930s, educators everywhere have strived for the ideal curriculum, including chemists (Bunce, 2001).

It is not known how students are best able to learn, so adjustments must be made in the classroom to accommodate multiple learning styles. What are the best adjustments? Only conveying simple concepts that are easy to understand leaves possibly important topics unexplored. Offering complex topics in a general, simplified manner may drive students to algorithmic methods of solving. Lecturing simply gives students a sampling of the teacher's understanding of the material, leaving them no choice but to memorize if they do not comprehend the material. As students continue to do poorly, withdraw from, or retake classes, instructors may be inclined to simplify lessons rather than to strive for greater comprehension in the class.

Perhaps before the best adjustments in teaching style are found, and before those adjustments are made, an understanding of student preparation is necessary. According to Herron on Piaget, over 50% of students coming into college as freshman directly out of high school are not functioning on a formal level. Whether they are chemistry or music majors, about half of students surveyed were not yet on this "formal" level of understanding that Piaget sets forth as his final step of development (Herron, 1975). Could the intellectual age of incoming college freshman really be less than 12 years old? The study concludes that, speaking of college students as a whole; about 25% of students function in a fully-formal capacity and are able to reason through problems based on a conceptual understanding of material. Here it becomes evident that Piagetian theory has some inherent flaws. According to Herron, not only does Piaget underestimate the abilities of some young children, but he also overestimates those in the final stage of development.

Central to this idea of education reform are the misconceptions held by educators past and present. These can be summarized by (1) student achievement is measured by increased content knowledge in the field and (2) students ought to be able to learn from traditional, rich sources like textbooks, lectures, and presentations alone. More recently, however, curricula have been less based around the principles of the field they intend to serve and more around how students learn (Pienta, 2005). From this stems what educators would have loved to know one hundred years ago; that education is not the regurgitation of facts, but a means to teach students how to think independently. Teachers should not be striving solely to convey the information associated with success in their field, but also should be working to improve student intellect.

Education reform has been at the forefront of state legislation recently, as Governor Bob McDonnell pushes his Commission on Higher Education Reform, Innovation, and Investment. In this proposal, over the next 15 years Virginians can expect one hundred thousand additional undergraduate degrees (bachelor or associate's) to be awarded through increased enrollment, improved graduation

and retention rates, and student assistance with college credit to complete degrees through public and

private institutions in Virginia. In this proposal, additional focuses are placed on attainment of degrees

in high-demand, high-income fields that will be instrumental to the 21st century economy. These include

degrees in the STEM (science, technology, education, and math) and healthcare fields. The governor has

laid out a plan to make available a not-for-profit partnership with educators and professional

communities to bring students into these fields. Tuition assistance and incentives for students who

show an interest in early hands-on collaboration and commercially-viable research are also part of his

proposal. In order to meet the demand of increased degrees in these areas, we must strive to attract,

retain, and produce highly-qualified students (McDonnell, 2010).

Figure 1 shows the grade distributions for recent offerings of CHEM 111 (Fundamentals of

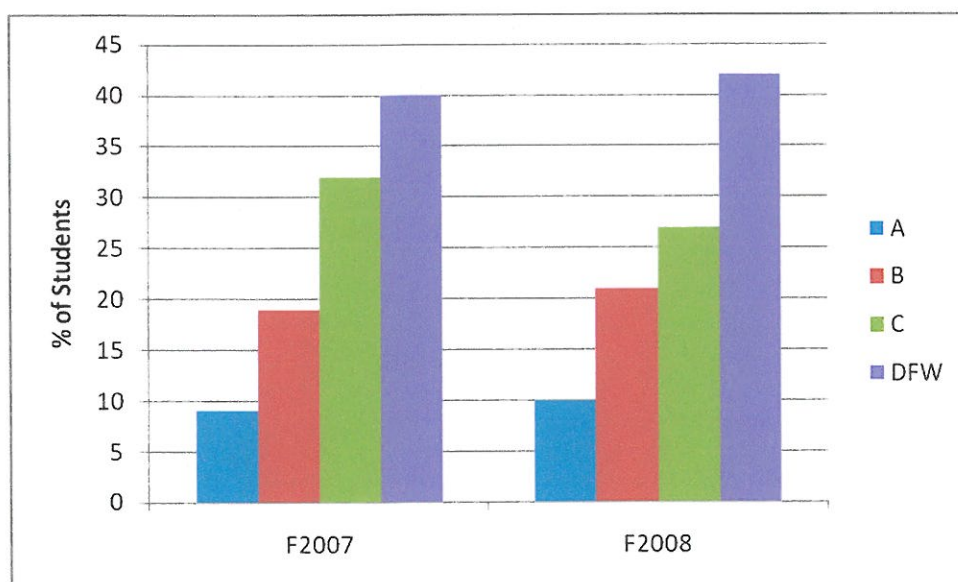Chemistry I), a required course that all Longwood freshmen science majors must successful complete.

