Comparing the Process Oriented Guided Inquiry Learning (POGIL) Method to an Independently Developed Guided Inquiry Method (InDGIM) in a High School Academic Chemistry Course.

Scott Zgraggen
Arcadia University, zgrags86@arcadia.edu

Follow this and additional works at: https://scholarworks.arcadia.edu/grad_etd

Part of the Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, and the Science and Mathematics Education Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Graduate Works at ScholarWorks@Arcadia. It has been accepted for inclusion in Graduate Theses & Dissertations by an authorized administrator of ScholarWorks@Arcadia. For more information, please contact gordonl@arcadia.edu.
Comparing the Process Oriented Guided Inquiry Learning (POGIL) Method to an Independently Developed Guided Inquiry Method (InDGIM) in a High School Academic Chemistry Course.

Scott M. Zgraggen

Arcadia University
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Abstract

A nonequivalent, control group, pretest-posttest design was used to investigate student achievement in secondary chemistry. This study investigated the effects of process oriented guided inquiry learning (POGIL) pedagogy in high school chemistry classes compared to that of an independently designed guided inquiry method (InDGIM). Data were collected from chemistry students from four college-prep chemistry classes in the same high school, over the course of the academic year, using the Particulate Nature of Matter (ParNoMA2) test, the Group Assessment of Logical Thinking (GALT), unit tests, and the final exam. Data were analyzed using a 2 x 2 Factorial Analysis of Covariance. This ANCOVA examined the main effects of group and gender on test results. The results show that there are no statistical differences in outcomes between the POGIL and the InDGIM group, nor are there statistical differences in performance between males and females with either approach, although females perform better than males overall. There are no interaction effects between group and gender. Students in the lowest quartiles, as per their Keystone Algebra scores (an end-of-course assessment designed by the Commonwealth of Pennsylvania to assess proficiency in multiple subjects) show no differences in performance with either methodology, whereas students in the upper quartile in the InDGIM group perform better than those in the POGIL group. Students perform better on almost all of the unit tests and the final exam when taught with the InDGIM rather than the POGIL approach. The results of the Keystone Biology test is a better indicator of student success in chemistry compared to the results on the Keystone Algebra test. Regression analysis indicates that students with higher Keystone Biology scores are 34.8% more likely to earn a higher score
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

on the chemistry final exam and are 20% more likely to earn a higher final grade in the chemistry course.

*Keywords:* active student-centered pedagogy, chemistry education, cooperative learning, group assessment of logical thinking, GALT, inquiry, independently developed guided inquiry method, particulate nature of matter, ParNoMA2, process oriented guided inquiry learning, POGIL, socialization.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Dedication and Acknowledgement

This has been an arduous yet fulfilling journey. It could not have happened if not for the support I received along the way. I must thank my wonderful wife, Marty, a woman who has constantly encouraged me to reach higher than I thought possible and has seen more in me than I have seen in myself. She and my two sons, Nick and Derek, kept me going when I lost sight of the purpose of this pursuit. Your belief in me and encouragement helped me more than you will know.

Along the way I found compadres who modeled how to juggle family life, life changes, long classes, and educational jobs, all while joining me on this trek. They kept a level head through it all. I think I might have given up sooner if not for Cyndi, Shawn, Wendy and Ryan, my modern Breakfast Club.

The research couldn’t have been collected if it weren’t for Scott, a fellow educator, willing to help put two guided inquiry methods to the test. I appreciate you providing me a space to do this research. Also thanks to Dr. Nancy Hacker, Superintendent, and Dr. Chuck Rittenhouse, Principal, for allowing me to conduct the research at “Falls Fallows High School.”

It all came together with the guidance of a mentor who taught me that this research didn’t need to be “perfect,” that it just needed to reflect my best abilities and a committee that made sure that happened. Thank you Dr. Steve Gulkus for all of your efforts and guidance; Dr. Tanya Santagelo for your analysis; and Dr. Rebecca Craik for your support and friendship.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Table of Contents
Dedication and Acknowledgement........................................................................................................ 4
List of Tables ........................................................................................................................................ 8
List of Figures ..................................................................................................................................... 8
Chapter One: Introduction ...................................................................................................................... 9
  Background ........................................................................................................................................ 9
    Definition of Inquiry.................................................................................................................... 10
    Comparing Inquiry to Direct Instruction ............................................................................ 16
    Inquiry and Gender ................................................................................................................ 19
    Which Inquiry Method? ........................................................................................................... 21
  Statement of Problem .................................................................................................................. 21
  Conceptual Framework .................................................................................................................. 24
  Purpose of the Study ..................................................................................................................... 30
  Research Questions ....................................................................................................................... 31
  Significance of the Study ............................................................................................................... 32
  Summary of Design ....................................................................................................................... 34
  Summary of Introduction ............................................................................................................... 35
Chapter Two: Literature Review ........................................................................................................... 37
  Process Oriented Guided Inquiry Learning ................................................................................. 37
  Effectiveness of POGIL .................................................................................................................. 38
    Colleges and Universities ............................................................................................................ 38
    High School ............................................................................................................................. 42
  Challenges and Strategies for Implementing POGIL ............................................................... 44
  Perceptions and Attitudes Resulting from POGIL Implementation ...................................... 47
  POGIL and Gender ....................................................................................................................... 49
  Summary ....................................................................................................................................... 50
Chapter Three: Methods ....................................................................................................................... 52
  Research Questions and Hypotheses ............................................................................................. 52
  Hypotheses ..................................................................................................................................... 54
  Research Design ........................................................................................................................... 55
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Fidelity ...................................................................................................................................... 58
Settings and Participants ........................................................................................................... 63
Instrumentation .......................................................................................................................... 68
Procedure ................................................................................................................................... 72
  Experimental Phase/Intervention ..................................................................................... 73
  Comparison Group (InDGIM) ......................................................................................... 77
Data Analysis ................................................................................................................................. 78

Chapter Four: Results ................................................................................................................... 82
Descriptive Statistics for the Pretest and Posttest ........................................................................... 84
Fidelity of Implementation ........................................................................................................... 87
  Internal Consistency Reliability ....................................................................................... 91
Results of Group Differences ........................................................................................................ 92
Covariate Effect ............................................................................................................................. 94
  Null Hypothesis and Research Question One ............................................................... 95
  Null Hypothesis and Research Question Two ............................................................... 97
  Null Hypothesis and Research Question Three ............................................................. 98
  Research Question Four ................................................................................................ 100
  Null Hypothesis and Research Question Five ............................................................. 106
Multiple Regression: Final Examination Scores ............................................................................. 109
Summary ................................................................................................................................. 113

Chapter Five: Conclusion ........................................................................................................... 115
Research Question One and Null Hypothesis One .................................................................. 115
Research Question Two and Null Hypothesis Two .................................................................. 116
Research Question Three and Null Hypothesis Three ............................................................ 117
Research Question Four .......................................................................................................... 117
Research Question Five and Null hypothesis Four ................................................................. 118
Discussion ............................................................................................................................... 119
  Theoretical Framework .................................................................................................. 120
  Similarities and Differences in InDGIM and POGIL Approaches ................................ 122
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

List of Tables

Table 3.1. Nonequivalent pretest posttest comparison group design ............................................ 56
Table 3.2. The different effects examined by ANCOVA ............................................................. 80
Table 4.1. Descriptive statistics of pretest and posttest by group ............................................... 85
Table 4.2. Descriptive statistics of pretest and posttest by gender ............................................... 86
Table 4.3. Rater agreement to determine if activity should be classified as a POGIL ............. 88
Table 4.4. Rater agreement to determine if activity should be classified as a non-POGIL ....... 89
Table 4.5. Degree of agreement between two raters on “inquiry” based activities ....................90
Table 4.6. Cronbach’s alpha reliability estimates for the ParNoMA2 and the GALT tests .... 92
Table 4.7. Covariance for ParNoMA2 posttest scores for grouping and gender ................. 93
Table 4.8. Covariance for GALT posttest scores for grouping and gender ............................ 94
Table 4.9. Statistics for groups on the ParNoMA2 and GALT posttests from the ANCOVA .... 96
Table 4.10. Statistics for gender on the ParNoMA2 and GALT posttests from the ANCOVA .... 97
Table 4.11. Statistics for groups*gender on the ParNoMA2 posttest from the ANCOVA ....... 99
Table 4.12. Statistics for groups*gender on the GALT posttest from the ANCOVA .......... 100
Table 4.13. Statistics of pre and posttest by lowest quartile for both groups ......................... 101
Table 4.14. Statistics of pre and posttest by lowest quartile for both genders ....................... 102
Table 4.15. Statistics of pre and posttest by highest quartile for both groups ....................... 104
Table 4.16. Statistics of pre and posttest by highest quartile for both genders ..................... 105
Table 4.17. Comparison of the outcomes of the assessments given to both groups .......... 107
Table 4.18. Means, standard deviations, and intercorrelations for the Final Exam grade ..... 110
Table 4.19. Simultaneous multiple regression analysis for gender, instructional treatment, Keystone Algebra Scores, and Keystone Biology Scores predicting the Final Exam grade ..... 111
Table 4.20. Means, standard deviations, and inter-correlations for the final grade .......... 112
Table 4.21. Simultaneous multiple regression analysis for gender, instructional treatment, Keystone Algebra Scores, and Keystone Biology Scores ............................................... 112
Table 5.1. Biology Keystone assessment anchors and eligible content .................................. 137
Table 5.2. Holistic developmental rubric on the use of inquiry-guided learning ............... 142

List of Figures

Figure 1. 1. Four levels of inquiry teaching based on the level of ownership ......................... 14
Figure 1. 2. Seven levels of inquiry, the level of sophistication and level of control ........... 15
Figure 1. 3. A rubric that can be used to determine the level of inquiry of an activity ........ 16
Figure 3. 1. Research design and implementation ................................................................. 62
Figure 4. 1. A scatter diagram illustrating the linear relationship between the residual and the predicted values for the final grade (as a percentage) .................................. 109
Chapter One

Introduction

Background

The National Science Education Standards (NSES) (Bransford, Brown, & Cocking, 1996) and the science standards established by the Commonwealth of Pennsylvania (Pennsylvania Department of Education, 2015) set clear expectations for student learning, indicating that students should have an integrated understanding of the concepts and processes of science and a firm grasp of the nature and structure of the discipline (Rhoton, 2001). Traditional teacher led instructional practices have not always been successful in delivering these student outcomes. Research has revealed teaching strategies that support the goals and instructional practices outlined in NSES (Rhoton & Bowers, 2001). Such practices include:

- using inquiry and problem-solving lessons, active student participation, and frequent teacher–student interactions;
- creating learning environments in which risk is supported and open discussion and use of student ideas takes place;
- implementing lessons that provide an accurate portrayal of content knowledge, the nature of science, and the structure of the discipline;
- selecting and adapting curriculum to meet the needs of all students;
- implementing learning environments that challenge students’ misconceptions;
- using discrepant events to facilitate student learning;
- and using a variety of techniques to assess student learning (Rhoton, 2001, p. 16).

Current research, described in this dissertation, supports this shift in teaching habits towards more inquiry-based learning.
Many science educators have come to embrace the idea of inquiry or inquiry-based instruction when it comes to teaching students (Coburn, 2000). The NSES place the development of an inquiry-based science program at the heart of all effective science teaching (Bransford et al., 1996) and Project 2061’s Benchmarks for Science Literacy discusses scientific inquiry throughout its 448 pages (American Association for the Advancement of Science, 1993) as it applies to all students, regardless of race, gender or socioeconomic background. The term “inquiry-based” has become synonymous with quality science education.

**Definition of inquiry.**

Perhaps one of the most confusing things about inquiry is its definition since it can be used to describe both teaching and performing science. NSES notes this dichotomy as it states that scientific inquiry refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work (Bransford et al., 1996). Inquiry can be used alternatively to describe instructional approaches and curriculum materials as put forth by Furtak, Seidel, Iverson, and Briggs (2012). In this paper, inquiry is referred to as a learning process in which students are engaged, or are active, in the learning process. It is “something that students do, not something that is done to them” (Anderson, 2002, p. 3) and it is created through “a classroom where students are engaged in essentially open-ended, student-centered, hands-on activities” (Colburn, 2000, p. 42). Some educators see current inquiry learning and teaching to be on a continuum from the more traditional, structured inquiry, through guided inquiry, to the ultimate open-ended student-led inquiry or discovery learning (Furtak, et al., 2012). The continuum is bordered on one side by
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

“traditional, direct instruction in which students are told the answers they are expected to learn by their teacher and at the other end of the continuum, students design and conduct their own investigations into phenomena that are not known to the teacher in what can be called open-ended scientific inquiry” (Furtak, 2006, p. 454). Operationally, these differences can be looked at through the role of the teachers, students and the students’ work.

With the proper implementation of inquiry-based instruction, teachers become facilitators, modeling the learning process, and assisting students in processing information as they learn to interpret, explain, hypothesize, and design their own activities centered on scientifically oriented questions (Anderson, 2002). Collaborative activities are guided by learners’ curiosity and interests, through which the students learn process skills, e.g., critical thinking, and provide opportunities for rich interaction with the material; thus, students achieve a deeper understanding of the content and become better able to apply their knowledge (Saunders-Stewart, Gyles, & Shore, 2012). This allows the learner to give priority to the evidence in responding to questions, formulate explanations from the evidence, connect those explanations to scientific knowledge, then communicate and justify those explanations. This inquiry approach can be presented by the instructor as either structured inquiry, guided inquiry, open-ended inquiry, or as part of the discovery learning cycle or self-directed inquiries (Colburn, 2000). Llewellyn (2011) defines each category based on whether the teacher and the student have low/passive or high/active ownership of the material being learned.

In structured inquiry, the teacher provides the students with a hands-on problem to investigate, as well as the procedures and the materials, but does not inform them of expected outcomes (Colburn, 2000). Students are allowed to discover relationships between variables or
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

generalize from the data collected. These types of investigations are similar to those known as cookbook activities, “although a true cookbook activity generally includes more direction than a structured inquiry activity about what students are to observe and which data they are to collect” (Colburn, 2000, p. 42). As an example of this approach, in physical science, students are given a step-by-step procedure, including diagrams, for making various electrical circuits. Questions prompt students to remove individual bulbs from each circuit and record their observations. Chances are that structured inquiry activities will result in all students having the same data and reaching the same conclusions.

In guided inquiry, the teacher provides only the materials and problems to investigate and the students are asked to devise and implement their own procedures to solve the problem (Colburn, 2000). Following the electricity example, students are given batteries, bulbs, wires, and other materials and procedures which instruct them to make a bulb light up as many ways as they can, using the supplies provided. Later, students are instructed to make two bulbs light, again, using different combinations of materials. Finally, students are asked to note what happens when they remove individual bulbs from their circuits. Thus with guided inquiry activities, students are given some instruction and a problem to solve, however, not all students will go about obtaining the solution in the exact same way.

The approach to open-ended inquiry is similar to guided inquiry, but with the addition that students also formulate their own problem to investigate (Colburn, 2000). Open-ended inquiry, in many ways, is analogous to ‘doing science,’ which means that students use observations and experimentation to describe and explain natural phenomena. Science fair activities and senior thesis research projects are examples of open-ended inquiry in that students
set up experiments to collect data that supports or rejects a given hypothesis. In the electricity example, students are given batteries, bulbs, wires, and other materials and they are instructed to investigate how bulbs light in electrical circuits. Students are not given much procedural detail, therefore, results vary with open-ended inquiry, but there are somewhat similar conclusions among the students.

Finally, with the discovery learning cycle or self-directed inquiries, students engage in an activity where the teacher introduces a new concept. Students take ownership of the concept by applying it in a different context which is why it is often considered the pinnacle of discovery learning (Colburn, 2000). Students might follow guided inquiry or open-ended inquiry procedures, but then the teacher discusses their findings and students are allowed to make larger connections. In the circuitry example, concepts such as series and parallel circuits are introduced to students who have some experience with the concepts. They eventually return to the lab to apply what they have learned to a new situation. For example, students could be given additional equipment, such as ammeters or voltmeters, to quantitatively investigate current and voltage in circuits, collecting data and developing new connections or hypotheses (Colburn, 2000).

According to Borthick and Jones (2000), the end result of discovery learning is that “participants learn to recognize a problem, characterize what a solution would look like, search for relevant information, develop a solution strategy, and execute the chosen strategy” (p. 181). Figure 1.1 shows the four different categories of inquiry proposed by Llewellyn (2011) and the different levels of ownership the instructor or the student have in each one.
Figure 1.1. Four levels of inquiry teaching based on the level of ownership both the instructor and the student have in the style of teaching. Adapted from *Differentiated Science Inquiry* by D. Llewellyn, 2011, p. 12. Copyright 2011 by Corwin.

Wenning (2005, 2007) breaks down the spectrum of inquiry teaching even more finely, using seven levels, basing his levels not only on the level of student or instructor control, but also the level of intellectual sophistication the student needs to have to be able to accomplish and function at the different levels of inquiry teaching. Wenning’s (2005) seven levels of inquiry are discovery learning, interactive demonstrations, inquiry lessons, guided inquiry labs, bounded inquiry labs, free inquiry labs, and pure or applied hypothetical inquiry. Figure 1.2 shows the seven levels of inquiry described by Wenning (2005, 2007), differences in the levels of control that the teacher or the student has for each level, and the level of intellectual sophistication students need to function at each level. Wenning (2007) recommends, that during a high school science course (in this case physics), the instructor progress from the low end of the inquiry spectrum to the high end of the inquiry spectrum as students advance in their knowledge and their inquiry ability.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

| Discovery Learning | Interactive Demos | Inquiry Lessons | Guided Inquiry Labs | Bounded Inquiry Labs | Free Inquiry Labs | Pure Hypothetical Inquiry
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Intellectual Sophistication</td>
<td>Locus of Control</td>
<td>High</td>
<td>Teacher</td>
<td>Student</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1.* This figure shows the seven levels of inquiry: the level of sophistication the student needs to function and the level of control the student has over the teacher increases from left to right along the chart. Adapted from “Levels of Inquiry: Hierarchies of Pedagogical Practices and Inquiry Processes” by C. J. Wenning, 2005, *Journal of Physics Teacher Education Online*, 2(3), p. 10. Copyright 2005 Illinois State University Physics Department.

No matter what variation, level, or category of inquiry an educator chooses it is commonly understood, according to St. Omer (2002), that unless students understand the language associated with science knowledge, they will not be able to retain information and will fail to gain mastery. St. Omer also stated that it is during inquiry-based laboratories and activities that students have the best opportunity to understand and retain this science knowledge. To assist science teachers, Bruck, Bretz, and Towns (2008) created, tested and validated, a rubric to characterize the level of inquiry used in college laboratory activities or exercises. The level of inquiry for the laboratory or activity is based on six different characteristics that each of the laboratories or activities contain and whether or not the information for that characteristic is provided by the instructor. The six different characteristics are problem/question, theory/background, procedure/design, results analysis, results communication, and conclusions; the fewer characteristics that are provided by the instructor the higher the level of inquiry. Bruck et al., (2008) also classify five levels of inquiry for the laboratory: confirmation, structured inquiry, guided inquiry, open-ended inquiry, and authentic inquiry. Each of these levels of
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

inquiry and the characteristics provided by the instructions in the laboratory can be seen in Figure 1.3. If this information cannot be provided by the instructor, then students need to determine that information for themselves. As one progresses from left to right across figure 1.3 a higher level of inquiry is involved. Based on this rubric, any laboratory that an instructor is using can be placed at a specific level of inquiry.

**Comparing inquiry to direct instruction.**

Regardless of which definition of inquiry is chosen, accumulating evidence indicates that direct instruction, by systematically guiding the student to solve one predetermined question, is insufficient in developing critical and scientific thinking, appropriate dispositions and attitudes (Sadeh & Zion, 2009). Sadeh and Zion’s (2009) quasi-experimental approach divided 50 11th and 12th grade students into lower and higher levels of inquiry. The rubric used to determine these levels is presented in Figure 1.3.}

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Confirmation</th>
<th>Structured Inquiry</th>
<th>Guided Inquiry</th>
<th>Open Inquiry</th>
<th>Authentic Inquiry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem/Question</td>
<td>Provided</td>
<td>Provided</td>
<td>Provided</td>
<td>Provided</td>
<td>Not Provided</td>
</tr>
<tr>
<td>Theory/Background</td>
<td>Provided</td>
<td>Provided</td>
<td>Provided</td>
<td>Provided</td>
<td>Not Provided</td>
</tr>
<tr>
<td>Procedures/Design</td>
<td>Provided</td>
<td>Provided</td>
<td>Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
</tr>
<tr>
<td>Results Analysis</td>
<td>Provided</td>
<td>Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
</tr>
<tr>
<td>Results Communication</td>
<td>Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
</tr>
<tr>
<td>Conclusions</td>
<td>Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
<td>Not Provided</td>
</tr>
</tbody>
</table>

Less Inquiry | More Inquiry

*Figure 1.3.* Rubric used to determine the level of inquiry in any laboratory exercise or activity based on information provided by the instructor. Adapted from “Characterizing the Level of Inquiry in the Undergraduate Laboratory,” by L. B. Bruck, S. L. Bretz and M. H. Towns, 2008, *Journal of College Science Teaching*, 38(1), p. 54. Copyright 2008, National Science Teachers Association.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

12 grade students into two equal groups, one taught by guided inquiry and the other taught by open-ended inquiry. Qualitative data in the form of interviews, students’ inquiry summary papers, logbooks, and reflections were collected and analyzed using a dynamic inquiry performance index. The quantitative data looked at “student performance according to four criteria: changes occurring during inquiry, learning as a process, procedural understanding, and the affective point of view, as well as a total summary score” (Sadeh & Zion, 2009, p. 1153) or what the authors labeled a dynamic inquiry skill level. A unidirectional MANOVA analysis of dynamic inquiry levels, performed by the Sadeh and Zion, revealed a significant difference between the two groups. The study concluded that open-ended inquiry students used significantly higher levels of performances in the criteria changes during inquiry than those students in the guided inquiry group and that “open-ended inquiry students possess deeper procedural understanding and can often express this when facing problems and difficulties in class, whereas students learning with the guided inquiry process, only follow the teachers’ instructions when planning and conducting inquiries” (Sadeh & Zion, 2009, p. 1155).