Figure 1 – Grade distribution for CHEM 111 for Fall 2007 and Fall 2008 (Instructor=Melissa Rhoten)

For each of these cohorts, 40% of the students earned grades of D, F, or withdrew from the course.

How will we accomplish the task of increasing the number of STEM degrees awarded if 40% of the

students cannot successfully pass their introductory chemistry course? Perhaps the manner in which the material is presented in the course needs to be investigated.

In the mid-nineties, a new teaching pedagogy called POGIL (Process-Oriented Guided Inquiry Learning) emerged in chemistry. This approach is based on research indicating that a) teaching by telling does not work for most students, b) students who are part of an interactive community are more likely to be successful, and c) students enjoy themselves more and develop greater ownership over the material when they are given an opportunity to construct their own understanding (POGIL.org, 2010). A POGIL classroom or lab consists of students working in small groups on specially-designed guided inquiry materials. These materials supply students with data or information followed by leading questions designed to guide them towards the formulation of their own valid conclusions. The instructor and student assistants serve as facilitators, observing and periodically addressing individual and classroom-wide needs.

For the 2009-2010 academic year, the chemistry division at Longwood University implemented the POGIL methodology for teaching some of its general chemistry students. In Fall 2009 there were six sections of CHEM 111, the introductory chemistry course for all science majors. Dr. Melissa Rhoten fully utilized the POGIL approach in her two sections as she has good background in its strengths and limitations. I served as a student assistant in this section, which required me to come to all class meetings (TR 9:30-10:45 AM). Dr. Rhoten often gave a brief background lecture on the topic of the day (5 minutes or less) prior to dispatching the students into their groups for further investigation. Dr. Rhoten, Catherine Swandby (the other student assistant), and I move around the room listening to the groups discuss the topic. We also answered questions from the groups and corrected any misconceptions that we deemed necessary. Dr. Sarah Porter used a hybrid-POGIL system in her two sections of CHEM 111, where she gave traditional lectures two days per week and used POGIL activities once per week. Dr. Keith Rider used a fully traditional lecture setting (e.g., no POGIL group work). This

is where my interest in this project peaked, and Dr. Rhoten and I planned to further investigate this new system.

These types of teaching methods are not exclusive to chemistry departments. During his time at Longwood University, Dr. James Moore, a physicist, also used an inquiry-based learning system. He achieved impressive results across many levels of physics education. Moore used a method pioneered by Eric Mazur at Harvard University. The Peer Instruction techniques differ depending on the level of the course, but in general students spend time working in groups on guided-inquiry activities designed to confront common misconceptions and to correct them. The literature suggests that this is far more effective as compared to passive lecture or questioning without discussion. To validate his use of this technique, Dr. Moore used various national Concept Inventory Tests to assess student learning gains. For example, in PHYS 202 Dr. Moore used the Conceptual Survey of Electricity and Magnetism as a means of assessing the Peer Instruction techniques used in the course. A normalized gain of 35% was obtained, which is significantly above the national average for traditionally taught courses and statistically equivalent to gains achieved at institutions adopting similar pedagogies. This shows that this methodology can be easily adapted to fit many lesson plans at various levels of science education.

Colleagues at UNC-Asheville found increased graduation rates among Science, Technology, Engineering, and Mathematics (STEM) majors that had POGIL in 1st semester general chemistry relative to those in traditional general chemistry courses. Does a POGIL experience in 1st semester general chemistry affect retention of science majors at Longwood University? The overall goal of my project was to initiate this study with the cohort of students in the 2009-2010 general chemistry sequence. Initially, I planned to track the success of this cohort in subsequent chemistry classes and to assess whether there is a correlation between a single POGIL experience and increased graduation rates of those science majors with this unique experience. However, it became apparent very quickly that the

scope of the project was too large and that other more immediate questions should also be addressed

(*vide infra*).

Methods

The Fall 2009 cohort was initially comprised of 136 students. Forty-eight of these students were enrolled in the traditional (lecture only) course, whereas the remaining 88 were enrolled in POGIL sections (either all POGIL or hybrid POGIL, *vide infra*). For the traditional sections, 78% of the students were biology majors, 12% physics majors, and 10% other non-science majors (i.e., UNDC, KINS, ATTR). For the POGIL sections, 66% of the students were biology majors, 16% chemistry majors, 11% physics majors and 7% other non-science majors (i.e., UNDC, KINS, ATTR). Students selected their sections randomly at registration with no knowledge of the teaching style to be used.

The Fall 2010 cohort was initially comprised of 149 students. Seventy-seven of these students were enrolled in the traditional (lecture only) course, whereas the remaining 72 were enrolled in fully POGIL sections. For the traditional sections, 80% of the students were biology majors, 6% chemistry majors, 4% physics majors, and 9% other non-science majors (i.e., UNDC, KINS, ATTR). For the POGIL sections, 68% of the students were biology majors, 6% chemistry majors, and 15% other non-science majors (i.e., UNDC, KINS, ATTR).

Student scores on common in-class exam questions as well as student scores on the First Semester General Chemistry American Chemical Society standardized exam (Fall 2009 cohort) or the First Term General Chemistry Paired Questions American Chemical Society standardized exam (Fall 2010 cohort) were collected. For the Fall 2010 cohort additional metrics such as GALT score, pre- and post-chemical concept inventory test data, SAT math and verbal data were collected as well.

The GALT (Group Assessment of Logical Thinking) is a twelve-item test developed by Roadrangka et al. (1982) to measure logical reasoning skills in both college and pre-college level students. At Longwood, 10 of the 12 items were administered via Blackboard during one of the first laboratory sessions of the semester. The remaining two items were administered in pencil-and-paper format

during the same laboratory session. Students were given approximately 30 minutes to complete the test. The instrument was scored according to the rubric provided with the test (Roadrangka, 1982).

Students in the Fall 2010 cohort were also given a chemical concept inventory test at the start and end of the semester to determine the "value added" by taking to course. This exam was developed in the mid-1990s as a master's thesis by Douglas R. Mulford of Purdue Universtiy. This 22-question exam covered the majority of the concepts presented in CHEM 111. Students were given this test during an initial and final laboratory period and were given approximately 45 minutes to complete it in each case.

Data on graduation rates of science majors from 2005 to 2008 at Longwood University were obtained through collaboration with Office of Assessment & Institutional Research. During this time, CHEM 111 was only taught using the traditional lecture-based format. This will allow eventually for the comparison of the 2009-2010 POGIL group graduation rates.

Results and Discussion

Initially the chemistry faculty and I were interested in determining if students exposed to POGIL were more successful in CHEM 111 than those students that were not exposed to this teaching pedagogy. To determine this, we analyzed performance across all six sections on common multiple choice, in-class test questions on a standardized final examination published by the American Chemical Society (ACS). This information was gathered at the end of last year as a way of telling whether or not this project had grounds to continue; does POGIL *look* successful? Table 1 shows the exam data collected for the Fall 2009 cohort. Based on these numbers, we can make some inferences about student performance in the POGIL versus non-POGIL environment.

	Test	N	Mean Score (%)	Std. Deviation (%)
Traditional	1	48	70	19
POGIL		88	78	16
Traditional	2	43	56	19
POGIL		85	65	13
Traditional	3	43	56	18
POGIL		79	66	13
Traditional	ACS Final *(raw scores)	44	*32	10
POGIL		83	*37	9

Table 1. Fall 2009 test score comparisons based on common questions

The mean scores and standard deviations seen in Table 1 were derived from common multiple choice questions on each exam. In-class tests varied slightly between sections as syllabus schedules tend to be

flexible and instructors do not all teach in the same order or at the same speed. The ACS first semester exam was published in 2005. Two-sample t-tests were conducted to determine if the average for the POGIL sections for each test was significantly different from the average for the traditional sections. The average for each in-class test and the final exam were statistically different at the 95% confidence level. Differences in N values in the table are due to the fact that we were advised by POGIL colleagues to consider Dr. Porter's class (POGIL once per week) a POGIL class (Hunnicutt, 2009). From these data, we concluded that the POGIL method did have an impact on student learning, and that collection of data from subsequent cohorts was warranted.