Berg, Bergendahl, Lundberg, and Tibell (2003) compared 190 university students’ outcomes of an open-ended inquiry activity and structured laboratory activity. Using interviews, questions asked during the experiment, and students’ self-evaluations, the authors of this qualitative study concluded that the open-ended inquiry activity showed more positive gains regarding learning outcomes, preparation time, time spent in the laboratory, and students’ perception of the experiment (Berg et al., 2003).

A qualitative analysis of high school laboratory manuals by Germann, Haskins, and Auls (1996) claimed that recipe-like activities often short circuit opportunities to stimulate thinking by
students and that guided inquiry can be used to help students make the transition from a structured direct instruction to an open-ended inquiry level. Interestingly, as students move from teacher directed to guided-inquiry, they transform the data into complex and abstract forms, such as graphs and concepts maps, more so than in structured activities, as was determined in a qualitative study designed to investigate how 15 pre-service science teachers reported on their observations of two organisms by Lunsford, Melear, Roth, Perkins, and Hickok (2007). The open-ended inquiry process appears to guide students to construct their own knowledge in ways that direct instruction cannot. It motivates students when they are confronted with an “authentic problem and they must take risks to solve the problem and it encourages students to conceptualize the dynamic and ever-changing nature of the scientific process” (Sadeh & Zion, 2009, p. 1138).

Blanchard, Southerland, Osborne, Sampson, Annetta and Granger (2010) examined the relative effects of direct instruction and inquiry-based teaching techniques for achievement, on standardized tests in a quasi-experimental quantitative study involving 1,700 students in 12 middle schools and 12 high schools. The data from the pre-, post-, and delayed posttests were analyzed using a Hierarchical Linear Model involving students’ scores, teacher, level of school, teacher observation scores, and school socioeconomic status (Blanchard, Southerland, Osborne, Sampson, Annetta & Granger, 2010). The findings, from their analysis, showed that those students receiving guided/open-ended inquiry-based laboratory instruction on concepts related to the course had stronger gains in various types of knowledge and generally better long-term retention (as measured with pre-and posttests) over time than students who received traditional, teacher directed laboratory instructions (Blanchard et al., 2010) regardless of grade level or
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

socioeconomic level of the school.

Saunders-Stewart et al. (2012) in their meta-analysis of the literature on student outcomes in inquiry environments, found that student participation in inquiry activities resulted in an increase “in a variety of different outcomes, including cognitive, metacognitive, affective, personal, and societal” (p. 5). When engaging in inquiry, students were able to “describe objects and events, ask questions, construct explanations, test those explanations against current scientific knowledge, communicate their ideas to others, identify their assumptions, use critical and logical thinking, and consider alternative explanations” (Saunders-Stewart et al., 2012, p. 6). In this way students are able to develop actively their understanding of science by combining knowledge with critical thinking skills.

Inquiry and gender.

Many science educators, including this researcher, agree that both guided and open-ended inquiry can be efficient in developing inquiry skills and critical thinking for both males and females. While refining his skills and pedagogy as a student-teacher this researcher was encouraged, by his master teacher and his professors, to incorporate the findings of the work by Eccles and Blumenfeld (1985) and Wilkinson and Marrett (1985) regarding gender equality in the classroom into his teaching, and they have become a strong conceptual framework for him. Through this work it was determined that females generally received less attention from teachers than males regardless of the subject or age of students (Wilkinson & Marrett, 1985) and that males were asked more in depth follow up questions and given more comments on the ideas represented in their work, than females, and that females were more frequently complemented for the neatness of their work (Eccles & Blumenfeld, 1985). Although females typically asked
more questions than males, teachers gave less feedback (positive or negative) on the females’ answers (Eccles & Blumenfeld, 1985) and males received more disciplinary attention (negative feedback) than females for verbalizing their responses without being called upon (Morse & Handley, 1985). When the underrepresentation of females in science was reported by Roychoudhury, Tippins, and Nichols in 1995 it was a concern for educators, and as reported by Riegle-Crumb, King, Grodsky, and Muller (2012), it still persists as an issue over fifteen years later.

Gender differences in cognitive skills can be explained by the experiential differences in the classroom (Morgan, Farkas, Hillemeier, & Maczuga, 2016). A strong high school science curriculum “provides more opportunities for concrete experiences of interest and competence and thus provides a partial antidote to gender stereotyping and the discouragement of girls’ interest in STEM fields” (Legewie & DiPrete, 2014, p. 263). Hands-on laboratory experiences incorporating inquiry, which, according to Chiappetta (1997), are relatively infrequent in many high school science classes, continue to be related to all students’ performance, but especially to females’ according to the large and nationally representative, longitudinal study, conducted by Burkham, Lee, and Smerdon (1997) for the National Center for Educational Statistics. The methods teachers use to teach science clearly have an effect on how students perceive the subject. Science pedagogy can re-enforce females’ negative attitudes about science by devaluing the contributions of female students and overemphasizing rote learning over critical thinking (Blickenstaff, 2005). Their findings illustrate the importance of the active involvement of students in the science classroom as a means to promote higher thinking among all students and as a means to promote gender equity.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Which inquiry method?

Evidence, collected over the past two decades, has supported the effectiveness of open-ended inquiry learning in developing skills for inquiry and autonomous learning compared to traditional lecture methods (Sadeh, & Zion, 2009) for all students, however, which type of inquiry is more relevant to the teaching and learning facilities available in high schools remains controversial among educators (Yerrick, 2000; Zion, 2008). Some teachers prefer using guided inquiry whereas others prefer open-ended inquiry (Zion, Cohen, & Amir, 2007). Those who prefer open-ended inquiry claim that this method achieves a higher inquiry level, and the students become more familiar with the nature of science and develop inquiry skills and sharper mental processes (Berg et al., 2003; Krystyniak & Heikkinen, 2007; Yen & Huang, 2001), whereas guided inquiry-based teaching helps students learn science content, master how to do science, and understand the nature of science better (Quintana, Zhang, & Krajcik, 2005; Tabak, Sandoval, Smith, Agganis, Baumgartner, & Reiser, 1995; Trautmann, MaKinster, & Avery, 2004). Whether inquiry is described as guided or open-ended, there still needs to be further investigation into the different inquiry learning practices (Berg et al., 2003; Chin & Chia, 2006; Crawford, 2000; Roth, 1999; Yerrick, 2000).

Statement of Problem

Most of the literature on inquiry in high school science classes tends to compare different inquiry methods (guided or open) to non-inquiry methods (lecture). According to Lazonder and Harmsen (2016) “early research by Bittinger (1968) and Hermann (1969) found inquiry-based learning to be more effective compared with expository forms of instruction” (2016, p. 3). Minner, Levy, and Century (2010) performed a meta-analysis of inquiry strategies between 1984
and 2002 and found that there were significant gains in conceptual understanding across a broad range of science topics, such as air quality, gas laws, kinetic molecular theory, motion, global climate change, and genetics when compared to lecture-based courses. Alfieri, Brooks, Aldrich, and Tenenbaum (2011) reported in their meta-analysis across multiple domains and settings that guided inquiry-based methods resulted in more student learning when compared to those who are taught the same content by expository methods. Furtak et al. (2012) conducted a meta-analysis on 37 experimental and quasi-experimental studies published between 1996 and 2006 and reported that students involved in inquiry lessons outperformed students taught by more traditional methods, such as lecture. This outcome was supported by Carolan, Hutchins, Wickens, and Cumming (2014) in a meta-analysis that compared controlled conditions with little to no learner freedom in Science, Technology, Engineering and Math (STEM). They related army training activities to treatment conditions manipulating more learner freedom as measured on differences between pre and posttests. The results illustrated that approaches that give students the option of exploring a problem were more effective for cognitive skill learning than mere rote memorization or situations where students were told information to remember (Carolan, Hutchins, Wickens, & Cumming, 2014). Although these studies illustrate the strength of inquiry learning, they do not go further to compare the different types of inquiry methods available to science teachers.

There have been limited analyses comparing inquiry methods. One such study compared and contrasted the characteristics features of Problem Based Learning (PBL), Process Oriented Guided-Inquiry Learning (POGIL), and Peer-Led Team Learning (PLTL) in order to enable teachers to decide which approach or combination of approaches would suit a particular situation.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

(Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson, & White, 2008). According to the authors of the normative comparison study, based on their analysis of the research, all three pedagogical approaches promote higher-order thinking skills, help students learn to reason through problems instead of using algorithmic approaches, build conceptual understanding through active engagement with the material, and foster growth in teamwork and collaborative problem-solving skills through an educational constructivist approach (Eberlein et al., 2008).

Many science courses at the high school and college level involve lecture classes, so the additional logistics of creating, scheduling, supervising, monitoring, and training associated with PBL, POGIL, or PLTL courses makes scalability a significant issue according to Eberlein, Kampmeier, Minderhout, Moog, Platt, Varma-Nelson, and White (2008). The POGIL model appears to have some advantages over the PBL and PLTL models in this regard. Because POGIL has a rather “structured format with multiple groups working simultaneously on the same tasks, some scale-up is possible with a corresponding sacrifice in the group contributions to whole class reporting and extra demands on monitoring and attending to the needs of individual groups, however, there are a number of strategies that have been successfully used for implementing POGIL in large classrooms” (Eberlein et al., 2008, p. 269).

According to Eberlein et al. (2008), student perspectives on the use of POGIL showed “overwhelmingly positive attitudes, with fewer than 8% being negative about the method, when compared with 30% who expressed negative attitudes toward the traditional lecture approach or the other inquiry approaches of PBL and PLTL” (p. 271). The students in the comparison study also reported significantly higher gains in their own process skills compared with those students whose classes were taught in either a lecture, PBL or PLTL fashion (Eberlein et al., 2008). Based
on the outcomes of this comparison study, the POGIL approach seems to have some significant strengths over other inquiry methods, however, it is not known how it compares with another inquiry method, the Independently Developed Guided Inquiry Method (InDGIM), created and taught by this researcher and adopted by the chemistry program in the school district being studied.

**Conceptual Framework**

Activities in science classrooms should follow the learning cycle structure of exploration, concept invention, and application in order to be considered inquiry-based (Coburn & Clough, 1997). The learning cycle divides instruction into three major phases: exploratory, content, and application. In the exploratory phase, students have relevant and concrete experiences with the content that will follow; in the content phase, students are introduced more explicitly to the science concepts in question; and in the application phase, students apply what they have learned to a new situation (Coburn & Clough, 1997). The POGIL process and the InDGIM process follow these phases of learning.

In the *exploration phase* of the POGIL method, students examine a model, search for patterns within it, and attempt to extract meaning from it (Eberlein et al., 2008). The model consists of any combination of pictures, tables, equations, graphs, prose, or other types of information. Often, the questions lead students to test hypotheses or explain the patterns and relationships found in the model. According to Eberlein et al. (2008) in the *concept invention or formation* (or term introduction) *phase*, a specific concept or relationship emerges and a term may be introduced to describe the newly developed concept or relationship. Alternatively, rather than being invented, the concept may be more fully developed or generalized during this phase.
Finally, the *application phase* gives students the opportunity to extend and apply the concept to new situations, augmenting their understanding of the concept (Eberlein et al., 2008). The sequence of questions in POGIL materials are carefully devised to help students progress properly through the phases, to guide them toward appropriate conclusions, and to develop desired process skills, such as problem solving, deductive reasoning, communication, and self-assessment.

In contrast the InDGIM, developed by this researcher from a constructivist model of learning, uses a wide variety of methods that can be used during the study of a topic to help students understand fundamental science concepts within many relevant contexts that relate to students’ lives including proper questioning strategies, science process skills, discrepant events, inductive activities, information gathering, and problem solving (Collette & Chiappetta, 1994). This constructivist model of learning “acknowledges this active role of the learner and the fundamental idea that knowledge cannot be transmitted from a textbook or the mind of the teacher to the mind of the learner” (Stefanich, 2001, p. 29). InDGIM is rooted in the approach that the science curriculum (and in this case specifically chemistry) should focus on developing deep understandings of a few concepts rather than a superficial coverage of many concepts. The goal of the constructivist teacher is to “help the student develop a meaningful, conceptual understanding of science and its value through descriptions, explanations, and predictions which come from the learner…and to help students develop their own explanation for the world around them in ways that incorporate concepts and thinking into their frameworks which emerge as students explain their own experiences and make sense of their world through interaction and problem solving” (Stefanich, 2001, p. 30). Conceptual understanding is at the heart of higher
level thinking, problem solving, and self-regulated learning and comes about through proper questioning strategies, discrepant events, and hands-on, minds-on activities (Stefanich, 2001).

The heart of the InDGIM inquiry process involves questions designed to stimulate thought and action of students (Chiappetta, 1997). From these questions, the teacher should be able to apply science processes to help guide student learning. These skills focus on thinking patterns used to “construct knowledge, represent ideas, and communicate information and help students pose questions, state problems, make observations, classify data, construct inferences, form hypotheses, communicate findings, and conduct experiments” (Chiappetta, 1997, p. 24). This concept has become the framework for the curriculum development of chemistry involving InDGIM.

To help enhance these skills the next phase is the use of a discrepant event. This is similar to the exploration phase of the POGIL process where students attempt to pull out an understanding of the demonstration. A discrepant event puzzled students, causing them to wonder why the event occurred as it did. It is a valuable way to increase motivation and make the invisible conceptions of the students more visible, for students are more inclined to “vocalize their interpretation/hypothesis of what they are seeing to the teacher or classmates” (Chiappetta, 1997, p. 25). This empirical-inductive approach, provides students with learning situations in which they can discover a concept or principle and gives students a concrete experience whereby they obtain sensory impressions and data from real objects and events. As a result, the learner can perceive certain stimuli and may be in a better position to make sense of a situation. Empirically obtained information can be acted upon cognitively by the student and organized in the
mind, where patterns may be discovered that are meaningful to the learner. The teacher helps bring into the discussion the appropriate terminology for naming the principle and defining it. (Chiappetta, 1997, p. 25)

Like POGIL this inductive step allows students to enter the concept formation phase of learning before getting involved in some form of deductive activity.

In POGIL activities the students follow processes with specific roles (such as manager, recorder, reflector, technician, and presenter), steps, and reports that, according to Kussmaul (2011) support metacognition. Using the groups allows the instructor to become a facilitator for learning and to provide directed attention to students and groups on an as-needed basis. This is an approach that is built on the foundational work in the areas of cognitive development, cooperative learning, and instructional design. This strategy moves away from a teacher-centered approach to teaching, and towards a student-centered learning method which is the favored pedagogy among science teachers (Dickson, 2013).

In InDGIM, activities are used to gather information in an exploration, formation, or application phase of the learning cycle. This might require students to read and answer questions during an analysis of a case study or search the Internet for information. It is from this point where students are asked to problem solve, which can make learning more meaningful and relevant (Chiappetta, 1997). Problem solving is often synonymous with inquiry and science process reasoning skills (Helgeson, 1994), both of which occur during the InDGIM and POGIL. During activities, following the InDGIM approach, students raise questions, plan procedures, collect information, and form conclusions either in small or large non-prescribed groups. Using
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

this approach, during an inquiry-based activity, assures both positive interdependence and individual accountability (Coburn, 2000).

Along with the phases of the learning cycle, both methodologies are structured around inquiry. There are four categories of inquiry that include “conceptual structures and cognitive processes that are used during scientific reasoning, epistemic frameworks used when scientific knowledge is developed and evaluated, social interactions that shape how knowledge is communicated, represented, argued and debated, and…procedural which describes the methods, asking scientifically oriented questions, designing experiments, executing procedures, and creating data representations” (Furtak et al., 2012, p. 305).

POGIL and InDGIM use an orienting framework of conceptual knowledge in that both utilize a student’s prior knowledge to assist with learning, but they also expect the student to develop “sophisticated understandings” (Furtak et al., 2012, p. 305) as a result of that instruction. When students examine and evaluate the quality of evidence, collected during a lab, they must interpret that evidence to develop an explanation for the phenomena being studied. This is the epistemic domain of inquiry. Both POGIL and InDGIM require students to go through this process, but with the InDGIM, a bridge is provided to the students so that they learn that their own process of collecting, evaluating, and interpreting evidence is similar to the practice of real scientists and subject to change in the light of new evidence or new interpretations of past evidence. This is a large part of inquiry (Furtak et al., 2012).

The next phase consists of the collaborative and communicative processes by which scientific knowledge is constructed, or the process by which students participate in the scientific process, the social domain. POGIL and InDGIM utilize this domain quite heavily by having the
students involved with communicating scientific ideas and understandings through collaborative groups where the students are able to make “public their ideas through argument, modeling, and other modes of representation to help them learn to examine and evaluate their developing understanding of science” (Furtak et al., 2012, p. 305).

The final phase of the process, the procedural phase, requires students to manipulate materials, collect their own data, and engage in the process of evaluating the data that they collected. This makes inquiry-based teaching more than just a hands-on situation. This is the phase the InDGIM appears to emphasize more than POGIL with the pinnacle of the process being students actually designing and implementing their own labs from start to finish.

Although InDGIM is based on many conceptual aspects of the learning cycle and inquiry-based learning, there is no current research available that compares this approach to any other inquiry-based models when it comes to teaching science (since it is an approach that was developed and instituted by this researcher for the past 20 years). There are no data present to indicate that one model of inquiry, POGIL or InDGIM, is better than the other when it comes to learning, as measured by increased test scores, which is why the results of this research would be important. Saunders-Stewart et al, (2012) surmised, that those programs that had activities through which a student’s learning process skills can be generalized across subject domains are stronger inquiry-based programs. But from a pedagogical position what must be provided with these programs?

Zachos, Hick, Doane, and Sargent (2000) developed and tested, using least square models, 29 different Scientific Inquiry Capabilities or learning traits that high school students must illustrate in order to be considered successful with discovery learning. The scales were
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

“ordered, to represent increasing competence with regard to some aspect of knowledge, skill or disposition related to scientific inquiry” (Zachos, Hick, Doane, & Sargent, 2000, p. 957) and were based on theories generated from multiple educational theorists such as Piaget, Klopfer, Campbell, Schwab, and Inhelder. The higher the capability score, the better the student was able to comprehend and perform a discovery learning or higher inquiry-based task. Their analysis supported that students who were able to search for a necessary underlining principle, had proportionality reasoning, were able to coordinate theory with evidence, were able to formulate composite variables, identify sources of errors in taking measurements, and record clear observations, showed significant gains in the discovery tasks (Zachos et al., 2000). If the POGIL approach shows to have a greater impact on student learning than InDGIM, it could be surmised that POGIL is constructed in a manner that either allows for a stronger development of these characteristics or allows students the opportunity to illustrate these inquiry traits more than InDGIM.

The results of this research could be used to evaluate changes that may need to take place in the chemistry curriculum at Falls Fallow High School (a pseudonym for the site where this researcher works) so that students can learn using the best educational approach possible.

Purpose of the Study

Although some inquiry pedagogical methodologies appear to be stronger than others for use in the high school or college classroom, how do they compare to classes where the instructor is already using different levels of inquiry to teach his/her students? Would the POGIL method, which is an established inquiry methodology, outperform other inquiry methods not discussed in the current literature, such as those developed and implemented by individual instructors? The
purpose of this research was to quantitatively compare two inquiry-based programs, the POGIL method, with one that has been created by a small group of teachers, labeled the InDGIM.

**Research Questions**

This study addresses one overarching question. Will there be a difference between the outcomes of students exposed to the POGIL process, where students work cooperatively in self-managed teams, using carefully designed materials developed through the POGIL Project, and those of students exposed to the non-POGIL inquiry process or InDGIM? The measured outcomes will include scores on the Group Assessment of Logical Thinking (GALT) test and the Particulate Nature of Matter Assessment (ParNoMA2).

**Research question 1:** What measurable differences does the use of POGIL, a student-centered cooperative learning instructional model used in the teaching of chemistry, have on high school students understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared with InDGIM, the currently used inquiry-based instructional model?

**Research question 2:** Are there overall measurable differences between female and male high school chemistry students on the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) regardless of the instructional model used to teach chemistry?

**Research question 3:** Is there a difference in the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between females and males taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model?

**Research question 4:** What is the differential pattern of performance in the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between those students falling in the
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

lowest performance quartile (as determined by their Keystone Algebra scores) taught using POGIL, a student-centered cooperative learning instructional model when compare to InDGIM, the currently used inquiry-based instructional model?

Research question 5: Will there be a difference in unit performance, as measured on unit tests, and the final exam, between those students taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model?

Significance of the Study

POGIL was designed to replace the lecture approach in the classroom involving students in discussing the course material, rather than just hearing about it. The POGIL process (developed and offered by an independent non-profit organization), which is a group-learning, research-based instructional strategy, guides students (in teams of three to five) to construct new understanding while they simultaneously develop key process skills, including critical thinking, problem solving, and collaboration (Dickson, 2013). Typically, the teams of learners work on carefully scripted inquiry activities and investigations designed to help them construct their own knowledge, often by modeling the original process/phases of discovery and research: an exploration phase, a concept invention phase, and an application phase (Kussmaul, 2011). In contrast the InDGIM uses a wide variety of methods that can be used during the study of a topic to help students understand fundamental science concepts within many relevant contexts that relate to students’ lives including proper questioning strategies, science process skills, discrepant events, inductive activities, information gathering, and problem solving (Collette & Chiappetta, 1994).
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

The school district, which employs this researcher, has made some major changes that will impact the learning of all students by establishing six goals of educational equity designed to eliminate performance gaps between all students. One of the goals is to make sure that there exists unobstructed entrance into, involvement of and full participation of all learners in all programs and activities within our schools, and that the patterns of interaction between individuals and our school communities are characterized by acceptance, valuing, respect, support, safety and security, such that all students feel challenged to become invested in pursuing excellence without fear of threat, humiliation, danger or disregard (School District of Springfield Township, 2016, para. 1).

To make this happen, the school district must create learning opportunities for every child, regardless of background, gender, abilities, and identified needs, so that he/she is presented with the challenge to reach high standards and is given the requisite pedagogical, social, emotional and psychological supports to achieve these standards of excellence. All educational stakeholders have been charged with accepting responsibility and holding themselves, and others, responsible for every learner having “full access to quality education, qualified teachers, challenging curriculum, full opportunity to learn, and appropriate, sufficient support for learning so they can achieve at excellent levels in academic and other student outcomes” (School District of Springfield Township, 2016, para. 2).

To help achieve these goals, the district has opted to eliminate all self-contained classrooms for students with Individual Educational Programs (IEPs) and have placed these learners in heterogeneous classrooms with multiple teachers. This could potentially present some issues for these students because of the importance of having a science curriculum that is
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

accessible by all learners, regardless of abilities or gender. The problem is that current literature is sparse regarding a strong comparison between different inquiry methodologies and the POGIL method as it applies to logical reasoning, chemistry understanding and the differential performance between genders, as well as struggling learners at the high school level. This study addressed the need for research that compares inquiry methods, embraced and practiced by current science teachers, or InDGiM, to that of the POGIL method, as it applies to the high school curriculum. This study looked at the impact of both approaches on gender as well as the low-performing, or struggling, learners. Similarities and differences between the methods were analyzed and the strongest methodology for use in the high school curriculum, for both genders, and struggling learners, was determined. It is unknown if the results will mirror those of Bilgin (2009), Brown, P. (2010), Brown, S. (2010), Geiger (2010), Kussmaul (2011), or Minderhout and Loertscher (2007) where achievement, as measured by final grades, tended to increase with the implementation of the POGIL method at the collegiate level.

Summary of Design

This study utilized a nonequivalent group design (NEGD) to investigate student achievement in multiple Academic Chemistry courses, taught by the same teacher, at a small suburban high school, Falls Fallows High School. This quasi-experimental approach, without true participant randomization, is modeled from multiple studies that have examined the effectiveness of POGIL such as those carried out by Kaundjwa, A. O. T. (2015); Sen, Yilmaz, and Geban (2015); Barthlow and Watson (2014), and Villagonzalo (2014). All of these studies compared high school student achievement under POGIL methods, versus a traditional, teacher-centered approach to chemistry instruction. Only one study, Barthlow and Watson (2014)
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

included the effects of POGIL on gender. This study had students take the Group Assessment of Logical Thinking Test (GALT) developed by Bunce, VandenPlas, Neiles, and Flens (2010), and the second iteration of the Particulate Nature of Matter (ParNoMA2) test developed by Yezierski and Birk (2006) at the beginning of the course. The same instructor administered the GALT and ParNoMA2 at the end of the course, after students were exposed to the concepts in chemistry, as a posttest. Grades on the assessments were compared between the POGIL and the Non-POGIL groups, and between genders.

Summary of Introduction

Teaching science through science inquiry is the cornerstone of good teaching. Unfortunately, according to Luft, Bell, and Gess-Newsom (2008), many teachers are still striving to build a shared understanding of what science inquiry means, and at a more practical level, what it looks like in the classroom. During inquiry, learners engage in scientifically oriented questions; give priority to evidence in responding to questions; formulate explanations from the evidence; connect explanations to scientific knowledge; and communicate and justify explanations. POGIL and InDGIM are just two possible inquiry methods that can be used by teachers in a secondary setting.

Inquiry, a process important at all grade levels, requires students to engage in higher-level thinking skills of summarizing, analyzing, and evaluating. By providing real-world data in a classroom-friendly format, guided inquiry provides teachers with methods that support students summarizing knowledge, analyzing data, and evaluating their findings. Teachers use inquiry methods to promote learning through student investigation, following the same process used by
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

scientists. By using data sets from working scientists, students focus their efforts on analysis and evaluation.

Both POGIL and InDGIM use student-centered guided inquiry, incorporating a cycle of exploration, concept invention and application, as the bases for materials that students use to guide them to construct new knowledge and develop higher-order thinking skills, but the question posed by this researcher is whether or not one method is statistically stronger than the other method, and will the results vary by gender? This will be determined with a quasi-experimental study using a NEGD in a small, suburban high school.

It should be noted that gender and prior achievement levels could be a possible source of variance in the scores of students on the GALT and ParNoMA2. The average scores for females are lower than their male counterparts on science tests at the secondary and post-secondary level, and it has been suggested that females benefit especially by the use of active pedagogies and when they are allowed to express ideas in words and discussions, whereas males prefer to work independently (Lorenzo, Crouch, & Mazur, 2006). Along with that, students experiencing lower mathematics achievement in 8th grade is a major predictor of their performance in science at the high school level (Morgan et al., 2016). Knowing that the POGIL method incorporates a great degree of group interaction the results of the research might be skewed in favor of one gender or towards lower achieving students.
Chapter Two

Literature Review

When the National Science Foundation first proposed a shift from educator-centered to student-centered teachings in the mid-1990’s, some science educators at the collegiate level in the United States took up the challenge to develop new methodologies to replace the age-old method of teaching by telling (Eberlein et al., 2008). It was from this effort that POGIL was developed. Research on POGIL in the natural sciences has revealed it to be a highly effective and advantageous pedagogy (Mauer, 2014). Its effectiveness has been evaluated against traditional teaching methods, such as lectures, using both qualitative, quantitative and mixed-methods approaches.

Process Oriented Guided Inquiry Learning

POGIL is a research-based, student centered philosophy and science pedagogy in which students work in small groups to engage in guided inquiry using carefully designed materials that direct and guide students to build and rebuild their content knowledge (Farrell, Moog, & Spencer, 1999; Moog & Spencer, 2008). POGIL simultaneously teaches both content and key process skills of science, but the approach can be used in other disciplines as well. POGIL activities focus on core concepts and processes of science as it encourages and fosters a deep understanding of the course material while developing higher-order thinking skills.

POGIL is active and student-centered and is based on the learning cycle (Farrell, Moog, & Spencer, 1999; Hansen, 2006; Moog & Spencer, 2008). POGIL instruction utilizes carefully written guided inquiry student learning documents available at http://www.pogil.org, with each document offered as a free download. Embedded in the POGIL student learning documents are models designed to help students visualize abstract concepts and submicroscopic phenomena as
they relate to chemistry. POGIL student learning documents are thoughtfully and intentionally developed with the purpose of each student experiencing the learning cycle.

Most of the literature regarding the use of the POGIL method illustrates one of four major themes: 1) the effectiveness of POGIL; 2) the challenges of implementing the POGIL method; 3) strategies for overcoming these challenges; and 4) how students’ attitudes regarding science can change with the POGIL process. This literature review will look at each theme.

**Effectiveness of POGIL**

**Colleges and universities.**

Unanimously, whether being adopted in science (P. Brown, 2010; S. Brown, 2010; Hein, 2012; Jin & Bierma, 2011; Meyers, Monypenny, & Trevathan, 2012; Radhi, 2013), business (Hale & Mullen, 2009) or in art classes (Mitchell & Hiatt, 2010), as a part of a lecture (Meeks, 2015; Meyers et al., 2012), or as a stand-alone session (Hein, 2012; Mitchell & Hiatt, 2010), POGIL, a form of guided inquiry learning, according to the vast majority of the POGIL-specific articles reviewed herein, has been reported to have a very encouraging impact on students’ learning. The studies reported a measurable increase in student engagement (Bilgin, 2009; Jin & Bierma, 2011; Mitchell & Hiatt, 2010; Moore, Black, Glackin, Ruppel, & Wasston, 2015; Williamson, Metha, Willison, & Pyke, 2013), attention and comprehension in class (Meyers et al., 2012; Vacek, 2011); knowledge retention (Bilgin, 2009; Meyers et al., 2012; Vaughan, 2010); process skills development (Straumanis & Simons, 2008) and possibly achievement (P. Brown, 2010; Degale & Boisselle, 2015; Hale & Mullen, 2009; Hein, 2012; Ucar & Trundle, 2011; Vacek, 2011), as compared with the traditional lecture-based method of instruction. POGIL is further suggested to boost the rapport between the students, and between students and
instructors, thus causing them, including the high-ability students, to feel more confident in class (Meyers et al., 2012).

POGIL effectiveness has been measured at different types of institutions in a variety of collegiate courses, with chemistry being the most common (Moog & Spencer, 2008). The results of these studies consistently show that for the chemistry courses taught using the POGIL method compared to lecture instruction, student retention improved, mastery of the content increased, preference for the POGIL methodology was higher, and there was significant lowering of absenteeism (Eberlein et al., 2008). In addition, POGIL produced greater student engagement and more higher-order thinking (Mohamend, 2008).

The retention data was based on the comparisons of successful students with course grades of A, B, or C to unsuccessful students with course grades of D, F, or W (course withdrawal). In many cases, the percentage of successful collegiate students significantly increased (Bilgin, 2009; S. Brown, 2010; Chase, Pakhira, & Stains, 2013; Conway, 2014; Daubenmire, Bunce, Draus, Frazier, Gessell & van Opstal, 2015; Kode & Cherukuri, 2014; Murphy, Picione, & Holme, 2010; Soltis, Verlinden, Kruger, Carroll & Trumbo, 2015). Content mastery was measured by common final exams given to POGIL and lecture sections.

In the study by S. Brown (2010), what had been a B-C grade distribution in the course became an A-B grade distribution after POGIL implementation. The average scores on summative exams shifted from 86% of students scoring in the B-C grade range in the non-POGIL sections to 82% of students scoring in the A-B grade range for POGIL students. It was reported that this shift was remarkable when considering how similar these groups were with regard to aptitude for the subject matter being taught and tested (S. Brown, 2010). Of interest is
the fact that instructors in the various sections of the courses in this study changed their summative assessments each year (S. Brown, 2010). The summative assessment documents were analyzed to determine the Bloom’s taxonomy level of each question and it was found that the examinations given in the POGIL sections consisted of fewer questions from Bloom’s level 1 (knowledge) and more questions from level 2 (application) (S. Brown, 2010). According to S. Brown, the students in the POGIL sections took more difficult exams than the exams given to the students in the non-POGIL sections of the course and yet, were earning better scores (2010). An extensive study of POGIL implementation in organic chemistry at seven tertiary institutions compared final exam results within each institution for POGIL and lecture sections (Straumanis & Simons, 2008). The percentage of successful students with grades of A, B, or C in the POGIL sections was significantly higher than in lecture sections. This study illustrated that the POGIL methodology not only reduced attrition without lowering standards, but it also improved student learning and promoted the development of key process skills such as critical thinking, teamwork and self-assessment.

The same pattern of effectiveness has emerged in multiple courses other than chemistry. In a study by P. Brown (2010) involving anatomy and physiology, the mean score on the final exam improved from 68% to 88%, with a decrease in the percentage of students earning a grade of D or F in the course, decreasing by half in the first two semesters and then dropping to 0% in the third semester. A similar finding was reported in a study by Simonson and Shadle (2013) of biomechanics where the number of A’s earned as course grades increased by 10%, the number of B’s by 13%, and the fraction of students earning a C was reduced from 36.5% to 18.6%, “indicating that POGIL may have benefited the mid-level students the most” (p. 61). In
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

engineering, grades increased significantly (Douglas & Chiu, 2012); in nursing the participating POGIL group mean on a national standardized test were higher than the non-participating group (Roller, 2015); in marketing courses students in the POGIL classroom displayed better attendance and better command of the information as assessed in class (Hale & Mullen, 2009); in finance courses there was an increase in overall scores for the POGIL method as compared to lecture (Maurer, 2014); and in environmental health the “POGIL units appeared to be associated with a small increase in content mastery, which was an unexpected result because less class time was devoted to this content” (Jin & Bierma, 2013, p. 25).

Hien (2012), Vacek (2011), Straumanis and Simons (2008), Shroeder and Greenbowe (2008), and Perry and Wight (2008) used a similar quasi-experimental pre- and posttest mixed-methods approach when studying POGIL and lecture-based curriculums, but came up with different conclusions. Hien (2012), Vacek (2011), Straumanis and Simons (2008), Shroeder and Greenbowe (2008) concluded that performance on post-POGIL tests pertaining to particular content improved compared to the students exposed to traditional lecture strategies. Perry and Wight’s results (2008) demonstrated that there was no statistical difference in the overall final performance between the comparison or lecture group, and the class that infused POGIL activities throughout the course. An item analysis for each question did yield insight into the overall learning process for the course.

Bailey, Minderhout, and Loertscher (2012) reported on biochemistry students’ understanding of concepts from general chemistry and biology using a diagnostic pretest and posttest at the beginning and the end of the semester. Their instrument “measured student understanding of seven different concept areas from prerequisite courses that are related to core
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

crosses in biochemistry” (p. 5). The mean score showed a statistically significant increase from 9.1 points to 12.5 points thus lending support to the idea that this pedagogical methodology is more beneficial for students than full time lectures.

Vanags, Pammer, & Brinker (2013) compared 354 first-year undergraduate psychology students’ learning in physiological psychology using different methods to isolate the most important aspects of the POGIL approach. The results illustrated that the comparison and the POGIL groups showed no improvement between pretest and posttest, but the POGIL group, showed the smallest drop in knowledge two weeks after the posttest of the class when they were given another assessment (Vanags et al., 2013). This was considered statistically valid (p = 0.05).

Simonson and Shadle’s (2013) mixed methods study found that POGIL was effective in improving instructor-student interactions. It was determined that the students and the instructor interact more frequently through “instructor observation, answering questions, and providing feedback than in regular lecture classes” (Simonson & Shadle, 2013, p. 60). This would allow for the instructor’s assessment of the students’ grasp of the material via the questions the students ask and how they manipulate the information as they move through the application questions (Simonson & Shadle, 2013). This type of interaction enhances the instructor’s ability to subjectively evaluate the students’ grasp of the material, as well as other skills, such as cooperation and group leadership.

High schools.

Unfortunately, few empirical studies are available that have examined science achievement and in-depth studies of complex science skills, as POGIL provides, at the high school level. A quasi-experimental study performed on 318 high school students in Georgia
determined that there were statistical gains between the pre- and posttest on a chemistry concept test, the Particulate Nature of Matter Assessment (ParNoMA2), of the group experiencing the POGIL method as compared to the non-POGIL method, labeled as traditional lecture and lab-based methodology. Results indicated that the gains made by students in the experimental group were related to the teaching method (Barthlow & Watson, 2014). There were gender and race gains as well. Barthlow and Watson (2014) found that male and female students experiencing the POGIL process posted better posttest scores than the comparison or lecture group and that the African-Americans and Hispanic-Americans in the POGIL group exhibited achievement gains consistent with European-Americans and Asian-Americans. This means that POGIL as a teaching strategy does not appear to favor gender or race.

Similar results using the ParNoMA2 were recorded by Villagonzalo (2014) but with a much smaller sample (n = 41). There was little difference in the overall performance of the two comparison groups on the pretest but there was a significant difference in the mean posttest scores between the students receiving the traditional teaching method as compared to the POGIL method (Villagonzalo, 2014). This suggests that the use of POGIL instruction is significantly better in enhancing students’ academic performance compared with the traditional teaching method. Results from an ANCOVA indicated that over half of the differences in scores could be attributed to the POGIL instruction (Villagonzalo, 2014).

Sen et al. (2015) conducted a study whose purpose was to investigate the effect of the POGIL method compared to traditionally designed chemistry instruction (teacher-center approach) on 11th grade students’ self-regulated learning skills. The study, involving 115 students, used a nonequivalent comparison group design, with two comparison groups and two
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

experimental groups. Students were given two questionnaires at the end of the course, the Achievement Goal, and the Motivated Strategies for Learning, in order to assess the growth of their self-regulated learning skills. Results revealed that POGIL improved students’ “mastery approach, task value, control of learning beliefs, self-efficacy for learning and performance, critical thinking, help seeking, peer learning, metacognitive self-regulation, effort regulation, and time management skills” (Sen et al., 2015, p. 61).

Fishback and Daniel (2011) collected data over a course of two school years in a number of high school college preparatory chemistry classes and reported that the students subjected to the POGIL process posted higher averages on twenty-six out of the thirty-three assessments given in the year long course. The findings suggest that students performed better on 78% of the assessments, although this was not statistically verified (Fishback & Daniel, 2011). Students exposed to the POGIL method scored the same or higher in five of the six subcategories on the Core Subject Theme Test with “significant increases in the Chemical Reactions, Conservation of Matter, and Stoichiometry subcategories” (p. 30). The overall scale score increased between POGIL and lecture groups by 7.1 points, but, it is unclear if this is statistically significant.

Challenges and Strategies for Implementing POGIL

In spite of the widely reported positive impact on students’ skills and performance the implementation of the POGIL methodology is not without its problems. The issues include student acceptance, team building, and proper teacher implementation.

Many studies illustrated that students were initially uncertain about an unfamiliar teaching and learning style (Al Awadh, 2012; Douglas & Chiu, 2012; Kussmal, 2015; Mitchell & Hiatt, 2010; Mulligan, 2014). Douglas and Chiu, (2012) noted that students normally tend to
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

resist active learning because they expect to be taught by an expert whom they assume is the main source of knowledge. Some students “felt uncomfortable not being told the answers to the worksheet questions and suggested that the instructor offer the answers to all questions so they knew they were getting them correct” (Douglas & Chiu, 2012, p. 255). This overestimation of the instructor’s role in the learning process, the two authors asserted, contradicts the nature of student-centered learning. Other students have reported skepticism for this methodology being used for an entire course (Mitchell & Hiatt, 2010). A study by Al Awadh (2012) illustrated that the POGIL method may be a little too abstract or removed from textbook learning for students which is why they may hesitate to embrace this method. According to Dickson (2013), this process is “known to push students into higher-level critical thinking or thinking ‘outside of the box’ for which there may be no fixed answer...all of which makes some students, particularly those who were performing extremely well under the previous system, uncomfortable” (p. 280). Students have commented that there is more pressure on them since the teachers, using this constructivist approach, are requiring them to do more tasks, but not showing the connections with the texts, which they still rely on outside of school (Dickson, 2013). Allowing student feedback throughout the process will help alleviate these uncomfortable feelings (Douglas & Chiu, 2012; Kussmal, 2015; Minderhout & Loertscher, 2007; Mitchell & Hiatt, 2010; Mulligan, 20104).

Student grouping is also a challenge. Mitchell and Hiatt (2010) stated that grouping students can be difficult when students are absent because this creates a lack of cohesiveness over a short period of time. Luxford, Crowder, and Bretz (2012) illustrated that when some students working in groups have some prior knowledge of the activity, the group members rely
more heavily on the students who seemed more confident in their knowledge and whom
provided the relevant information to other less confident students. Another issue was that when
students in the group rely too heavily on one student, they increase the likelihood that all
students might end up with all the same incorrect answers (Johnson, 2011). To combat this
concern changing the groups on a regular basis forces students to work more cooperatively
(Luxford, Crowder, & Bretz, 2011; Meyers et al., 2012) and peer evaluations should help
eliminate reliance on only one student.

Several factors complicate the use of POGIL in large lecture settings as was determined
by Yezierski, Bauer, Hunnicutt, Hanson, Amaral, and Schneider (2008). One is the physical
formation of student teams; it is typically difficult to prearrange team membership, and students
usually cannot always move chairs or desks to facilitate team communication. Second, it can be
difficult to disseminate evidence to be used in a POGIL exercise, particularly if the evidence
comes from laboratory-based investigations by the students. Finally, it is difficult for the
instructor to monitor and interact with students as part of the POGIL process. However,
recommendations have been developed for the use of POGIL in large classes (Yezierski et al.,
2008). Some of the recommendations most relevant to incorporating POGIL into a course
include the use of clickers, keeping POGIL activities relatively short, and distributing materials
electronically or graphically in class rather than on paper (Jin & Bierma, 2013).