Similar analyses were carried out with the 2010-2011 cohort. Surveying the literature convinced us that we needed to collect some additional data on the cohort in order to address some additional interesting questions. Table 2 shows SAT math and verbal scores, pre- and post-concept inventory test scores, and GALT test scores.

	SAT Math	SAT Verbal	GALT	Pre-CIT	Post-CIT
Traditional	515 ± 55	520 ± 66	7 ± 3	33 ± 15	36 ± 16
POGIL	513 ± 71	516 ± 62	8 ± 2	31 ± 13	36 ± 17

Table 2. 2010-2011 array of testing scores

In order to better understand our science major population the GALT (Group Assessment of Logical Thinking) test was administered to the 2010/11 cohort at the start of the fall semester. This instrument is a twelve-item test developed by Roadrangka et al. (1982) to measure logical reasoning skills in both college and pre-college level students. The test includes questions on the conservation of mass and volume, proportional reasoning, correlational reasoning, control of experimental variables, probabilistic reasoning, and combinatorial reasoning. The two items related to the conservation of mass

and volume are used to assess the concrete operational stage (achieved during the 7-12 year age range), whereas the remaining 10-items correspond to the formal operational level (begins at age 12). This test is designed to determine the Piagetian cognative level of the responder. GALT scores of 0-4 are characteristic of concrete thinkers, 5-7 of individuals in the concrete to formal transitional stage, and 8-12 of formal thinkers (Roadrangka, 1982).

At Longwood, 10 of the 12 items were administered via Blackboard. Each item is comprised of two parts; the first was a multiple choice question not necessarily related to an explicit chemistry topic, whereas the second part, also multiple choice, asked for a justification for the answer given in the first part of the question. An example of one of the first 10-items on the GALT test is shown in Figure 2.

Metal Weights

Linn has two jars. They are the same size and shape. Each is filled with the same amount of water.
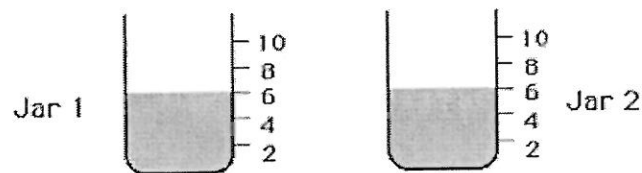
She also has two metal weights of the same volume. One weight is light. The other is heavy.

heavy metal weight   light metal weight
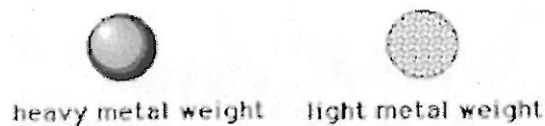
She lowers the light weight into jar 1. The water level in the jar rises and looks like this:
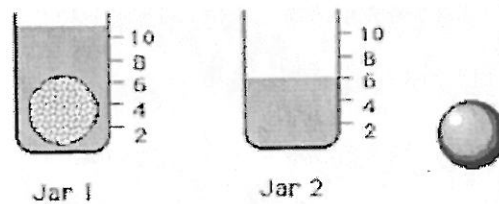
IF THE HEAVY WEIGHT IS LOWERED INTO JAR 2, WHAT WILL HAPPEN?

    A. The water will rise to a higher level than in jar 1.

    B. The water will rise to a lower level than in jar 1.

    C. The water will rise to the same level as in jar 1.

SELECT THE REASON FOR YOUR ANSWER:

    1. The weights are the same size so they will take up equal amounts of space.

    2. The heavier the metal weight, the higher the water will rise.

    3. The heavy metal weight has more pressure, therefore the water will rise.

    4. The heavier the metal weight, the lessr the water will rise.

Figure 2. An example of one of the first 10-items on the GALT test.

The last two questions on the GALT are free-response, asking students questions about different combinations or arrangements of objects and having them list all possibilities. On average there is no difference in overage GALT score for students in the traditional versus POGIL sections for this cohort. Interestingly, of the 65 students in the POGIL sections, 14% were classified as concrete thinkers, 31% transitional and 55% formal. Of the 65 students in the traditional sections 12% were operating at a concrete level, 39% at a transitional stage, and 49% had reached the formal operational level. The findings above are comparable to those obtained by Bird (2010) and McConnell et al. (2005). Even though our results agree with the standard spread of GALT scores, it is still disconcerting that approximately 50% for the students enrolled in CHEM 111 fall below the formal operational level as mastery of many topics in this course requires formal reasoning ability.