Other questions arise such as how will an instructor know whether students are benefiting
from POGIL, or how can an instructor get the most out of the initial and subsequent
implementations of POGIL? Assessment must include a feedback loop that allows the instructor
to identify the strengths and areas to focus on for improvement. This assessment information can
come from self-analyses, student assessments, and peer assessment from other instructors. The different assessments can be used for three different levels of analysis: a specific activity (a guided inquiry worksheet, a particular demonstration, a lab experiment); a general component of the course (lab, group work, lecture, the text, etc.); and the course in general (Cole & Bauer, 2008). The literature has illustrated that using the American Chemical Society standardized final exams, GALT, and the ParNoMA2 offer convenient measurement of content mastery and provide a basis for more rigorous statistical studies comparing POGIL with lecture instruction (Geiger, 2010; Lewis & Lewis, 2005; Perry & Wright, 2008).

Implementation issues need to be addressed in the transition of POGIL from colleges to high schools. It was determined by Trout, Padwa and Hanson (2009) that the issues include the range of topics covered, the rigor of the materials, and the perceived difficulties associated with the use of cooperative learning in high schools. High school teachers have found that they need to alter college level materials or write new materials in order to effectively use POGIL in their classes (Trout et al., 2009). The work load for the instructor in the preparatory phase of POGIL is intensive (Vacek, 2011). But Hale and Mullen (2009) contend that once the activities are prepared, instructors will need to spend less preparatory time and will be able to give more consistent instructions. Attending a professional development session also enhances the confidence of teachers who use POGIL.

**Perceptions and Attitudes Resulting from POGIL Implementation**

Perceptions can have an impact on the success of a newly implemented program. Hinde and Kovac (2001) found that the students’ response to the attempt to introduce active inquiry learning into their physical chemistry courses were positive, stating that the “students felt that the
active-learning sessions were valuable learning experiences...perhaps they felt more positive about physical chemistry” (p. 97). Shroeder and Greenbowe (2008) found that at the conclusion of the term, when surveyed, many students thought the POGIL approach made the class easier than what they initially expected. This result was replicated two years later in P. Brown (2010) and Mitchell and Hiatt’s (2010) studies which showed that student satisfaction with the method was high, and that most students perceived the value of the POGIL approach. Students found working in group environments was made more enjoyable due to the fact that class assessments were based on individual work, thus removing the pressure of group interactions and that the active nature of learning the concepts and skills was preferable to other methods (Mitchell & Hiatt, 2010).

Studies from Bailey et al., (2012) and by Jin and Bierma (2013) used student affective surveys at the end of each course to collect data on the perception of their students. In the study by Bailey et al., (2012) 83.5% of students strongly agreed that the POGIL activities helped them in the course and 86% of the responders felt that it should be used in future classes. These findings are supported with data from the Jin and Bierma’s (2013) study with 84% of students strongly agreeing that the POGIL modules presented the material in a more interesting format than that of the traditional modules and 83% agreed that the POGIL methodology contributed to a greater understanding of the material.

Students’ positive free responses regarding the benefits of the POGIL approach included statements that the approach kept them focused and broke up the time between lectures and “it made for an easy, unintimidating environment to ask questions” (Bailey et al., 2012, p. 5). However, this has not been the case with all students. Some had negative responses which
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

suggested that the POGIL activities were too short, some groups finishing before others. Some students felt that the whole-group project-based team learning was “highly unnecessary and that the instructor should teach instead of expecting the students to learn on their own” (Bailey et al., 2012, p. 5). Clearly not all students perceived this approach to be beneficial.

In the study by Vanags et al., (2013) POGIL students’ self-assessment of knowledge was consistent with the comparison group but the POGIL students had reduced confidence at the posttest two weeks later even though they actually scored higher. Finally, with over 1000 surveys collected from seven institutions, a study by Straumanis and Simons (2008) illustrated that only 8% of all students had negative feelings about the POGIL process.

POGIL and Gender

There is limited information about differences in the performance of genders under the POGIL process. Barthlow and Watson (2014) considered the learning by gender and determined if there was a difference in the “achievement between male and female students taught using POGIL methods and materials to teach physical and chemical changes in matter related to particle theory in secondary chemistry when compared to with traditional chemistry pedagogy” (p. 251). The authors concluded that the main effect of gender was not significantly related to posttest scores with females posting a slightly lower estimated marginal mean than their male counterparts, but the difference was not considered to be significant (Barthlow & Watson, 2014). Thus, any achievement on the posttest scores were not due to gender.

The study from Lorenzo et al., (2006) demonstrated that interactive engagement, which is typical for POGIL instruction, effectively reduced the gender gap in performance in physics. “Although both genders benefit and achieve similar high normalized gains, females improve
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

their performance most, and overcome a considerable pre-instruction gender disparity” (p. 121).
The investigators attributed the observed reduction of the gender gap to the instructional
approach that included peer instruction, tutorial, and cooperative quantitative problem solving
activities which allowed students to interact and explain their ideas during both lecture and lab
sections and receive frequent feedback on their understanding (Lorenzo et al., 2006). The teacher
approach, which alternated between structured teaching and peer discussion, emphasized
conceptual reasoning, promoted collaboration among peers, and created a less competitive
classroom culture (all characteristics found in the POGIL methodology) helped reduce the
gender gap in the understanding of physics (Lorenzo et al., 2006).

Summary

All of the literature cited shows the significant gains that the POGIL approach fostered
both in knowledge and positive perceptions regarding content. POGIL has been found to be
effective in promoting students’ understanding, engagement, retention, and performance. Some
issues that have resulted from implementation of this program have included student resistance
to both the new change and the instructor-selected grouping, reliance on educationally stronger
students and potentially longer preparatory work by instructors. Management strategies might
include provision for feedback by students, peer evaluations and reflective/response teaching.

It would appear that many studies have been performed on the effectiveness of the
inquiry style of teaching in general, and the POGIL process specifically. Researchers have
examined a variety of different components to assess effectiveness, such as the scores students
earned on standardized exams, the test scores students earned in the classroom, the types of
questions posed by the teachers and students, the retention level of the students, and the attitudes
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

held by both educators and students about the POGIL methodology. It must be stated that, almost all the studies compared either lecture-based courses in college with the POGIL approach, with only a small fraction of studies comparing the POGIL process to traditional teacher-center courses at the high school level.

Based on a comprehensive search of the literature, there does not seem to be any quantitative studies that compare POGIL to another guided inquiry approach, hence the significance of this current study. POGIL is a well-established curriculum that has been quantitatively and qualitatively researched at many different levels. InDGIM is an independently developed (constructivist) curriculum that has been implemented and modified over the past 15 years without any data to illustrate that it is more or less effective than any other inquiry approach. This study filled the need for research that compares current inquiry methods embraced and practiced by chemistry teachers at Falls Fallow High School, InDGIM, to that of the POGIL method as it applies to the high school curriculum.
Chapter Three

Methods

There is little evidence comparing the success of POGIL over other inquiry-based models for teaching science, and none comparing POGIL to InDGIM. Thus, the intent of this study was to compare the POGIL approach with another inquiry-based approach. The comparison approach, INDGIM, includes demonstrations of discrepant events, lectures, tutorials, web-based activities, hands-on inquiry-based labs, quizzes, and tests, developed independently by chemistry teachers at Falls Fallow High School. This study compared these two approaches utilizing a nonequivalent comparison group, pretest-posttest design (Campbell & Stanley, 1963) to investigate student achievement in a college preparatory chemistry course. This quasi-experimental approach was modeled from multiple studies that have examined the effectiveness of POGIL such as those carried out by Kaundjwa (2015); Sen, Yilmaz, and Geban (2015); Barthlow and Watson (2014), and Villagonzalo (2014).

The purpose of this research was to determine if there would be statistical differences between students, exposed to POGIL and those exposed to a self-developed guided inquiry approach, InDGIM, and whether these differences varied according to genders or between students performing at different levels of achievement.

Research Questions and Hypotheses

This study addressed one overarching question. Will there be a difference between the measured outcomes of students exposed to the POGIL process, where students work cooperatively in self-managed teams, using carefully designed materials developed through the
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

POGIL Project, and those of students exposed to the non-POGIL inquiry process, or InDGIM?
The measured outcomes included scores on the GALT test and the ParNoMA2.

Research question 1: What measurable differences does the use of POGIL, a student-centered cooperative learning instructional model used in the teaching of chemistry, have on high school students’ understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared with InDGIM, the currently used inquiry-based instructional model?

Research question 2: Are there overall measurable differences between female and male high school chemistry students on the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) regardless of the instructional model used to teach chemistry?

Research question 3: Is there a difference in the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between females and males taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model?

Research question 4: What is the differential pattern of performance on the ParNoMA2 and GALT between those students falling in the lowest quartile of the class (as determined by their scores on the Keystone Algebra exam) taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model?

Research question 5: Will there be a difference in unit performance, as measured on unit tests and the final exam, between those students taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-
Hypotheses

There are four research null hypotheses for this study:

- **Ho1 (RQ1):** After adjusting for pretest differences, there will be no statistically significant differences on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between high school chemistry students taught using POGIL, a student-centered cooperative learning instructional model and students taught using InDGIM, the currently used inquiry-based instructional model.

- **Ho2 (RQ2):** After adjusting for pretest differences, there will be no statistically significant differences on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between female and male students taking a high school chemistry course.

- **Ho3 (RQ3):** After adjusting for pretest differences, there will be no statistically significant difference on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) of female and male students who were taught using POGIL, a student-centered cooperative learning instructional model and female and male students taught using InDGIM, the currently used inquiry-based instructional model.

- **Ho4 (RQ5):** There will be no statistically significant differences on unit assessments and the final exam measuring the understanding of chemistry content of those students taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

To test the first three null hypotheses, differences were analyzed using a 2 x 2 Factorial Analysis of Covariance procedure (ANCOVA). The main effects due to treatment (POGIL vs InDGIM) and gender (male vs female) tested the effects of the first two null hypotheses, Ho₁ and Ho₂ respectively. The third null hypothesis (Ho₃) examined a potential interaction effect between treatment and gender. Pretest performance on the GALT and ParNoMA2 was used as covariates to adjust the posttest scores to account for potential pretest differences between the groups. If the covariates were significant, adjusted mean posttest scores were used to test the main effects and the interaction effect. The fourth null hypothesis (Ho₄) examined the differential effects of the two different strategies (POGIL vs. InDGIM), results of unit tests and the final exam were compared using an Independent Samples t-test.

Research Design

A non-equivalent pretest-posttest comparison group design was used to collect data in this study. All student participants were placed in their respective Academic Chemistry classes based on recommendations, the work of their guidance counselors, and their schedules. In addition, they were recommended to take the course from their previous science teacher and had taken up through Algebra I in their math sequence. There were no selection criteria based on any variables such as gender, race, age, socioeconomic status, grade point average, pre-scholastic assessment test scores (PSATs), Keystone Biology or Algebra scores, or class rank. The volunteer instructor randomly selected which sections were taught using the POGIL method and which ones were taught using the InDGIM approach. There were four Academic chemistry sections involved in the study: two were taught using POGIL and the other two were taught using the InDGIM approach. This design is appropriately used when it is difficult to randomly
assign students, so intact, already established groups of subjects were used (McMillian & Schumacher, 2010).

Students completed both the Group Assessment of Logical Thinking Test (GALT) and the Particulate Nature of Matter Assessment (ParNoMA2) at the beginning of the study and again at the end of the study. The GALT measures a students’ logical thinking using a 22-item questionnaire that “tests both students’ ability to answer a question correctly and select the correct reasoning for the answer” (Daubenmire et al., 2015, p. 75). The 20-item multiple-choice ParNoMA2 targets misconceptions surrounding phases of matter (Yezierski & Birk, 2006). As in studies by Barthlow and Watson (2014), and Villagonzalo (2014), this test was administered to collect data assessing student’s understanding of the chemistry/content development of participants receiving the POGIL and InDGIM treatment. Table 3.1 illustrates how the groups were arranged in a non-equivalent pretest-posttest comparison groups design.

<table>
<thead>
<tr>
<th>Table 3.1.</th>
<th>Nonequivalent Pretest Posttest Comparison Group Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Pretests (Co-variates)</td>
</tr>
<tr>
<td>POGIL</td>
<td>O&lt;sub&gt;GALT&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>O&lt;sub&gt;ParNoMA2&lt;/sub&gt;</td>
</tr>
<tr>
<td>InDGIM</td>
<td>O&lt;sub&gt;GALT&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>O&lt;sub&gt;ParNoMA2&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

In order to test for the possible disparity between different achievement levels in science and treatment conditions, POGIL vs. InDGIM, subjects were dichotomized into two categories of achievers based on categories of proficiency on the Pennsylvania Keystone Exams in Algebra.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

and Biology. Students were grouped by quartiles using the ratings from the Pennsylvania Department of Education as Advanced, Proficient, Basic, and Below Basic. The students in the Basic and Below Basic quartiles were labeled as low performers for the purpose of this research. This is how student achievement was differentiated for the start of this study.

As it was, less than 25% of the students in the Academic Chemistry classes fell within the parameters of the Below Basic categorization by Keystone Algebra scores. Due to this, the participating students were divided up into quartiles within the class so that students in quartile 1 (Q1) had the overall lowest scores on the Keystone exams, of the students in that class, consisting of Below Basic and Basic scores. They were considered Basic for the purposes of this research. Those placed in quartile 2 and 3 were labeled as moderate to high achievers, and those in quartile 4 were labeled as the highest achievers.

The individual module scores for the Pennsylvania Keystone Exams in Biology and Algebra were collected along with unit tests, quizzes (sample Appendix I) and final exam scores (sample Appendix J). According to the Pennsylvania Department of Education (PDE) the Keystone Exams are end-of-course assessments designed to assess proficiency in the subject areas such as Algebra and Biology (Pennsylvania Department of Education, 2014). Across the administration of the exams, the overall decision accuracy ranges from 0.78 to 0.82 and the decision consistency ranged from 0.69 to 0.75. According to the PDE the decision accuracy of the Basic/Proficient cut scores ranged from 0.90 to 0.94 and the decision consistency ranged from 0.86 to 0.91. These results indicate that at least 90% of students meeting or exceeding the Proficient cut score would receive the same classification if their true scores were known. If a parallel test were administered, at least
86 percent or more of students meeting or exceeding the Proficient cut score would be classified in the same way…(T)he exact inter-rater agreement percentages (for Algebra and Biology) ranged from 85 to 100 percent (PDE, 2014, p. 215).

Thus, these scores were used to establish a proficiency level for students taking Academic Chemistry.

**Fidelity.**

The instructor for both comparison groups was Mr. B. (not this researcher) as to eliminate the introduction of the variable of multiple teachers using both the InDGIM and POGIL methods. He used the POGIL method with classes in two Academic Chemistry sections and InDGIM in the other two sections of Academic Chemistry. Also, Mr. B. is a certified POGIL instructor and he has taught InDGIM at Falls Fallow High School since 2012, further enhancing the fidelity of implementation.

Mr. B. used a fidelity checklist as a form of self-evaluation, thus providing an assessment of himself. The checklist evaluation, a Science Teacher Inquiry Rubric or STIR (Appendix G), was developed to serve as a self-assessment tool for elementary schools teachers to understand how they implement the essential features of inquiry into their classroom instruction (Bodzin & Beerer, 2003). This rubric was designed to identify and classify inquiry-based activities for each of five essential features of classroom inquiry and their variations based on the amount of learner self-direction and directions from materials (Bodzin & Beerer, 2003). The continuum describes the instruction of classroom learning environments that range from teacher-centered instruction on one end to student-centered learning on the other. This rubric was validated by having three science educators, with expertise in teaching and learning with inquiry, review and evaluate the
content for “accuracy, importance, and validity of the content, then they provided feedback and suggestions” eventually coming to a unanimously agreed upon product (Bodzin & Beerer, 2003, p. 45). The further analysis of the rubric indicates that it can be used as an effective observation tool for supervisors at multiple levels of science instruction not just at the elementary level (Bodzin & Beerer, 2003). The STIR was used for the InDGIM lessons.

There are no developed rubrics to evaluate a teacher’s implementation of POGIL, in fact, according to the authors of the POGIL Project, there is no single way to implement POGIL, since every classroom implementation has unique characteristics that can influence how and whether particular goals are achieved (POGIL Project, 2014). However, there are a few core characteristics that must be present for a classroom environment to be considered a POGIL implementation, according to the POGIL Project (2014):

- Students are expected to work collaboratively, generally in groups of 3 or 4.
- The activities that the students use are POGIL activities, specifically designed for POGIL implementation.
- The students work on the activity during class time with a facilitator present.
- The dominant mode of instruction is not lecture or instructor-centered; the instructor serves predominately as a facilitator of student learning.
- Students have assigned roles within their groups.
- The activity is designed to be the first introduction to the topic or specific content.
- Students are not expected to have worked on any part of the activity prior to class meeting time.
• Groups are expected to complete all of the Critical Thinking Questions (or equivalently designated questions) during class (in no more than about 40 minutes of actual working time), but they are not expected to work on any of the Exercises or Problems.

A checklist of the guidelines was developed (Appendix H) and used as a self-evaluation by Mr. B. and this researcher when evaluating the lesson. The POGIL Implementation Guidelines were used when Mr. B. was performing a POGIL lesson.

To increase the integrity of the treatment implementation and to establish inter-rater reliability, this instructor, certified in POGIL methodology, stopped by weekly to make sure that Mr. B. was implementing both POGIL and InDGIM correctly. This researcher used the fidelity checklist to evaluate Mr. B.’s implementation of POGIL. Before beginning both the observer and Mr. B. came to agreement about the definitions on the rubric, thus establishing content agreement (validity) of the instrument being used. The STIR checklist was used for those classes where InDGIM was implanted and the POGIL Implementation Guideline checklist was used when POGIL lessons were used. This researcher played no other role than to monitor fidelity and collect data from Mr. B.

The research design used in this study can have multiple threats to internal and external validity; however, there were attempts to identify some of the potential sources of invalidity. It was determined, after the pretest, whether the two groups appeared to be similar on key measures before the study began and a judgment was made about the existence of a threat to selection on prior knowledge differences through the use of statistical analysis. In addition, the researcher attempted to control for attrition by assuring that the respondents continually participated in the
study, keeping the dropout rates low. It should be emphasized that there was a high degree of proximal similarity between these groups and those found in other high school academic chemistry classes, increased the validity of the study (Trochim & Donnelly, 2006). There were also potential threats to external validity with regard to people, places, and times (Trochim & Donnelly, 2006). Sources such as variation in student ability and experiences, differences in classroom setting, or the unique location of the study were potential sources of external invalidity. In addition, although instructor variation was controlled for and kept constant in this study, variations in instructor implementation of the pedagogy was a potential source of external invalidity. However, fidelity of implementation was monitored when using any procedure whether in a controlled research condition or in the typical classroom.

Figure 3.1 represents the overall research design and implementation illustrating those controlled variables (CV), independent variables (IV), and the dependent variables (DV) used in this particular design. The controlled variables are those aspects of the research design that could influence the outcomes. Controlling extraneous/confounding variables is an important part of any experimental design. It’s important to either keep them fixed or eliminated in order to clearly identify the relationship between the independent and dependent variables. For this research only one teacher was involved, which curtailed the influence of different personalities or teaching styles on the outcome. However, the tradeoff was that it narrowed the generalizability of the results to stylistically similar groups of teachers. The content, chemistry, was the same for all groups, and all participants had the same amount of time in the course every day for one complete semester. Finally, all participants were subjected to the exact same evaluation tools
The comparison of two different guided inquiry methods in high school chemistry classes

Figure 3.1. Research design and implementation

throughout the course including the same unit quizzes, tests, and final exam thus eliminating potential variables in testing instrumentation or content emphasis. Controlling confounding variables is important because slight variations in the experimental set-up could strongly affect the outcome being measured (Teddlie & Tashakkori, 2009). This protected the integrity and accuracy of the research.
Settings and Participants.

Settings.

This research was conducted in a small suburban high school, Falls Fallow High School (a pseudonym), located in the suburban community of a major Mid-Atlantic metropolitan city. The setting was selected because of a current relationship with the school, where this researcher is employed, and the availability of both teacher and student participants. This school, which emphasizes a liberal arts education, had roughly 635 students (52% male and 48% female) in grades nine through twelve and was served by 60 full time teachers and three administrators at the time of the study. Seventy-three percent (73%) of the students enrolled in the school were European Americans or White; 18% were African Americans or Black; 5% were Asian American; 3% were Latin American or Hispanic; less than 1% were American Indian, Alaskan Native, Hawaiian Native, or Pacific Islander; while roughly 1% of students claimed mixed racial heritage, and 20% of the population was considered economically disadvantaged (U.S. News and World Report, Education Rankings and Advice, 2015). Eight two percent (82%) of the population was considered proficient in literature, whereas 73% was proficient in mathematics, according to the performance on state exams, and there was a 47% participation rate with regards to Advanced Placement exams and less than a 1% dropout rate (U.S. News and World Report, Education Rankings and Advice, 2015).

Falls Fallow High School is part of a small district, the Falls Fallow School District (FFSD) which includes four schools that serve approximately 2,229 students in Kindergarten through 12th grade; has 166 teachers and 273 other staff members (nurses, counselors, paraprofessionals, support personnel); and a student to teacher ratio of 13.45:1 (National Center
for Educational Statistics, 2015). There were 31 English as a Second Language learners and 468 students, or 21% of the total district population, that were served by both Gifted and Non-gifted Special Education, according to the latest statistics from National Center for Educational Statistics (NCES, 2015). The district spends roughly $19,096 per pupil with 60% of the cost earmarked for instructional expenditures, 11% for student and staff support, 11% for administrative costs, and 18% on operations, food services, and other services while it receives 83% of its yearly budget from local sources (NCES, 2015).

The high school worked on a Block schedule so that a full year course was taught in 90 days. Students met for roughly 90 minute intervals and had four courses on any given day (A, B, C, and D Block). Students were enrolled in an Academic Chemistry course during any of the four blocks and were enrolled in the course during the fall or spring semester.

**Participants.**

**Students.**

The participants were selected from existing intact classrooms, as a sample of convenience. Convenience sampling allows for the selection of participants who meet the criteria and are available to participate in the research. The participating teacher, however, randomly assigned his classes to either the comparison group or the treatment group.

In order to select a somewhat homogeneous group of students across classrooms, the students in this sample were already enrolled in a college preparatory Academic Chemistry course (not a high-ability or low-ability class) which met for 84 minutes per day for 90 days. As stated by Ary, Jacobs, Razavieh, and Sorensen (2006) “in a typical school situation, schedules cannot be disrupted nor classes reorganized to accommodate a research study. In such a case it is
necessary to use groups as they are already organized into classes” (p. 341). The majority of the students were in 11th grade, but there were some 10th and 12th grade students enrolled in the course as well. Based on historical data collected over the past five years, there was a high probability that there would be an approximately equal percentage of males and females in the sample, i.e., average in the past 5 years, 51.1% to 48.9% respectively, however, this was not the case. The gender break down for the study was 63.8% females and 36.2% males. All of the students had taken and passed Environmental Science, Biology, and Algebra I, which are the prerequisites for the course. Differences in grade level were also noted. Some of the students might have taken Physical Science before the Chemistry course (along with Environmental Science and Biology). The actual demographics of the students were provided once the data collection was complete.

*Instructor.*

The teacher who was invited to assist with this study, Mr. B., is credentialed in Pennsylvania to teach Biology, Chemistry, Environmental Science, and Physics; is a certified trainer in the POGIL process; has taught the InDGIM approach to chemistry for three years; and is considered an innovator within the Science Department. He collected, recorded, and reported the data using numerical codes for each student to maintain confidentiality. He agreed to assist with this endeavor with the disclaimer that none of the data collected and analyzed would be used for any evaluative purpose.

The selection of Mr. B. was based on convenience as well as his ability to use inquiry in his classes. What are the characteristics of teachers that use inquiry? According to Colburn (2000) an inquiry-based teacher “possesses certain attitudes and skills to encourage student
success” (p. 43). Through previous observations and discussions with him, as his mentor and Department Coordinator, it was observed that Mr. B. supports inquiry-based instruction in that he uses inquiry and problem-solving lessons, has active student participation, and frequent teacher-student interactions, while he selects and adapts the curriculum to meet the needs of all of his students. These are inquiry-based methodologies (Rhoton, 2001). He uses computers and computer-based simulations as well as hands-on labs to enhance the experiential learning of the students, all part of the InDGIM process. He encourages the use of multiple methods to facilitate student learning from problem-based laboratory exercises to discrepant events as a means of challenging students’ misconceptions.

Colburn (2000) states that to be a successful inquiry-based instructor, “the teacher needs formal operational thinking abilities, knowledge of the subject students are investigating, and some understanding of how students learn (to be able to respond effectively to student statements)” (p. 44). Other characteristics of a successful inquiry-based instructor is one that asks open-ended, or divergent questions; waits a few sections after asking a question, giving students time to think; responding to students by repeating and paraphrasing what they have said without praise or criticism; avoids telling the students what to do; and maintains a disciplined classroom (Colburn, 2000). Based on his professional evaluations by the current administrative staff, which he has shared with this researcher, it was concluded that Mr. had the characteristics of an inquiry-based instructor, receiving proficient and distinguished ratings in establishing a culture for learning and engaging students in learning.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

**Consent and Assent.**

**Consent.**

For the purpose of this study, and informed parental consent form (Appendix A) was utilized. Two weeks prior to the administration of the GALT and ParNoMA2, the informed consent form and a brief summary of the purpose of this study (Appendix B) was electronically mailed to the homes of all eligible Academic Chemistry students in order for parents to have adequate time to receive, read, and respond to the informed consent. The consent forms were handed out to the students in all of the eligible Academic Chemistry classes after this researcher explained the purpose of the study to the students. Students took the forms home to be signed by their parents and/or guardians. In addition to parental consent, student assent was also required.

**Assent.**

Along with receiving parental consent, students also agreed, or assented to participate in the study. In order for students to participate in the electronic assessments, the GALT and the ParNoMA2, students were required to complete and return the assent form (Appendix C).

Once informed consent and student assent was granted the electronic version of the GALT and ParNoMA2 was shared on the password protected Google® drive thus controlling which students were granted access to the tests. Mr. B. was asked to set aside time during the second week of class for students to take the GALT during a 20-minute period of the class. Twenty minutes was also required during a second class period to have the students take the ParNoMA2 on the computer. A short synopsis of the study was provided for the students in the introduction of each of these assessment (Appendices D & E). The students were asked to select a box to acknowledge that they had read and understood the parameters of their participation.
before beginning each pretest. This was repeated at the end of the course as well when students took the posttests.

**Instrumentation**

Two instruments were used to compare the differences between POGIL and InDGIM in the chemistry classroom. Two independently produced instruments were used to collect data, the GALT (logical reasoning) and the ParNoMA2 (understanding of chemistry concepts). Both instruments were administered via computer within the first two weeks of class.

The GALT has been used before and after implementing the POGIL treatment in many studies (Bilodeau and Rhoten, 2011; Bunce et al., 2010; Daubenmire et al., 2015). The 22 question GALT (Appendix D) test, developed by Roadrangka, Yeany and Padila (1983), is a 15-20 minute Piagetian test of logical thinking that, used to measure logical reasoning skills in pre-college (as low as 6th grade) and college level students, focuses on six modes of reasoning, “one of which is concrete operational (conservational reasoning) and five of which are formal operational (proportional reasoning, controlling variables, probabilistic reasoning, correlational reasoning and combinatorial logic)” (Bunce & Hutchinson, 1993, p. 184). The first of these skills (mass/volume conservation) is typically mastered at the concrete operational level, whereas all others correspond to the formal domain. The concept of the conservation of matter and energy is a basic building block to the understanding of chemistry in that it makes it possible to compute the masses of substances involved in physical and chemical processes, which result in the formation of new substances. The amount of these substances depends on the number and types of masses of elements in the reactants, as well as the efficiency of the transformation. The other five modes focus on abstract reasoning which is a key skill for chemistry (Geiger, 2010). Results
on this test are used to determine the operational level of the responder. Students are awarded a point for each correct section and a resulting score between 0-4 is characteristic of concrete thinkers, 5-7 of individuals in a transitional stage, and 8-12 of formal thinkers (Roadrangka et al., 1983). This test has a reported Cronbach alpha reliability coefficient of 0.85 (Roadrangka et al., 1983); a Spearman-Brown prophecy of 0.79 (Bird, 2010); and the internal consistency of each subtest ranged from 0.37 to 0.85 (with only two of the subtests had reliabilities below 0.58) with the item discrimination indices averaging 0.44 and the mean inter-correlation coefficients for the subtests averaging 0.49. This shows that the GALT has reliability and adequately measures six logical operations (Roadrangka, Yeany & Padila, 1983). These measures were tested on multiple student samples and could “provide a means to assess the cognitive development of a large number of students within a single class period” (Roadrangka et. al, 1983, p. 153). In addition, “the influence of a science curriculum on the developmental progress of the skills measured by GALT” (Roadrangka et. al, 1983, p. 153) can be used to assess the level of development.

The GALT has also been demonstrated to predict critical thinking abilities and potential grades assigned by teachers in science and math (Bird, 2010; Bitner, 1991; Bunce & Hutchinson, 1993). From the data collected by Bird (2010), “it is evident that, in terms of final grade, the mode for students operating at a formal level is A, for those at a transitional level is B and for students at a concrete level is C” (p. 543). It was also determined that among the six logical reasoning modes, probabilistic reasoning is the single best predictor of student performance in general chemistry (Bird, 2010). Bunce and Hutchinson (1993) showed that there was a significant correlation between the math SAT scores and the GALT scores (2-tailed significance) thus suggesting that both predict similar critical thinking abilities. Finally, the GALT was used
by Bunce et al. (2010) in a nationwide study to investigate the differences in student learning from both POGIL and non-POGIL general chemistry students.

The second instrument used was the ParNoMA2 (Appendix E). The 20-item multiple-choice Particulate Nature of Matter Assessment targets misconceptions surrounding phases of matter (Yezierski & Birk, 2006). A study by Yezierski and Birk (2006) created a second version of a test on the particulate nature of matter (ParNoMA2) which was administered before and after treatment to measure differences in content development. The ParNoMA2 had a Cronbach alpha of 0.78 when tested on college students (Yezierski & Birk, 2006). The first version of the instrument was used as a pretest and posttest to remediate the misconceptions held by 690 middle school, high school, and college students and has been determined to be valid for testing a student’s understanding of the particular nature of matter (Yezierki, 2003). This second version was later used separately by Yezierski and Birk (2006) with 719 college students; Tang and Abraham (2016) with 170 college students; and Aydeniz and Kotowsk (2012) with 87 middle and high school students to determine if the intervention improved the understanding of chemistry for those students. According to the analysis of the researchers, the test can illustrate if a given treatment can help students construct connections between more scientific concepts than those in a comparison group (Aydeniz & Kotowski, 2012; Tang & Abraham, 2016; and Yezierski & Birk, 2006).

The items on the ParNoMA2 are based solely on published misconceptions. Four of the items are based on the work of Osborne and Cosgrove (1983) and relate to the composition of bubbles in boiling water and particulate descriptions of evaporation and condensation. The item about gas molecules under different pressures relates to the findings of Benson, Whittrock, and
Baur (1993) and incorporates a misconception about pressure changing the size of molecules. The items related to energy, shape, arrangement, structure, and weight of atoms/molecules and phases are based on the findings summarized by Griffiths and Preston (1992) and Garnett, Garnett, and Hackling (1995). The ParNoMA2 was reviewed by three college chemistry instructors and the answers were deemed valid based on their 100% agreement (Yezierki & Birk, 2006). The test, with a Cronbach alpha of 0.83, was piloted with college students, since, according to Yezierki and Birk (2006) it was likely that these students would have the fewest misconceptions regarding the particulate nature of matter. Since the mean of the ParNoMA2 was a 5.78 out of a possible score of 12 (a 48.2%), the instrument did not illustrate a ceiling effect in the pilot sample (Yezierki & Birk, 2006).

Other instruments used were the unit tests (Appendix I) and final exam (Appendix J) given at the end of the course. The unit tests were first developed and modified by this researcher and Mr. B. since their inception over 15 years ago. The final exam was developed by this researcher and another chemistry teacher at Falls Fallow High School within the last six years. These teacher-developed assessments were based on the Academic Standards developed by the Pennsylvania Department of Education. These standards describe what students should know and be able to do in science courses such as chemistry. The standards clarify the targets for instruction and student learning which allowed the teachers to develop curriculum that could meet the needs of local students (PDE, 2015).

In an attempt to reduce bias by the instructor, a general grading rubric was established by the chemistry teachers at Falls Fallow High School. When students were asked the following question: If the climbers, ascending Mt. Everest, carry 10.0 liter tanks with an internal pressure
of 40.00 atm, what will be the volume of the gas when it is released from the tanks at 0.197368 atm?, they were given credit for mathematical answers as long as the problem was set up correctly; the work to get the answer was shown; the answer was mathematically correct; the answer was written in scientific notation; and the work had the correct number of significant digits. Students earned all the points or a portion of the points depending on what was present. Just writing a correct answer would earn the student 20% of the potential points.

**Procedure**

The teacher participant, Mr. B. selected two of his Academic Chemistry classes to serve as controls where he taught chemistry using the InDGIM approach and the remaining two classes received the POGIL approach. To reduce the effects of curricular inconsistencies both comparison groups had identical syllabi, textbooks, management software (class website and grading program) and assessment tools and were taught by the same teacher during equivalent teaching blocks. The only difference was the curricular activities between the InDGIM and the POGIL approach. The POGIL activities were presented in accordance with POGIL guidelines for each class in the experimental group and used in substitution of the inquiry activities currently used in the chemistry curriculum. Both comparison groups were demographically similar.

Two teaching methods were compared; the traditional Falls Fallow High School InDGIM approach versus the POGIL approach. This factor or independent variable was Treatment (teaching method) and the dependent variable was the students’ understanding of chemistry and logic, as measured by the ParNoMA2 and the GALT, respectively. Since it was not possible to randomly assign students to the different comparison groups, intact chemistry classes were
randomly assigned to either the POGIL or InDGIM group. Some degree of random assignment had occurred in that students were placed in their classes without any consideration of this study. Students were enrolled in their classes by their counselors without knowledge of group treatment assignment or non-participating classrooms.

After the assignment of classes to the comparison groups, and consent forms were signed and returned, pretests were given (the GALT and ParNoMA2) to measure the baseline of the students. The academic achievement, represented by grades on these assessments, were collected along with the demographic data and gender of the participants.

Mr. B. was responsible for using the STIR and the POGIL Implementation Guidelines Checklists when he was being observed by this researcher.

**Experimental Phase/Intervention.**

**POGIL philosophy.**

The philosophy of POGIL is that students learn complex concepts best when they are actively engaged in the learning process. This philosophy is expressed in the POGIL objectives which are accomplished during POGIL activities designed to focus on core concepts and processes of science that encourage a deep understanding of course material while developing higher-order thinking skills. The objectives of POGIL (Moog & Spencer, 2008) are to:

- develop process skills in the areas of learning, thinking, and problem solving,
- engage students to take ownership of learning,
- increase student-student and student-instructor interactions,
- improve attitudes toward chemistry and science,
- enhance learning with information technology, and
• develop supporting process skills in teamwork, communication, management, and assessment that are essential for the workplace.

Each POGIL session allows students to engage in conversations as they explain their answers to questions or explore possible answers. In these discussions, students are found to employ higher order thinking skills as they engage in critical thinking, discovery learning, and inquiry (P. Brown, 2010; S. Brown, 2010).

**POGIL resource.**

The online POGIL Implementation Guide, which is a resource developed through the 3-year High School POGIL Initiative project, was used as the main source for intervention. According to the POGIL website (www.pogil.org) this guide has suggestions, tips, video clips, and material files to assist trained teachers with implementing POGIL. Another resource, *POGIL Activities for High School Chemistry* was also used. The students exposed to the POGIL approach followed the activity as outlined in the POGIL Activities for High School Chemistry Manual (Trout, 2012) which consists of: a) a starting question; b) a model developed from experimental data with questions (data may or may not be generated by students); c) a short section of reading with questions to answer; d) a second model followed by more reading and questions; and e) a series of extension questions. This manual contains several features including learning objectives, prerequisites, assessment questions, optional extension questions for differentiated instruction, and teacher tips about facilitating the lessons (Trout, 2012). Because this manual was not an exact fit for the chemistry course (it contains activities for topics such as Thermochemistry, Equilibrium, Acids and Bases, and Oxidation and Reduction that are not
typically covered in the Falls Fallow High School Academic Chemistry course) not all exercises or activities were implemented throughout the course.

**Typical POGIL lesson.**

As part of this intervention, almost all of the class time was spent working in groups of three to four. The membership of the group changed, frequently at first, and less frequently as the semester progressed. Each day, each member of the group was assigned one of four roles: manager (ensuring that members are fulfilling their role, the assigned tasks are being accomplished on time, and that all members of the group are participating in the activities), recorder (recording names and roles of the group members and the important aspects of group discussions, observations, insights, etc.), spokesperson/presenter (presents oral reports to the class for the group, responsible for asking questions and getting feedback from the teacher), and strategy analyst/reflectror (observing and commenting on the group dynamics and behavior with respect to the learning process) (Moog & Spencer, 2008).

In a POGIL classroom, a teacher is a guide in the process of student learning, in developing student skills, and in developing student understanding (POGIL Project, 2014). The instructor was asked to follow the POGIL protocol where he created the learning environment by developing and explaining the lesson; determining the objectives; defining the expected behaviors and criteria for success; and establishing the organization, such as goals, teams, and time structure. According to the POGIL Project, during the class, the instructor monitored the progress of groups and responds to questions, but he did not answer these questions directly, instead he guided the students to answers by asking questions that lead them in the correct direction. The instructor asked questions of groups or individuals to check understanding or
made sure that the group was working as a team. The instructor provided closure to the lesson by asking team members to report answers, summarize the major points, and to explain strategies, actions, and results of the team (POGIL Project, 2014).

Typically, there was a five-minute review given at the beginning of the each class based on the material presented in the previous class meeting. The POGIL Project (2012) states that this review assessment has multiple purposes:

- to give the instructor some immediate feedback about how well a specific concept was learned,
- to reinforce the concept in the student’s mind,
- to encourage the student to attempt and complete the exercises and problems before the review assessment is given,
- to partition the course material into small, manageable sections,
- to encourage the students to engage with the course content as it is being introduced, rather than the night before a test,
- to develop good study habits.

After the review assessment was administered, a POGIL activity was implemented. Every POGIL activity has three components including model development; critical thinking questions; and exercises or problems. The model, which can be a figure, equation, a table, prose, or any combination, provided the basis to develop an important concept (POGIL Project, 2014). Answers to critical thinking questions reveal fundamental relationships or concepts inherent in the model. The questions were written to lead the student to make inferences and to draw
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

accurate conclusions. Finally, the problems were designed to give the students practice in applying the concepts developed in the activity and integrating those concepts to solve problems (POGIL Project, 2014).

**Comparison group (InDGIM).**

The current curriculum for the chemistry course has each unit beginning with demonstrations or a discrepant event that poses a problem, followed by lab activities (requiring manipulation of variables in a lab or through a computer simulation, collecting and analyzing data, and reaching conclusions) that engage students and encourage student-teacher interactions; with teacher discussion as needed. Students worked together in groups of two to four to collect and analyze the data, but without the specific roles found in the POGIL process. The instructor guided the students during the analysis phase in this process, providing mathematical manipulation and modeling of chemical behavior (this is different than POGIL where students are asked to wrestle with the chemical model before performing any lab). During the class, the instructor monitored the progress of groups and responded to questions, and may or may not have answered these questions directly. The instructor guided the students to answers by asking questions that lead them in the correct direction, which is a major aspect of inquiry. The instructor asked questions of groups or individuals to check understanding or to make sure that the group was working as a team. The instructor provided the students with problems that involved the model and mathematics studied.

The curriculum itself encourages the use of multiple methods to facilitate student learning from problem-based laboratory exercises to discrepant events as a means of challenging
students’ misconceptions. According to Rhoton (2001) this can be classified as an inquiry approach since it engages and encourages the student to become active learners.

Each chemistry section, regardless of teaching methodology, used the same textbook, were assigned the same homework problems, took the same quizzes and exams, and were given the same final exam at the end of the course. The only differences between the sections was whether they were taught chemistry using the POGIL method or the InDGIM and the time of day they had chemistry.

**Data Analysis.**

Comparisons were made between the treatment groups on posttest scores on the ParNoMA2 and the GALT. The posttest results of the groups were compared using an Analysis of Covariance (ANCOVA) similar to Kaundjwa (2015); Barthlow and Watson (2014); Villagonzalo (2014); and Vanags et al., (2013). This method was chosen to compare the nonequivalent comparison (Non-POGIL or InDGIM) and treatment (POGIL) groups in order to control for pretest differences. The current study used the same procedure to control for the possible confounding effects of pretest score differences on posttest means. A randomized block design would have been a more desirable option in order to minimize specific threats to internal and external threats to validity; however, for this particular study it was not possible to randomly assign students to various teaching methods groups. The study was conducted during the normal scheduled teaching sessions of the school day, therefore randomly assigning students to teaching groups could have negatively affected the normal teaching routine at Falls Fallow High School. In addition, differences between males and females were also be explored. A 2 x 2 factorial Analysis of Covariance (ANCOVA) was used to analyze the data from the ParNoMA2 and the
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

GALT to test the three major hypotheses, the main effects due to Treatment, and Gender, as well as the interaction effect of Treatment by Gender. The impact on logical reasoning (GALT) and the understanding of chemical concepts (ParNoMA2) due to the POGIL or InDGIM treatments or due to gender was compared (Ho1 and Ho2). These main effects along with the interaction effect was assessed (Ho3). Pretest performance from the GALT and the ParNoMA2 were used as covariates in order to control for the potential inequality between the groups on posttest GALT and ParNoMA2 means. The ANCOVA tested whether the factors would still impact the posttest scores after the influence of the covariate (pretest scores on the GALT and ParNoMA2) were removed. Adjusted posttest mean scores were analyzed only if the covariates were statistically significant. Each null hypothesis was tested using $p \leq 0.05$ as the critical cutoff to evaluate statistical significance.