We also decided to give the students a pre- and post-concept inventory test. The exam chosen was developed in the mid-1990s as a master's thesis by Douglas R. Mulford of Purdue Universtiy. This 22-question exam covered the majority of the concepts presented in CHEM 111. The results of this

comparison were rather disappointing in that the students in either teaching environment did not, on

average, perform statistically better on the test despite taking the course (see Table 1).

We then looked at the averages on common questions from in-class tests for this cohort as well,

which surprising looked very similar (Table 3).

	Test 1	Test 2	Test 3	ACS Final
Traditional	65 ± 16	54 ± 14	60 ± 22	53 ± 19
POGIL	65 ± 14	66 ± 16	64 ± 21	53 ± 21

Table 3. Test score comparison for 2010-2011 group

Only Test 2 showed a significant difference in performance as determined by a t-test (95% confidence

level). However, looking at each exam on a question-by-question basis, there are some interesting

things to note. Figures 3-5 show the difference in testing scores calculated by subtracting the percent

of traditional students who answered correctly from the percent of POGIL students who answered
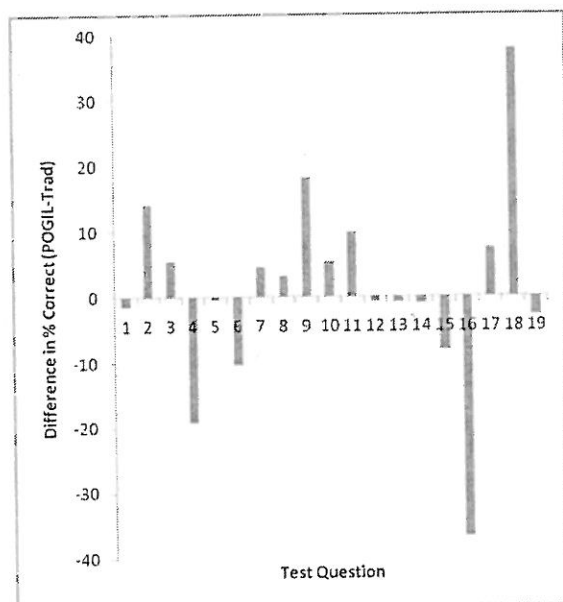
Figure 3. Test 1 question-by-question analysis.
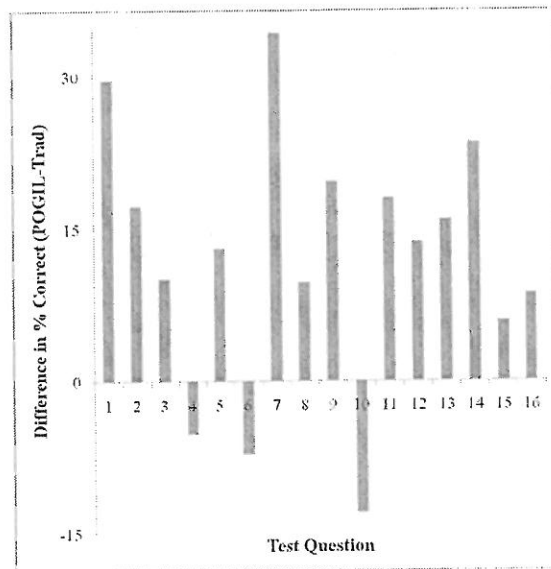
Figure 4. Test 2 question-by-question analysis.

correctly. In each case there are not that many questions where the traditional students outperformed the POGIL students.
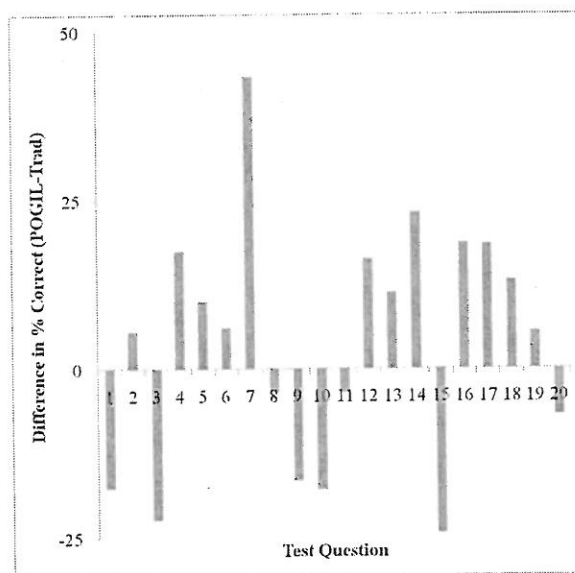