In addition, the assumption of the homogeneity of the regression slopes and the assumption of existence of a linear relationship between the covariate (pretest scores) and the dependent variables (posttest scores) were assessed. Like Barthlow and Watson (2014) a Levene’s test of equality of variance was performed to determine the level of homogeneity between groups. (If the resulting $p$-value of Levene's test is less than some significance level, typically 0.05, the obtained differences in sample variances are unlikely to have occurred based on random sampling from a population with equal variances.)

Conceptual understanding of the particulate nature of matter has been shown to be critical for success in learning chemistry and therefore the ParNoMA2 was used to measure students’ conceptual understanding of this major concept in chemistry. The GALT provided a means to access the cognitive development and logical reasoning of students and the results
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

could be used to “better understand the abilities of students and to match instruction accordingly” (Roadrangka et al., 1983, p. 153). It was thought that there should be growth in conceptual development through a content based course such as chemistry, however, it was not known how much development in formal reasoning would take place in that same full year course that takes place over a 90 day period on a Block schedule.

The ANCOVA analysis examined the effect of POGIL on chemistry students’ logical analytical skills according to the GALT results and understanding of chemistry according to the ParNoMA2, with significance level of 0.05 for this analysis. Table 3.2 illustrates which effects were studied.

Table 3.2. The Different Effects that are Examined by ANCOVA

<table>
<thead>
<tr>
<th>GALT (DV)</th>
<th>ParNoMA2 (DV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups (InDGIM vs. POGIL)</td>
<td>Groups (InDGIM vs. POGIL)</td>
</tr>
<tr>
<td>Gender (female vs. male)</td>
<td>Gender (female vs. male)</td>
</tr>
<tr>
<td>Interaction of group (InDGIM vs. POGIL) and gender</td>
<td>Interaction of group (InDGIM vs. POGIL) and gender</td>
</tr>
</tbody>
</table>

According to Tabachnick and Fidell (1996) ANCOVA is often used in “nonexperimental situations where subjects cannot be randomly assigned to treatments…used as a statistical matching procedure…to adjust group means to what they would be if all subjects scored identically on the covariate(s)” (p. 322). The advantage to using ANCOVA, according to Huck (2004), is that of increased power and the control of confounding variables. The means on the dependent variable will “automatically be adjusted to reflect differences among the group means on the covariate variable…and there will be an increase in the statistical power of the inferential tests” (Huck, 2004, p. 397). The data on the covariate and the dependent variable will be used to
compute the adjusted means on the dependent variable, with the focus of ANCOVA relying on these adjusted means. By reducing the error variance using the correlation that may exist between the covariate and the controlled factor, the sensitivity of the experiment is improved and the significance of the designed factors may be easier to determine (Ary, Jacobs, & Sorensen, 2010; Trochim & Donnelly, 2006). This is important since comparisons were made, between the POGIL and InDGIM groups, between females and males, and the interaction of group and gender. Standard significance criterion of $p \leq 0.05$ were used to test the null hypotheses.

In order to differentiate the effects of the treatment (POGIL) to the comparison (InDGIM) due to differences in content understanding as measured by performance on unit tests ($H_04$), one of two possible analyses were performed. If there was a gender effect then a 2 x 2 factorial ANCOVA was performed (as stated previously) and if there was no gender effect then a simple Independent Samples t-test was performed on each unit test and the final exam between the two comparison groups.
Chapter Four

Results

The purpose of this research is to determine the statistical differences on the measured outcomes between students, exposed to POGIL, a student-centered cooperative learning instructional model used in the teaching of chemistry, and those exposed to a self-developed guided inquiry approach, InDGIM, and whether these differences vary according to genders or between students performing at different levels of achievement as determined by quartile placement from their grades or from their Keystone Biology or Algebra scores. The measured outcomes included scores on the GALT, ParNoMA2, select unit tests, and the final exam.

The research questions and the null hypotheses for this study are:

Research question 1: What measurable differences does the use of the POGIL instructional model have on high school students understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared with the InDGIM, the currently used inquiry-based instructional model?

Null hypothesis 1: After adjusting for pretest differences, there will be no statistically significant differences on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between high school chemistry students taught using POGIL and students taught using InDGIM, the currently used inquiry-based instructional model.

Research question 2: What measurable differences exist between female and male high school chemistry students on the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) regardless of the instructional model used to teach chemistry?
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Null hypothesis 2: After adjusting for pretest differences, there will be no statistically significant differences on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between female and male students taking a high school chemistry course.

Research question 3: What measurable differences does the use of the POGIL instructional model or the InDGIM instructional model have on high school females’ understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared to high school males’ understanding of chemistry (ParNoMA2) and logical reasoning (GALT)?

Null hypothesis 3: After adjusting for pretest differences, there will be no statistically significant difference on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) of female and male students who were taught using the POGIL instructional model and female and male students taught using InDGIM, the currently used inquiry-based instructional model.

Research question 4: What is the differential pattern of performance in the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between those students falling in the lowest performance quartile (as determined by their Keystone Algebra scores) taught using POGIL, a student-centered cooperative learning instructional model when compare to InDGIM, the currently used inquiry-based instructional model?

Research question 5: What measurable differences does the use of the POGIL instructional model have on high school students’ performance on unit tests, and the final exam, when compared with the InDGIM, the currently used inquiry-based instructional model?
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Null hypothesis 4: There will be no statistically significant differences on unit assessments measuring the understanding of chemistry content of those students taught using the POGIL instructional model when compared to InDGIM, the currently used inquiry-based instructional model.

Descriptive Statistics for the Pretest and Posttest

Fifty-two (52) students agreed to participate in the research. Forty-seven (47) students completed all the parameters of the study, which included taking the pre and post GALT, ParNoMA2, all unit tests, and the final exam. The 5 students not included in the final sample did not complete the post GALT or the post ParNoMA2 tests. There were twenty-four (24) students in the POGIL group (14 females and 10 males) and twenty-three (23) students in the InDGIM group (16 females and 7 males). The gender representation of the 47 students included in the final sample was 30 females (63.8%) and 17 males (36.2%) overall.

The ParNoMA2 consists of 20 questions designed to determine a student’s misconception with regards to the particulate behavior of matter. A maximum score of 20 is possible on both the pre and the posttests. The GALT is a developmental test that measures a students’ logical thinking using a 22-item questionnaire to test both the ability to answer a question and select the proper reason for the answer. A maximum score of 22 is possible on both the pre and posttests. Descriptive statistics for the pretest and posttest scores by group and gender are presented in Tables 4.1 and 4.2 respectively.
The comparison of two different guided inquiry methods in high school chemistry classes

Table 4.1. Descriptive Statistics of Pretest and Posttest by Original Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Pretest</th>
<th></th>
<th></th>
<th>Posttest</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>ParNoMA2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison (InDGIM)</td>
<td>28</td>
<td>10.93</td>
<td>4.56</td>
<td>23</td>
<td>11.22</td>
<td>3.58</td>
</tr>
<tr>
<td>Treatment (POGIL)</td>
<td>24</td>
<td>9.25</td>
<td>4.66</td>
<td>25</td>
<td>9.54</td>
<td>4.80</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>10.09</td>
<td>4.61</td>
<td>47</td>
<td>10.38</td>
<td>4.19</td>
</tr>
<tr>
<td>GALT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison (InDGIM)</td>
<td>28</td>
<td>11.93</td>
<td>2.59</td>
<td>23</td>
<td>9.96</td>
<td>3.83</td>
</tr>
<tr>
<td>Treatment (POGIL)</td>
<td>24</td>
<td>11.08</td>
<td>3.87</td>
<td>24</td>
<td>10.17</td>
<td>4.63</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>11.51</td>
<td>3.23</td>
<td>47</td>
<td>10.07</td>
<td>4.23</td>
</tr>
</tbody>
</table>

According to the data in Table 4.1, both groups showed an overall increase in ParNoMA2 scores between pretest and posttest. The InDGIM group had a mean ParNoMA2 pretest score of 10.93 ($SD = 4.56$) and a posttest mean of 11.22 ($SD = 3.58$) which is an increase of 2.65%. The POGIL group had a mean ParNoMA2 pretest score of 9.25 ($SD = 4.66$) out of 20 and posttest mean of 9.54 ($SD = 4.80$) which is an increase of 3.14%. Although the POGIL group showed a greater percent increase, it should be noted that each group was not equally matched on the pretest scores, and each had the same overall gain in mean by only 0.29 points.
Both groups showed a marked decrease in GALT scores as can be seen by the data in Table 4.1. The InDGIM group had a mean GALT pretest score of 11.93 ($SD = 2.59$) out of 22 and a posttest score of 9.96 ($SD = 3.83$) which is a decrease of 16.51%. The POGIL group had mean GALT pretest score of 11.08 ($SD = 3.87$) and a posttest score of 10.17 ($SD = 4.63$) which is a decrease of 8.21%. The scores for both groups decreased from pretest to posttest; however, the scores of the InDGIM group indicated a greater decrease in performance.

Table 4.2. Descriptive Statistics of Pretest and Posttest by Gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>ParNoMA2</th>
<th>GALT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td></td>
<td>N  M  SD</td>
<td>N  M  SD</td>
</tr>
<tr>
<td>Female</td>
<td>31 9.42 4.470</td>
<td>30 10.40 4.207</td>
</tr>
<tr>
<td>Total</td>
<td>52 10.33 4.612</td>
<td>47 10.30 4.364</td>
</tr>
</tbody>
</table>

As seen in Table 4.2, the females showed an increase between the ParNoMA2 pretest and posttest whereas the male subjects showed a decrease. The female subjects had a mean ParNoMA2 pretest score of 9.42 ($SD = 4.47$) and a posttest mean of 10.40 ($SD = 4.21$), an increase of 10.40%. The male subjects had a mean ParNoMA2 pretest score of 11.25 ($SD =
4.75) out of 20 and posttest mean of 10.20 ($SD = 4.52$), a decrease of 9.33%. The females had a posttest mean for gain scores of 0.967 ($SD = 3.22$) and the males had a posttest mean for gain scores of -0.706 ($SD = 4.04$). Although the posttest scores showed a difference in direction, it should be noted that both genders had comparable posttest scores.

According to Table 4.2 the scores on the GALT dropped for both groups, with the scores for male subjects dropping more. The female subjects had a mean GALT pretest score of 11.00 ($SD = 3.10$) out of 22 and a posttest score of 10.07 ($SD = 4.27$) which is a decrease of 8.45%. The male subjects had a mean GALT pretest score of 12.38 ($SD = 3.34$) and a posttest score of 10.06 ($SD = 4.24$) which is a decrease of 18.86%. The females had a posttest mean for gain scores of -0.933 ($SD = 2.65$) and the males had a posttest mean for gain scores of -2.12 ($SD = 4.04$). Although the scores dropped for both, the final scores on the posttest were almost identical between genders.

**Fidelity of Implementation**

**Inter-rater reliability.**

In order to make sure that the teaching methodology programs were being implemented correctly the instructor was monitored on multiple occasions and checklists were used to determine if the activity being monitored was either a POGIL activity or not. The two checklists found in Appendices G and H were used by the instructor and the observer who came into the classroom and a degree of agreement was computed. The following data illustrates the comparison of the ratings for the 11 different POGIL activities administered in the treatment group and the 11 different activities that were administered in the comparison group on the same day of the POGIL activity. Because the premise of this study is to compare POGIL to another
type of guided inquiry pedagogy it was important to establish that the activities performed with the comparison group could be defined as guided inquiry. The instructor and observer also used the checklist (Appendix G) to define the activities as guided inquiry-based on where they fell on the spectrum between teacher centered and student centered.

At least two raters evaluated the activities. The rating scale was dichotomous. If the activity fit the definition of a POGIL activity, it was classified as present (1) if it was not POGIL then it was classified as not present (0). The same scale was used to determine if the alternative activities being presented to the comparison group were not POGIL based. This scale was used for ten different questions listed on the form in Appendix H. The results for this agreement for the POGIL activities are illustrated in Table 4.3 and the “Non-POGIL” activities are in Table 4.4.

Table 4.3.
Inter-Rater Agreement for Determining if Activity Should Be Classified as a POGIL Activity

<table>
<thead>
<tr>
<th>Lesson Number</th>
<th>Rater One’s Score</th>
<th>POGIL Y/N</th>
<th>Rater Two’s Score</th>
<th>POGIL Y/N</th>
<th>% Agreement Per Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9/10</td>
<td>Y</td>
<td>9/10</td>
<td>Y</td>
<td>100.00</td>
</tr>
<tr>
<td>2</td>
<td>8/10</td>
<td>Y</td>
<td>8/10</td>
<td>Y</td>
<td>100.00</td>
</tr>
<tr>
<td>3</td>
<td>9/10</td>
<td>Y</td>
<td>8/10</td>
<td>Y</td>
<td>88.89</td>
</tr>
<tr>
<td>4</td>
<td>8/10</td>
<td>Y</td>
<td>9/10</td>
<td>Y</td>
<td>88.89</td>
</tr>
<tr>
<td>5</td>
<td>9/10</td>
<td>Y</td>
<td>10/10</td>
<td>Y</td>
<td>90.00</td>
</tr>
<tr>
<td>6</td>
<td>8/10</td>
<td>Y</td>
<td>9/10</td>
<td>Y</td>
<td>88.89</td>
</tr>
<tr>
<td>7</td>
<td>9/10</td>
<td>Y</td>
<td>9/10</td>
<td>Y</td>
<td>100.00</td>
</tr>
<tr>
<td>8</td>
<td>10/10</td>
<td>Y</td>
<td>8/10</td>
<td>Y</td>
<td>80.00</td>
</tr>
<tr>
<td>9</td>
<td>9/10</td>
<td>Y</td>
<td>9/10</td>
<td>Y</td>
<td>100.00</td>
</tr>
<tr>
<td>10</td>
<td>9/10</td>
<td>Y</td>
<td>8/10</td>
<td>Y</td>
<td>88.89</td>
</tr>
<tr>
<td>11</td>
<td>8/10</td>
<td>Y</td>
<td>7/10</td>
<td>Y</td>
<td>87.50</td>
</tr>
</tbody>
</table>

Average Agreement Between Raters on Lessons 92.10

It was determined that the raters were in agreement, on whether the lessons being taught, were a POGIL lesson, 92.10% of the time and both raters were in 100.00% agreement that the
science content being taught utilized POGIL methodologies.

There were at least 11 observations when it was apparent that the POGIL group was exposed to material that wasn’t developed as a POGIL activity (lectures, PowerPoint presentations, worksheets). It was determined that the raters were in agreement on whether the lesson being taught was not a POGIL lesson 90.90% of the time; however, both raters were in 100.00% agreement that the activities taught to the comparison (InDGIM) group were not defined as POGIL activities.

Table 4.4.
Rater Agreement for Determining if Activity Should Be Classified as a Non-POGIL Activity

<table>
<thead>
<tr>
<th>Lesson Number</th>
<th>Rater One’s Score</th>
<th>POGIL Y/N</th>
<th>Rater Two’s Score</th>
<th>POGIL Y/N</th>
<th>% Agreement Per Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2/10</td>
<td>N</td>
<td>2/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>2</td>
<td>1/10</td>
<td>N</td>
<td>1/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>3</td>
<td>1/10</td>
<td>N</td>
<td>1/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>4</td>
<td>1/10</td>
<td>N</td>
<td>1/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>5</td>
<td>1/10</td>
<td>N</td>
<td>2/10</td>
<td>N</td>
<td>50.00</td>
</tr>
<tr>
<td>6</td>
<td>1/10</td>
<td>N</td>
<td>1/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>7</td>
<td>1/10</td>
<td>N</td>
<td>1/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>8</td>
<td>4/10</td>
<td>N</td>
<td>4/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>9</td>
<td>4/10</td>
<td>N</td>
<td>2/10</td>
<td>N</td>
<td>50.00</td>
</tr>
<tr>
<td>10</td>
<td>2/10</td>
<td>N</td>
<td>2/10</td>
<td>N</td>
<td>100.00</td>
</tr>
<tr>
<td>11</td>
<td>4/10</td>
<td>N</td>
<td>4/10</td>
<td>N</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Average Agreement Between Raters on Lessons | 90.90

Both groups were exposed to a variety of lab activities throughout the course. These activities were evaluated to determine where they exist on the inquiry-learning spectrum. These
Table 4.5.  
The Degree of Agreement Between Two Raters on the 25 Different “Inquiry” Based Lab Activities Performed in Academic Chemistry.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Rater One</th>
<th>Rater Two</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making Meringue</td>
<td>6/30</td>
<td>6/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Chemical Changes in a Bag</td>
<td>24/30</td>
<td>27/30</td>
<td>88.89</td>
</tr>
<tr>
<td>Physical vs. Chemical Changes</td>
<td>15/30</td>
<td>12/30</td>
<td>80.00</td>
</tr>
<tr>
<td>Solubility</td>
<td>9/30</td>
<td>9/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Density: Test Tube Challenge</td>
<td>24/30</td>
<td>27/30</td>
<td>88.89</td>
</tr>
<tr>
<td>Density: Solids &amp; Liquids</td>
<td>6/30</td>
<td>6/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Thickness of Aluminum Foil</td>
<td>6/30</td>
<td>6/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Properties of Matter Lab</td>
<td>21/30</td>
<td>24/30</td>
<td>87.50</td>
</tr>
<tr>
<td>Gas Laws Demonstrations</td>
<td>21/30</td>
<td>18/30</td>
<td>85.71</td>
</tr>
<tr>
<td>Syringe/Pressure Lab</td>
<td>18/30</td>
<td>18/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Atomic Scale Model</td>
<td>12/30</td>
<td>12/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Indirect Measurement and Size of the Atom</td>
<td>18/30</td>
<td>12/30</td>
<td>66.66</td>
</tr>
<tr>
<td>Isotopic Masses</td>
<td>15/30</td>
<td>9/30</td>
<td>60.00</td>
</tr>
<tr>
<td>Conservation of Matter</td>
<td>9/30</td>
<td>12/30</td>
<td>75.00</td>
</tr>
<tr>
<td>Electron Probabilities</td>
<td>6/30</td>
<td>6/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Flame Tests</td>
<td>15/30</td>
<td>15/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Gas Discharge Tubes</td>
<td>15/30</td>
<td>15/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Alkali &amp; Alkali Earth Metals</td>
<td>12/30</td>
<td>12/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Halogens</td>
<td>12/30</td>
<td>12/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Metal Reactivity’s</td>
<td>21/30</td>
<td>18/30</td>
<td>85.71</td>
</tr>
<tr>
<td>Jelly Bean Formulas</td>
<td>6/30</td>
<td>6/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Chemical Reaction Race</td>
<td>21/30</td>
<td>24/30</td>
<td>87.50</td>
</tr>
<tr>
<td>Chemical Reaction Demonstrations</td>
<td>15/30</td>
<td>15/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Empirical Formulas</td>
<td>12/30</td>
<td>12/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Stoichiometry</td>
<td>6/30</td>
<td>6/30</td>
<td>100.00</td>
</tr>
<tr>
<td>Nuclear Penny Flip</td>
<td>12/30</td>
<td>9/30</td>
<td>75.00</td>
</tr>
<tr>
<td>Mean</td>
<td>13.73</td>
<td>13.38</td>
<td></td>
</tr>
</tbody>
</table>

Average Level of Rater Agreement: 85.63
activities were rated using the Science Teacher Inquiry Rubric (STIR) (see Appendix G). Raters used a 5-point scale (1=no evidence of inquiry to 5=independent or authentic student inquiry) for six separate questions. The results for this inter-observer reliability are included in Table 4.5.

It was determined that the raters were in agreement on whether the lesson being taught was on the spectrum of inquiry 85.63% of the time; however, both raters were in 100.00% agreement that the activities taught to the comparison (InDGIM) group were classified as falling midway between fully teacher centered lessons to fully student centered lessons, which is within the realm of guided inquiry.

According to the rating scales from the Science Teacher Inquiry Rubric (see Appendix G) and the POGIL Guideline Checklist (see Appendix H) it can be determined that a) the activities presented to the treatment group were defined as POGIL activities; b) the activities presented to the comparison group were not defined as POGIL activities; and c) the inquiry-based activities/labs presented to both groups were considered to be guided inquiry type activities.

**Internal consistency reliability.**

Cronbach’s Alpha was used to measure internal consistency reliability on the 20 item ParNoMA2 and the 22 item GALT pre and posttests. Cronbach’s Alpha is a way to see how closely related items are as a group. This value can range from 0 to 1.00 with higher values representing a greater measure of reliability with those scores higher than .70 considered acceptable (McMillan & Schumacher, 2010). Previous studies have determined the internal consistency reliability of the ParNoMA2 to have a Cronbach Alpha of 0.78 when tested on college students (Yezierski & Birk, 2006) and the GALT to have a Cronbach Alpha reliability
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

coefficient of 0.85 (Roadrangka et al., 1983). Table 4.6 illustrates the internal consistency for each of the tests. Both tests exceeded the .70 criteria for the pretest and the posttest.

Table 4.6. *Cronbach’s Alpha Reliability Estimates for the ParNoMA2 and the GALT Tests Administered in Academic Chemistry Classes.*

<table>
<thead>
<tr>
<th>Test</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest (n = 52)</td>
</tr>
<tr>
<td>ParNoMA2</td>
<td>.844</td>
</tr>
<tr>
<td>GALT</td>
<td>.726</td>
</tr>
</tbody>
</table>

**Results of Group Differences**

Since this study used intact groups (classes), group performance on the pre and posttests could not be equally matched due to a lack of random selection of subjects or groups. This particular weakness could be controlled for, to a certain degree, using analysis of covariance (ANCOVA). In order to determine the impact of pretest scores on posttest scores, an ANCOVA was used to test for differences in posttest scores from each group (POGIL vs. InDGIM) after controlling for the differences between the pretests scores on each standard measure (ParNoMA2 and GALT). A 2 x 2 Factorial ANCOVA examined the main effects of Group and Gender on test results. The posttest was the dependent variable and the pretest results were entered as a covariate to adjust for any pretest differences in the InDGIM and POGIL groups on both the ParNoMA2 and GALT posttests. Gender and grouping (POGIL or InDGIM) were entered as fixed factors. Table 4.7 presents the results of the ANCOVA for the ParNoMA2 and Table 4.8 illustrates the results for the GALT.
Table 4.7.
Analysis of Covariance for ParNoMA2 Posttest Scores as a Function of Grouping (POGIL vs. InDGIM) and Gender, Using the Pretest Scores as a Covariate

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta²</th>
<th>Noncent. Parameter</th>
<th>Observed Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>400.611\textsuperscript{a}</td>
<td>4</td>
<td>100.153</td>
<td>9.469</td>
<td>.000</td>
<td>.474</td>
<td>37.875</td>
<td>.999</td>
</tr>
<tr>
<td>Intercept</td>
<td>104.567</td>
<td>1</td>
<td>104.567</td>
<td>9.886</td>
<td>.003</td>
<td>.191</td>
<td>9.886</td>
<td>.867</td>
</tr>
<tr>
<td>Pretest</td>
<td>367.153</td>
<td>1</td>
<td>367.153</td>
<td>34.712</td>
<td>.000</td>
<td>.452</td>
<td>34.712</td>
<td>1.000</td>
</tr>
<tr>
<td>Group</td>
<td>3.324</td>
<td>1</td>
<td>3.324</td>
<td>.314</td>
<td>.578</td>
<td>.007\textsuperscript{c}</td>
<td>.314</td>
<td>.085</td>
</tr>
<tr>
<td>Gender</td>
<td>11.024</td>
<td>1</td>
<td>11.024</td>
<td>1.042</td>
<td>.313</td>
<td>.024\textsuperscript{c}</td>
<td>1.042</td>
<td>.170</td>
</tr>
<tr>
<td>Group*Gender</td>
<td>.015</td>
<td>1</td>
<td>.015</td>
<td>.001</td>
<td>.971</td>
<td>.000\textsuperscript{c}</td>
<td>.001</td>
<td>.050</td>
</tr>
<tr>
<td>Error</td>
<td>444.240</td>
<td>42</td>
<td>10.577</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5891.000</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>844.851</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} R squared = .474 (Adjusted R Squared = .424)
\textsuperscript{b} Computed using alpha = .05
\textsuperscript{c} Partial Eta² measures effect size. According to Cohen (1988) 0.000 to 0.003 indicates no effect; 0.010 to 0.040 indicates a small effect; 0.060 to 0.110 indicates an intermediate effect; and 0.140 and higher indicates a large effect. According to Hattie (2009) the desired effect size should be greater than 0.039.
Table 4.8.
Analysis of Covariance for GALT Posttest Scores as a Function of Grouping (POGIL vs. InDGIM) and Gender, Using the Pretest Scores as a Covariate

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta$^2$</th>
<th>Noncent. Parameter</th>
<th>Observed Power$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>544.812$^a$</td>
<td>4</td>
<td>136.203</td>
<td>14.301</td>
<td>.000</td>
<td>.577</td>
<td>57.206</td>
<td>1.000</td>
</tr>
<tr>
<td>Intercept</td>
<td>4.734</td>
<td>1</td>
<td>4.734</td>
<td>.497</td>
<td>.485</td>
<td>.012</td>
<td>.497</td>
<td>.106</td>
</tr>
<tr>
<td>Pretest</td>
<td>525.298</td>
<td>1</td>
<td>525.298</td>
<td>55.157</td>
<td>.000</td>
<td>.568</td>
<td>55.157</td>
<td>1.000</td>
</tr>
<tr>
<td>Group</td>
<td>16.825</td>
<td>1</td>
<td>16.825</td>
<td>1.767</td>
<td>.191</td>
<td>.040$^c$</td>
<td>1.767</td>
<td>.255</td>
</tr>
<tr>
<td>Gender</td>
<td>22.097</td>
<td>1</td>
<td>22.097</td>
<td>2.320</td>
<td>.135</td>
<td>.052$^c$</td>
<td>2.320</td>
<td>.319</td>
</tr>
<tr>
<td>Group*Gender</td>
<td>9.508</td>
<td>1</td>
<td>9.508</td>
<td>.998</td>
<td>.323</td>
<td>.023$^c$</td>
<td>.998</td>
<td>.164</td>
</tr>
<tr>
<td>Error</td>
<td>399.997</td>
<td>42</td>
<td>9.524</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6566.000</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>944.809</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R squared = .577 (Adjusted R Squared = .536)
b. Computed using alpha = .05
c. Partial Eta$^2$ measures effect size. According to Cohen (1988) 0.000 to 0.003 indicates no effect; 0.010 to 0.040 indicates a small effect; 0.060 to 0.110 indicates an intermediate effect; and 0.140 and higher indicates a large effect. According to Hattie (2009) the desired effect size should be greater than 0.039.

Covariate Effect

The pretest scores (covariate) on the ParNoMA2 (p < .001) and the GALT (p < .001) had a significant impact on the posttest scores see (see Tables 4.7 and 4.8). Posttest means for each factor and the interaction were adjusted to take into account the pretest effect.

A high post hoc power (1.000), which IBM SPSS (Leech, Barrett, & Morgan, 2011) labels observed power, for both pretests usually indicates that the sample size was large enough to determine any influence that the pretest would have on the posttest. The post hoc power is
related to the $p$-value in that it is a direct relationship, the lower the $p$-value, the higher the power. With an observable $p$-value of $p = .05$ the observed power should be 50% (Lakens, 2014). It appears that the pretest did have a significant impact on the results of the ParNoMA2 and the GALT posttest, for this sample size, however, gender, grouping, or the combined effects were not determined to have any significant impact on how students performed on the ParNoMA2 or GALT posttest.

This was not the case for the other factors. Using the ANCOVA the observed power for the ParNoMA2 and group was .085; for gender it was .170; and for the combined factors it was a .050. The observed power, from the ANCOVA, for the GALT and group was .255; for gender it was .319; and for the combined factors it was .104. O’Keefe (2007) states that “a non-significant result does not mean that the population effect is in fact zero; it means only that a population effect of zero cannot be ruled out” (p. 296).

**Null hypothesis and research question one.**

This study was conducted to determine the measurable differences that using POGIL would have on high school students’ understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared with an independently developed inquiry-based instructional model, InDGIM. The first null hypothesis states that after adjusting for pretest differences, there would be no statistically significant differences on posttest scores measuring for the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between high school chemistry students taught using POGIL and students taught using InDGIM, the currently used inquiry-based instructional model.

A 2 x 2 ANCOVA was used to test the first hypothesis comparing the outcomes of the
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

POGIL group to the InDGiM group on the ParNoMA2 and GALT posttests. The adjusted marginal means are presented in Table 4.9.

The main effect for Group on the ParNoMA2 was not significant \((F(1,46) = .314, p = .578, \text{partial } \eta^2 = .007)\), see Table 4.7. The POGIL group posttest estimated marginal mean of 10.004 \((SD = .661)\) was fairly close to the InDGiM posttest marginal mean of 10.734 \((SD = .676)\), see Table 4.9. Power was low at .085 and the partial \(\eta^2 = .007\) suggests virtually no effect size. Similarly, no significant difference was found on the GALT measure on the Group factor \((F(1,46) = 1.767, p = .191, \text{partial } \eta^2 = .040)\) (see Table 4.8). Posttest adjusted marginal means and standard deviations are presented in Table 4.9. Based on the ANCOVA results, \(H_01\), which stated that after adjusting for pretest differences, there would be no statistically significant differences on posttest scores measuring for the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between high school chemistry students taught using POGIL and students taught using InDGiM, was not rejected.

Table 4.9.
Adjusted marginal means and Standard Deviation for Groups Main Effect on the ParNoMA2 and GALT Posttests from the ANCOVA

<table>
<thead>
<tr>
<th>Group</th>
<th>ParNoMA2 Posttest</th>
<th>GALT Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td><strong>Comparison</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(InDGIM)</td>
<td>23</td>
<td>10.734</td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(POGIL)</td>
<td>24</td>
<td>10.004</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>47</td>
<td>10.369</td>
</tr>
</tbody>
</table>
Null hypothesis and research question two.

The second research question considers differences by gender and asks if there are measurable differences between female and male high school chemistry students on the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) regardless of the instructional model used to teach chemistry? The second null hypothesis (Ho2) states that after adjusting for pretest differences, there will be no statistically significant differences on posttest

Table 4.10.
Adjusted Marginal Means and Standard Deviation of the Main Effect due to Gender on the ParNoMA2 and GALT Posttests from the ANCOVA

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ParNoMA2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>30</td>
<td>10.769&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.586</td>
<td>1.042</td>
<td>.313</td>
</tr>
<tr>
<td>Male</td>
<td>17</td>
<td>9.644&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.781</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>10.207</td>
<td>.684</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GALT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>30</td>
<td>10.454&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.555</td>
<td>2.320</td>
<td>.135</td>
</tr>
<tr>
<td>Male</td>
<td>17</td>
<td>9.375&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.740</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>9.960</td>
<td>.648</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Covariates appearing in the model are evaluated at the following values: Pre-ParNoMA2 out of 20.00 = 10.00, (the total points possible was 20.00 but the mean value was 10.00 points),

b. Covariates appearing in the model are evaluated at the following values: Pre-GALT out of 22.00 = 11.43, (the total points possible was 22.00 but the mean value was 11.43 points).
scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between female and male students taking a high school chemistry course. Referring back to Tables 4.7 and 4.8 the main effect due to gender was not significantly different on posttest scores for either the ParNoMA2 ($F(1,46) = 1.042, p = .313, \text{partial } \eta^2 = .024$) or the GALT ($F(1,46) = 2.320, p = .135, \text{partial } \eta^2 = .052$) with females having an estimated marginal mean of 10.769 ($SD = .586$) on the ParNoMA2 and a 10.454 ($SD = .555$) on the GALT. The males had estimated adjusted marginal means of 9.644 ($SD = .781$) on the ParNoMA2 and a 9.375 ($SD = .740$) on the GALT (see Table 4.10).

An Independent Samples t-test was also performed to see if there was a significant difference between genders in the gain scores on either the ParNoMA2 or the GALT. There was no statistically significant difference between females and males when comparing gain scores on the ParNoMA2 ($p=.126$) and the GALT ($p=.200$). For the ParNoMA2 males had lower gain scores ($M = -.706, SD = 4.04$) than females ($M = .967, SD = 3.22$) with a small effect size ($Hedges_g = .473$). The same pattern occurred with the GALT, males ($M = -2.12, SD = 3.55$) lost ground to the females ($M = -.933, SD = 2.65$) with a small effect size ($Hedges_g = .396$). However, for both tests the differences between males and females on gain scores were not significant. This further confirms that student differences on either the GALT or ParNoMA2 measures suggest no difference due to Gender.

**Null hypothesis and research question three.**

Is there a difference in the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between females and males taught using POGIL when compared to InDGIM? In order to determine if the gains made by the POGIL group were greater for one gender than the
other, the interaction of gender and group was tested. The null hypothesis (Ho3) states that after adjusting for pretest differences, there will be no statistically significant difference on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) of female and male students who were taught using either POGIL or InDGIM. The results can be seen in Tables 4.11 and 4.12.

Table 4.11. Adjusted Marginal Means and Standard Deviation for Group*Gender on the ParNoMA2 Posttest ANCOVA

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group*Gender</td>
<td></td>
<td></td>
<td></td>
<td>.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.971&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>InDGIM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
<td>11.045&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.813</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7</td>
<td>9.970&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.254</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>10.508</td>
<td>1.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POGIL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>10.438&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.882</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>9.436&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.029</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>9.937</td>
<td>.955</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. f and p values for the posttest ParNoMA2 are based on the effects of the interaction between group and gender.

b. Covariates appearing in the model are evaluated at the following values: Pre-ParNoMA2 out of 20.00 = 10.00 (the total points was 20.00 but the mean value was only 10.00 points).

The adjusted marginal mean scores (Tables 4.11 and 4.12) for females were higher than the males for both tests in both the comparison (InDGIM) and treatment (POGIL) groups, with the highest scores in the comparison group. The males scored slightly higher on the ParNoMA2 in the comparison group and on the GALT in the treatment group. Results presented in Tables 4.7 and 4.8 indicate that there was no interaction effect for the ParNoMA2 (F(1,46) = .001, p > .05)
or the GALT \((F(1,46) = .998, p > .05)\). The ParNoMA2 partial \(\eta^2 = .000\), and for the GALT the partial \(\eta^2 = .023\). The adjusted marginal mean scores (Tables 4.11 and 4.12) for females were higher than the males for both tests in both the comparison (InDGIM) and treatment (POGIL) groups, with the highest scores in the comparison group. The males scored slightly higher on the ParNoMA2 in the comparison group and on the GALT in the treatment group.

Table 4.12. Adjusted marginal means and Standard Deviation for Groups*Gender on the GALT Posttest ANCOVA

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group*Gender</td>
<td></td>
<td></td>
<td></td>
<td>.998a</td>
<td>.323a</td>
</tr>
<tr>
<td>InDGIM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
<td>10.306b</td>
<td>.753</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7</td>
<td>8.073b</td>
<td>1.147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>9.190</td>
<td>.950</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POGIL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>10.635b</td>
<td>.816</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>10.269b</td>
<td>.956</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>10.452</td>
<td>.886</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. \(f\) and \(p\) values for the posttest GALT are based on the effects of the interaction between group and gender.
b. Covariates appearing in the model are evaluated at the following values: Pre-GALT out of 22.00 = 11.43 (the total points possible was 22.00 but the mean value was 11.43 points).

\textbf{Research question four.}

What is the differential pattern of performance in the understanding of chemistry chemistry (ParNoMA2) and logical reasoning (GALT) for those students falling in the lowest performance quartile (as determined by their Keystone Algebra scores) taught using POGIL, a student-centered cooperative learning instructional model, when compare to InDGIM, the
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

currently used inquiry-based instructional model?

Because of the small n’s for each subgroup, a simple descriptive analysis was performed on this data. Table 4.13 illustrates the descriptive statistics for student performance on the ParNoMA2 and GALT tests when they are ranked in the lowest quartile, based on Keystone Algebra scores, based on groups. The InDGIM group showed a decrease in ParNoMA2 scores

Table 4.13. 
Descriptive Statistics of Pre and Posttest by Lowest Quartile as Sorted by Results on the Keystone Algebra Exam for Both InDGIM and POGIL Groups

<table>
<thead>
<tr>
<th>Test</th>
<th>ParNoMA2</th>
<th>GALT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison (InDGIM)</td>
<td>5</td>
<td>12.60</td>
</tr>
<tr>
<td>Treatment (POGIL)</td>
<td>8</td>
<td>5.63</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>9.12</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison (InDGIM)</td>
<td>5</td>
<td>13.60</td>
</tr>
<tr>
<td>Treatment (POGIL)</td>
<td>8</td>
<td>8.50</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>11.05</td>
</tr>
</tbody>
</table>
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

but the POGIL group showed an increase in scores. Both groups showed a decrease in their GALT scores. This data illustrates that the lowest quartile in the InDGIM group had a mean ParNoMA2 pretest score of 12.60 ($SD = 2.93$) and a posttest mean of 9.67 ($SD = .577$) which is a decrease of 23.25%. The POGIL group had a mean ParNoMA2 pretest score of 5.63 ($SD = 5.18$) out of 20 and posttest mean of 6.38 ($SD = 2.13$) which is an increase of 13.32%. The InDGIM group had a mean GALT pretest score of 13.60 ($SD = 1.52$) out of 22 and a posttest score of 12.33 ($SD = 5.03$) which is a decrease of 9.34%. The POGIL group had a mean GALT pretest score of 11.05 ($SD = 2.58$) and a posttest score of 9.73 ($SD = 4.73$) which is a decrease of 11.95%. The scores for both groups dropped.

Table 4.14. Descriptive Statistics of Pre and Posttest by Lowest Quartile as Sorted by Results on the Keystone Algebra Exam for Both Genders

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>ParNoMA2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>5.60</td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>10.00</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>7.80</td>
</tr>
<tr>
<td>GALT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>10.00</td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>10.75</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>10.38</td>
</tr>
</tbody>
</table>
Table 4.14 illustrates the performance of the lowest quartile, as sorted by the results of the Keystone Algebra scores, based on gender. The females showed an increase in scores on the ParNoMA2, whereas they had a drop in scores for the GALT. The males showed a drop in both the ParNoMA2 and the GALT. The females had a mean ParNoMA2 pretest score of 5.60 \((SD = 4.93)\) and a posttest mean of 7.60 \((SD = 2.88)\) which is an increase of 35.71\%. The males had a mean ParNoMA2 pretest score of 10.00 \((SD = 4.78)\) and posttest mean of 7.00 \((SD = 2.10)\) which is a decrease of 30.00\%, but both genders had almost the same posttest scores. The same pattern did not occur with the GALT posttest scores. The females had a mean GALT pretest score of 10.00 \((SD = 4.18)\) and a posttest score of 9.20 \((SD = 6.22)\) which is a decrease of 8.00\%. The males had a mean GALT pretest score of 10.75 \((SD = 4.08)\) and a posttest score of 8.00 \((SD = 4.20)\) which is a decrease of 25.58\%. It appears that females performed better than males.

To see how these scores might compare to other students, the same statistics were performed on the students in the top quartile, based on their Keystone Algebra scores. The data is for these students is illustrated in Tables 4.15 and 4.16.

The results found in Table 4.15 illustrate that students in the upper quartile, based on their Algebra Keystone scores, did better in the comparison group compared to the variable group. The scores for the comparison group (InDGIM) on the ParNoMA2 posttest actually increased 9.13\% from 12.60 \((SD = 2.93)\) to 13.75 \((SD = 4.03)\), whereas the scores for the POGIL group decreased 2.60\% from 11.00 \((SD = 4.23)\) to 10.71 \((SD = 4.89)\). The upper quartile students also showed a decrease in their posttest GALT scores. The InDGIM group dropped by only 3.41\% from 13.20 \((SD = 3.11)\) to 12.75 \((SD = 3.80)\) and the POGIL group dropped 10.37\% from 12.43 \((SD = 3.60)\) to 11.14 \((SD = 4.14)\).
The pattern of the results is different than what was observed with the lowest quartile group. The changes that occurred in the ParNoMA2 scores between these two quartile groups are different as well, with the lower quartile group had higher percent changes (increase or decrease) than the highest quartile students. As for the GALT, the decreases that the upper quartile students showed, regardless of pedagogical approach, were far less than those of the lowest quartile.

Table 4.16 shows the differences in performance on the ParNoMA2 and the GALT, by
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

gender for the upper quartile students, as determined by their Keystone Algebra scores. The females in the highest quartile, like their counterparts in the lowest quartile, increased their scores on the ParNoMA2, and decreased their scores on the GALT. The pattern for the males was a little different.

Table 4.16.
*Descriptive Statistics of Pre and Posttest by Highest Quartile as Sorted by Results on the Keystone Algebra Exam for Both Genders*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>ParNoMA2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>11.17</td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
<td>12.00</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>7.80</td>
</tr>
</tbody>
</table>

| GALT      |     |       |      |     |       |      |
| Gender    |     |       |      |     |       |      |
| Female    | 6   | 11.00 | 3.464  | 5   | 9.60  | 4.037 |
| Male      | 6   | 14.50 | 2.074  | 6   | 13.50 | 2.739 |
| Total     | 12  | 10.38 | 4.087  | 11  | 8.60  | 5.208 |

This data illustrates that the highest quartile among the females had a mean ParNoMA2 pretest score of 11.17 ($SD = 5.57$) and a posttest mean of 11.60 ($SD = 6.27$) which is an increase of 3.85%. The males had a mean ParNoMA2 pretest score of 12.00 ($SD = 3.52$) and posttest mean of 12.00 ($SD = 3.41$) which shows no change. Both genders of the highest quartile showed
a decrease in their posttest GALT scores, which is similar to the genders in the lowest quartile, however, not to the same degree. The males dropped from 14.50 ($SD = 2.07$) to 13.50 ($SD = 2.74$), a drop of 6.90%, and the females dropped from 11.00 ($SD = 3.46$) to 9.60 ($SD = 4.04$), a drop of 12.73%. The females in this upper quartile outperformed the males on only one test whereas the males showed better retention on the second test.