Figure 5. Test 3 question-by-question analysis

The final exam administered to this cohort was different than the one used in Fall 2009. Although this exam, the Paired Conceptual General Chemistry Exam, is published by the American Chemical Society, it is designed to allow for comparison of student performance on conceptual versus algorithmic questions that are paired. It was anticipated that the POGIL students would perform better on the conceptual questions than the traditional students simply because of the manner in which they learned the material, i.e., in a more conceptual framework. However, inspection of Figures 6 & 7 show that there was no statistical difference in performance amongst the two groups.
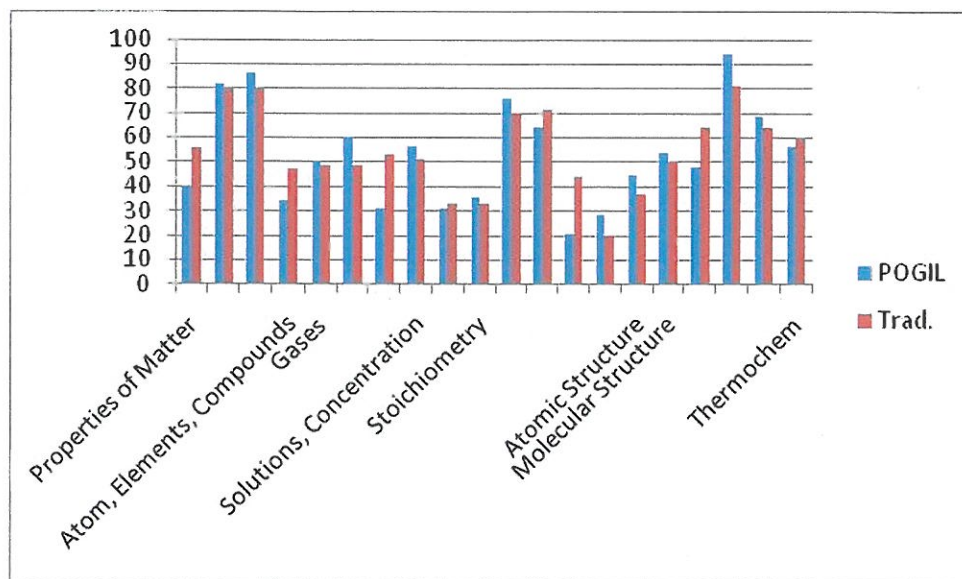
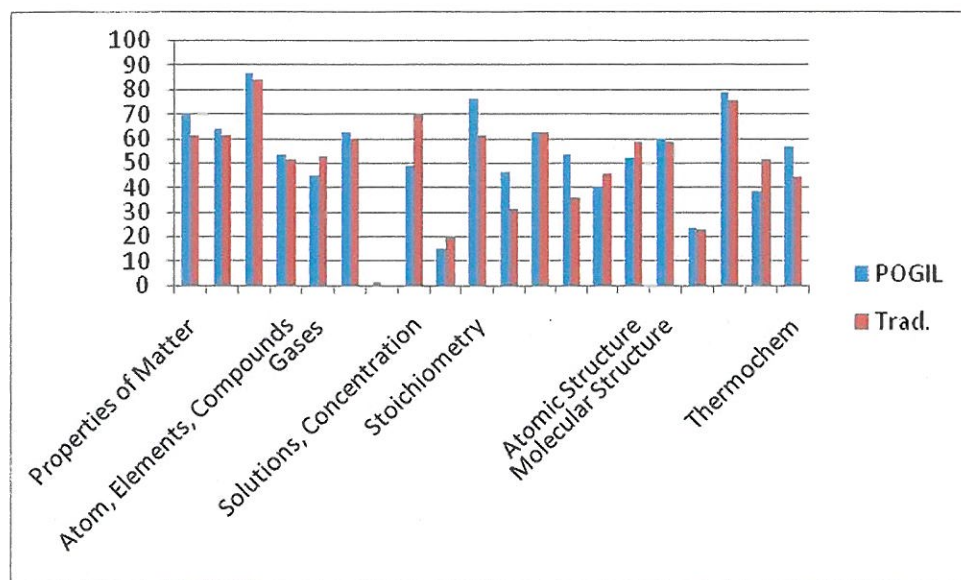Figure 6. Comparison of the conceptual question performance on the ACS exam

Figure 7. Comparison of the algorithmic question performance on the ACS exam

Conclusions and Future Work

There is now a solid foundation on which this project can move forward. Data compilation is a time consuming process and should be kept up with as closely as possible. Waiting two or three years and then attempting to catch up will inundate the researcher. Keeping all information on each student condensed in one spreadsheet makes data analysis and calculation very simple. So far, it seems that the POGIL classroom does make for better test scores, but there is no correlation between other variables present (i.e. SAT scores, etc.) and student test performance. From this, it can be gathered that this project has the potential to demonstrate POGIL's power to convey new material and foster success on the path to an undergraduate science degree.

An interesting change on the horizon is that Dr. Porter will be teaching the general chemistry class alone in the fall; a situation that has not been encountered so far in this study. This raises further questions as to what teaching method will be employed and how sections will be broken up. In the past, we know she has utilized POGIL teaching methods. This could potentially be a semester to compare class success against years prior to 2009, before the POGIL option was offered. Assessment & Institutional Research will keep receiving new data as the next three years go by, and this will be an opportunity to collect all retro data on the classes of 2011, 2012, and 2013. It appears in the next 3-5 years there will be sufficient data to support a preliminary claim, but research should continue beyond the bare minimum. With data compiled as of right now, there is a marked difference in the overall teaching styles of POGIL and Traditional lecture. It seems possible that those differences in teaching style could manifest themselves in overall student performance, which is why further tracking this group is required. When concluded, findings could vastly improve the chemistry program both here at Longwood University but chemistry departments throughout the United States.

In future work, use of SPSS or other statistical programming should be done in order to further investigate correlations between sets of data. For example; is there a correlation between SAT and

GALT scores and success in the general chemistry sequence? Perhaps there is a correlation between GALT scores and scores on the conceptual ACS exam. If these correlations exist, are they indicative of the students' learning environment (POGIL vs. Traditional)? It should be made quite clear that while these tests may not take very long to run, there is a learning curve involved before SPSS can be of any use, especially for these more complicated tests. While it was proposed that this step would already be done, we underestimated the amount of time needed for basic familiarity with the program and were unable to run the numbers.

Acknowledgements

References

1. Bird, L. 2010. Logical Thinking Ability and Student Performance in General Chemistry. *Journal of Chemical Education.* 87: 541-546.

2. Bodner, G. M. 1986. Constructivism: a theory of knowledge. *Journal of Chemical Education.* 63. 873-877.

3. Bunce, D. M. 2001. Does Piaget Still Have Anything to Say to Chemists? *Journal of Chemical Education.* 78: 1107-1119.

4. Butts, D. P., Karplus, R. 1977. Science teaching and the development of reasoning. *Journal of Research in Science Teaching.* 14: 169-175.

5. Caprio, M. J., Schnitzer, D. K. 1999. Academy Awards. *Educational Leadership.* 57: 46-49.

6. Herron, J. D. 1975. Piaget for Chemists: Explaining what "good" students cannot understand. *Journal of Chemical Education.* 57: 146-150.

7. Hunnicutt, S., personal communication, 2009.

8. Johnstone, A. H. 1997. Chemistry Teaching- Science or Alchemy? *Journal of Chemical Education.* 74. 262-269.

9. Johnstone, A. H. 2010. You Can't Get There from Here. *Journal of Chemical Education.* 87: 22-29.

10. McConnell, D. A., Steer, D. N., Owens, K. D., Knight, C. C. 2005. Challenging Students' Ideas About Earth's Interior Structure Using a Model-based, Conceptual Change Approach in a Large Class Setting. *Journal of Geoscience Education.* 53: 462-470.

11. McDonnell, Robert. Commonwealth of Virginia. *Preparing for the Top Jobs of the 21st Century.* Richmond: Governor's Commission, 2010. Print.

12. Pienta, N. J., Cooper, M. M., Greenbowe, T. J. in Chemists' Guide to Effective Teaching (Volume I), Prentice Hall: Upper Saddle River, NJ 2005.

13. POGIL: Process Oriented Guided Inquiry Learning. About POGIL. http://www.pogil.org/about

14. Roadrangka, V., Yeany, R. H., Padilla, M. J. Group Assessment of Logical Thinking Test (Masters Thesis); University of Georgia: Athens, GA, 1982.