**Null hypothesis and research question five.**

The final research question asked if there would be a difference in unit performance, as measured on unit tests, and the final exam, between those students taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model. The null hypothesis (Ho4) stated there will be no statistically significant differences on unit assessments measuring the understanding of chemistry content of those students taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model.

Since early analysis indicated that there was no influence of gender on the overall performance, independent samples t-tests were used to identify any differences found on the 19 unit assessments and the final exam that might exists between groups. These assessments were administered to both the comparison (InDGIM) and treatment (POGIL) groups. The following assumptions were tested and met: (a) groups are approximately the same size, (b) the variances of the two populations are equal, (c) observations were independent, and (d) the dependent variable was approximately normally distributed.
Table 4.17.
A Comparison of the Outcomes of the Assessments Given to Both the InDGIM (I) and POGIL (P) Groups

<table>
<thead>
<tr>
<th>Assessment Name</th>
<th>Max. Score</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Hedges g</th>
<th>t</th>
<th>Sig.</th>
<th>Holm-Bonferroni Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making Meringue</td>
<td>4.00</td>
<td>P = 24</td>
<td>2.64</td>
<td>1.12</td>
<td>.714</td>
<td>6.572</td>
<td>.013*</td>
<td>.234</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>3.39</td>
<td>.988</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety, Rules, Regulations</td>
<td>25.00</td>
<td>P = 11</td>
<td>20.62</td>
<td>2.06</td>
<td>.474</td>
<td>1.355</td>
<td>.256</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 14</td>
<td>21.66</td>
<td>2.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rxn in a Bag</td>
<td>5.00</td>
<td>P = 24</td>
<td>3.75</td>
<td>.897</td>
<td>.356</td>
<td>1.654</td>
<td>.204</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>4.07</td>
<td>.899</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes (P and C)</td>
<td>5.00</td>
<td>P = 24</td>
<td>2.87</td>
<td>1.51</td>
<td>.713</td>
<td>6.785</td>
<td>.012*</td>
<td>.228</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>3.81</td>
<td>1.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A, P, SN, SF</td>
<td>6.00</td>
<td>P = 24</td>
<td>4.70</td>
<td>1.77</td>
<td>.359</td>
<td>1.585</td>
<td>.214</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>5.20</td>
<td>.966</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solubility</td>
<td>5.00</td>
<td>P = 24</td>
<td>3.29</td>
<td>1.16</td>
<td>.107</td>
<td>.140</td>
<td>.710</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 27</td>
<td>3.41</td>
<td>1.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement</td>
<td>22.00</td>
<td>P = 24</td>
<td>15.88</td>
<td>2.78</td>
<td>.444</td>
<td>2.553</td>
<td>.116</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>17.18</td>
<td>3.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boyle’s Law</td>
<td>5.00</td>
<td>P = 24</td>
<td>3.40</td>
<td>1.03</td>
<td>.171</td>
<td>.415</td>
<td>.522</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>3.58</td>
<td>1.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charles’ Law</td>
<td>5.00</td>
<td>P = 24</td>
<td>4.08</td>
<td>.634</td>
<td>.643</td>
<td>5.439</td>
<td>.024*</td>
<td>.360</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>4.42</td>
<td>.418</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas Laws</td>
<td>18.00</td>
<td>P = 24</td>
<td>14.44</td>
<td>2.77</td>
<td>.714</td>
<td>6.283</td>
<td>.015*</td>
<td>.255</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>16.00</td>
<td>1.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atoms &amp; PT</td>
<td>16.00</td>
<td>P = 24</td>
<td>12.89</td>
<td>2.27</td>
<td>.524</td>
<td>4.149</td>
<td>.047*</td>
<td>.611</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>13.88</td>
<td>1.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atomic Theory</td>
<td>28.00</td>
<td>P = 24</td>
<td>20.11</td>
<td>4.20</td>
<td>.559</td>
<td>4.034</td>
<td>.050*</td>
<td>.600</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 28</td>
<td>22.09</td>
<td>2.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave Behavior</td>
<td>7.00</td>
<td>P = 24</td>
<td>5.25</td>
<td>1.28</td>
<td>.344</td>
<td>1.467</td>
<td>.232</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = 27</td>
<td>4.76</td>
<td>1.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.17 (continued)

<table>
<thead>
<tr>
<th></th>
<th>P = 23</th>
<th>I = 27</th>
<th>P = 24</th>
<th>I = 28</th>
<th>P = 23</th>
<th>I = 28</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DeBroglie &amp;</strong></td>
<td>7.00</td>
<td>5.45</td>
<td>.993</td>
<td>.642</td>
<td>5.111</td>
<td>.028*</td>
</tr>
<tr>
<td><strong>Wein’s Law</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.06</td>
<td>.911</td>
<td>19.85</td>
<td>4.14</td>
<td>.659</td>
<td>.022*</td>
</tr>
<tr>
<td></td>
<td>5.111</td>
<td>.028*</td>
<td>5.632</td>
<td>.352</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Quantum**     | 27.00  | 17.85  | 3.83   | .346   |
|                | 6.06   | .911   | 20.95  | .950   |
|                | 5.111  | .028*  | 3.234  | .078   |

| **Periodicity I** | 5.00 | 3.45 | .822 | .346 | .869 | .359 |
|                  | I = 16 | 3.76 | .950 |

| **Periodicity II** | 28.00 | 14.38 | 3.83 | .500 | 3.234 | .078 |
|                   | I = 28 | 20.50 | 4.39 |

| **Formulas**      | 13.00 | 10.01 | 1.80 | .399 | 1.985 | .165 |
|                  | I = 28 | 10.72 | 1.76 |

| **Empirical Formulas** | 9.00 | 7.24 | 1.39 | .126 | .205 | .653 |
|                       | I = 28 | 7.05 | 1.59 |

| **Final Exam** | 100.00 | 77.39 | 15.33 | .202 | .506 | .480 |
|                | I = 27 | 80.76 | 17.74 |

*Indicating tests with statistically significant outcomes.

Table 4.17 presents the standard p-value without a Bonferroni correction and p-values adjusted using the Bonferroni correction for the assessments that were administered to both groups. The comparison group (InDGIM) outscored the treatment group (POGIL) on almost all of the assessments except two (Empirical Formulas, and Waves). There were significant differences between the scores of eight different assessments reviewed above (see Table 4.17). Using the Holm-Bonferroni adjustment procedure increases the p-value which yields a more conservative interpretation of the results and reduces Type I errors (false positives), however, the adjustment increases the likelihood of Type II errors (accepting a null hypothesis that is actually false). Since the results of the events (assessments and final exam) are independent of each other and there is no overlap of the characteristic being measured, the Holm-Bonferroni correction may not be necessary. Those tests determined to be significant had an intermediate to large effect.
size between 0.500 and 0.800 (Cohen, 1988) which is within the zone of desired effect (Hattie, 2009). Therefore, there were additional practical differences based on effect size for those group comparisons that were significantly different. Based on this analysis, the null hypothesis was rejected for 8 of the 19 measures and practitioners would expect a practical impact on student scores based on those differences.

**Multiple Regression: Final Examination Scores**

Before a multiple regression can be performed, there are assumptions that must be met. The first assumption is that there must be an established linear relationship between the outcome variable (final exam scores) and the independent variables (gender, treatment, Keystone Algebra and Keystone Biology); second the errors between observed and predicted values should be normally distributed; third, the independent variables should not be highly correlated with each other (no multicollinearity); and finally, the differences between what is observed and what is expected should not show a pattern (Tabachnick et al, 1996).

According to Huck (2004) the easiest way to check on these assumption is to view a scatter diagram of the sample data and if “the data in the sample appear to conform to the

![Figure 4.1](image)

*Figure 4.1. A scatter diagram illustrating the linear relationship between the Residual and the Predicted Values for the Final Grade (as a Percentage) in the Academic Chemistry Course.*
linearity and equal variance assumptions, then the researcher can make an informed guess that linearity and homoscedasticity are also characteristics of the population” (p. 226). Figure 4.1 illustrates a scatter plot of the Final Grade (as a percentage) in the Academic Chemistry course. This illustrates an approximate normal distribution of the residual plots, thus providing evidence that the assumptions of linearity, normality, and homoscedasticity have been met.

When assessing for collinearity tolerances the Keystone Algebra and Biology were collinear. The Keystone Biology variable was retained in the equation because it was more highly correlated with the final exam score (r = .590) while the Keystone Algebra was lower (r = .443) and removed as a predictor.

The means, standard deviations, and inter-correlations for the final exam score (criterion) and predictor variables (gender, treatment, Keystone Algebra scores, and Keystone Biology

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Exam Score (DV)</td>
<td>79.65</td>
<td>16.61</td>
<td>.124</td>
<td>.135</td>
<td>.443*</td>
<td>.590**</td>
</tr>
</tbody>
</table>

**PREDICTORS:**

1. Gender
2. Treatment
3. Keystone Algebra
4. Keystone Biology

*p < .05; **p < .01
scores) can be found in Table 4.18. This combination of variables significantly predicted chemistry achievement, $F(4,43) = 6.619, p < .001$. The adjusted R squared value was .381, which indicates that 38.1% of the variation in final exam scores can be explained by the model.

Table 4.19.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SEB</th>
<th>$\beta$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>5.352</td>
<td>4.166</td>
<td>.159*</td>
<td>.206</td>
</tr>
<tr>
<td>Instructional Treatment</td>
<td>-3.141</td>
<td>4.245</td>
<td>-.095</td>
<td>.463</td>
</tr>
<tr>
<td>Keystone Algebra Score</td>
<td>.032</td>
<td>.109</td>
<td>.048</td>
<td>.773</td>
</tr>
<tr>
<td>Keystone Biology Score</td>
<td>.282</td>
<td>.080</td>
<td>.598**</td>
<td>.001</td>
</tr>
<tr>
<td>Constant</td>
<td>-401.911</td>
<td>124.546</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: $R^2 = .381; F(4,43) = 6.619, p < .001$
*p < .05; **p < .01

The beta weights in table 4.19 (Keystone Algebra included) indicate that out of all of the predictor variables, the Keystone Biology scores had the greatest impact on the overall final examination score in the course. It significantly contributed ($p<.01$) to the explanation of the variation in Final Exam Scores (DV). Gender, the treatment variable of either POGIL or InDGIM, and the Keystone Algebra score had a negligible impact and did not significantly ($p>.05$) contribute to the explanation of the variation on the overall final exam score. Roughly 34.8% ($r^2 = .590^2$) of the final exam grade can be explained based on student’s performance on the Keystone Biology exam taken before the course.
Table 4.20.
Means, Standard Deviations, and Inter-correlations for the Final Grade as a Percentage in the Academic Chemistry Course and Predictor Variables (N=48)

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Grade in Course (DV)</td>
<td>76.26</td>
<td>10.88</td>
<td>.406*</td>
<td>.236</td>
<td>.319*</td>
<td>.452**</td>
</tr>
</tbody>
</table>

PREDICTORS:

1. Gender
2. Treatment
3. Keystone Algebra
4. Keystone Biology

Note: R² = .394; F(4,43) = 6.984, p < .001
*p < .05; **p < .01

The means, standard deviations, and inter-correlations for the final grade (as a percentage) can be found in Table 4.20. This combination of variables significantly predicted

Table 4.21.
Simultaneous Multiple Regression Analysis for Gender, Instructional Treatment, Keystone Algebra Scores, and Keystone Biology Scores Predicting Final Grade as a Percentage in Chemistry (N = 48)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SEB</th>
<th>β</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>9.666</td>
<td>2.700</td>
<td>.439**</td>
<td>.001</td>
</tr>
<tr>
<td>Instructional Treatment</td>
<td>.670</td>
<td>2.751</td>
<td>.031</td>
<td>.809</td>
</tr>
<tr>
<td>Keystone Algebra Score</td>
<td>-.037</td>
<td>.070</td>
<td>-.086</td>
<td>.598</td>
</tr>
<tr>
<td>Keystone Biology Score</td>
<td>.161</td>
<td>.052</td>
<td>.522**</td>
<td>.003</td>
</tr>
<tr>
<td>Constant</td>
<td>-119.748</td>
<td>80.715</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: R² = .394; F(4,43) = 6.984, p<.001
*p < .05; **p < .01
chemistry achievement, \( F(4,43) = 6.984, p < .001 \), when all four variables are taken into consideration, however not all variables have the same impact on the final exam grade for the course. The adjusted R squared value was .394, which indicates that 39.4% of the final grade in the class was explained by the model.

The beta weights in table 4.21 (Keystone Algebra included) indicate that out of all of the variables, Gender \( (p<.01) \) and the Keystone Biology \( (p<.01) \) scores had a significantly greater impact on the overall grade in the course, with gender influencing the variation in final grade performance by approximately 16% \( (r^2=.407^2) \) and how well students performed on the Keystone Biology exam having about a 20% \( (r^2=.454^2) \) impact on the variation on final grade performance. The treatment variable of either POGIL or InDGIM and the Keystone Algebra score had a negligible impact on variation on the overall course grade.

Multiple regression illustrated that the one variable that had the largest impact on how students performed on the Final Exam for these Academic Chemistry courses was how they performed on the Keystone Biology exam administered at the end of their Biology course. The Final Exam is 20% of the final grade for the Academic Chemistry course. When the weight of the Final Exam is removed from the multiple regression the variables that had the greatest impact on how students performed overall in the course, regardless of which methodology was used, were a student’s gender and how well they performed on the Keystone Biology exam. Females that scored high on the Keystone Biology exam tended to perform well in the course.

**Summary**

Chapter Four has presented a detailed report of the statistical analysis of this study. Data were analyzed using IBM’s SPSS to perform the ANCOVAs. Descriptive and inferential
statistics were reported. The use of POGIL pedagogy to reduce the alternate conceptions of held by students was not statistically stronger than the comparison or InDGIM. The data illustrated that only one of the null hypotheses could be rejected. It appears that females did slightly better than males on the posttests (in both the lowest and highest quartiles, with placement based on performance on the Keystone Algebra test); the students in the highest quartile still showed improvement, regardless of pedagogical approach; and that Gender and the Keystone Biology exam were better indicators of success in Chemistry than the Keystone Algebra exam.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Chapter Five

Discussion

The purpose of this research is to compare, quantitatively, two inquiry-based programs, the POGIL method, with one that has been created by a small group of teachers, labeled the InDGIM. This study included 48 high school chemistry students enrolled in a suburban high school and utilized a nonequivalent, comparison group, pretest-posttest design with two different normed tests, the ParNoMA and the GALT. The data were analyzed using ANCOVA and revealed that overall, students subjected to POGIL did not statistically outperform those students subjected to InDGIM on the measures analyzed.

Summary of Research Questions and Results

Research Question One and Null Hypothesis One (Ho₁).

Research question one asked: What measurable differences does the use of the POGIL instructional model have on high school students understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared with the InDGIM, the currently used inquiry-based instructional model? The null hypothesis stated that after adjusting for pretest differences, there would be no statistically significant differences on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between high school chemistry students taught using POGIL and students taught using InDGIM, the currently used inquiry-based instructional model. Based on the results of the ANCOVA, the data failed to reject the null hypothesis.

Students in the experimental group, who were taught using POGIL activities, documents and methods, did not earn statistically significant higher posttest scores (after adjusting for
pretest differences) than the comparison group who were taught using the InDGIM approach. The InDGIM group had an increase from their ParNoMA2 pretest to posttest score of 2.65% whereas the POGIL group had an increase of 3.14%. Although the POGIL group showed a greater percent increase, this was not statistically significant and each group had the same overall gain in mean by only 0.29 points.

The InDGIM group showed a decrease in GALT scores by 16.51% while the POGIL group had a decrease of 8.21% in their GALT scores. The scores for both groups decreased from pretest to posttest; however, the scores of the InDGIM group indicated a greater decrease in performance, but this was not statistically significant.

**Research Question Two and Null Hypothesis Two (H₀²).**

The second research question asked: What measurable differences exist between female and male high school chemistry students on the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) regardless of the instructional model used to teach chemistry? The second null hypothesis stated that after adjusting for pretest differences, there would be no statistically significant differences on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) between female and male students taking a high school chemistry course.

The overall main effect due to Gender was not significantly different on posttest scores on either the ParNoMA2 or the GALT, although females scored, on average, higher than males. In addition, there was no statistically significant difference between females and males on gain scores of these two dependent measures. Thus, student differences on either the GALT or
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

ParNoMA2 measures indicated no difference on the Gender factor; therefore, failing to reject null hypothesis number two.

**Research Question Three and Null Hypothesis Three (H03).**

Research question three asked: What measurable differences does the use of the POGIL instructional model or the InDGIM instructional model have on high school females’ understanding of chemistry (ParNoMA2) and logical reasoning (GALT) when compared to high school males’ understanding of chemistry (ParNoMA2) and logical reasoning (GALT)? The third null hypothesis tested for the interaction effect and stated that after adjusting for pretest differences, there would be no statistically significant difference on posttest scores measuring the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) of female and male students who were taught using the POGIL instructional model and female and male students taught using InDGIM, the currently used inquiry-based instructional model. The estimated marginal mean scores for females were higher than the males for both tests in both the comparison (InDGIM) and treatment (POGIL) groups, with the highest scores in the comparison group. The males scored slightly higher on the ParNoMA2 in the comparison group and on the GALT in the treatment group. However, there was no significance between gender and the pedagogical approach for the ParNoMA2 or the GALT. Neither gender nor teaching methodology influenced the outcome of the posttest score on the ParNoMA2 or the GALT. Therefore, the results failed to reject the null hypothesis.

**Research Question Four.**

Research question four asked: What is the differential pattern of performance in the understanding of chemistry (ParNoMA2) and logical reasoning (GALT) for those students
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

falling in the lowest performance quartile (as determined by their Keystone Algebra scores) taught using POGIL, a student-centered cooperative learning instructional model when compared to InDGIM, the currently used inquiry-based instructional model?

Results of students who placed in the lowest quartile on the Pennsylvania Keystone Examination Algebra measure, showed an increase in their ParNoMA2 posttest scores with the POGIL teaching methodology (13.32%), whereas the comparison group, InDGIM, showed a decrease in their posttest scores (23.25%). Females showed a 35.71% increase in scores, whereas males showed a 30% decrease in scores.

For the GALT the lowest quartile students showed a decrease in their posttest scores for both students experiencing the comparison (InDGIM) and the treatment (POGIL) with a 9.34% and 11.95% decrease respectively. Females, regardless of the approach, dropped their posttest scores by 8% and the males dropped by 25.88%.

**Research Question Five and Null Hypothesis Four (Ho4).**

Research question five asked: What measurable differences does the use of the POGIL instructional model have on high school students’ performance on unit tests, when compared with the InDGIM, the currently used inquiry-based instructional model? The null hypothesis stated that there would be no statistically significant differences on unit assessments measuring the understanding of chemistry content of those students taught using the POGIL instructional model when compared to InDGIM, the currently used inquiry-based instructional model.

The Independent Samples t-test was used to analyze differences between the two instructional methodologies on each of the 8 unit assessments and the final exam. The comparison group (InDGIM) statistically (p < .05 to p < .012) outscored the treatment group
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

(POGIL) on almost all of the assessments except two (Empirical Formulas and Waves). Since the results of the events (assessments and final exam) are independent of each other and there is no overlap of the characteristic being measured, the Holm-Bonferroni correction is not necessary. Those tests determined to be significant had an intermediate to large effect size between 0.500 and 0.800 (Cohen, 1988) which is within the zone of desired effect (Hattie, 2009). Based on this analysis, the null hypothesis was rejected for 8 of the 19 measures and practitioners would expect a practical impact on student scores based on those differences.

The simultaneous multiple regression analysis for gender, instructional treatment, Keystone Algebra scores, and Keystone Biology scores performed determined that performance on the Keystone Biology test, taken at the conclusion of their biology course the previous year, was the most significant factor in determining the final exam scores and the final course grade. Gender significantly contributed to explaining the variance in final course grades along with the Keystone Biology, whereas educational methodology (POGIL vs. InDGIM) and the students’ Keystone Algebra scores had no significant contributions to the explanation final exam scores and the final course grade.

Discussion

Many questions were generated after looking at the data. When looking at the outcomes of the posttests, the unit quizzes and tests, the final exam scores, and the final course grades, why did it appear that the InDGIM approach, or comparison, was just as successful as the POGIL approach, or treatment? Why did females appear to outperform males on many of the measures collected throughout the study period? Why were there such drastic decreases in the GALT posttest scores? Why did those students in the upper quartile outperform the students in the other
quartiles on the measures used? Why is the Keystone Biology test a better indicator of overall performance in a chemistry class? Then, finally, why did students using the InDGIM approach outperform POGIL students on most of the assessments administered in class, whereas sometimes the POGIL approach earned better scores.

**Theoretical framework.**

The theoretical ideas that help establish an understanding of this research comes from a social constructivist’s view of learning and the emergent perspective (Becker, Rasmussen, Sweeney, Wawro, Towns & Cole, 2013). Social constructivism is the concept that humans learn knowledge through their interactions with others, thus it is socially situated. In the emergent perspective, “the classroom micro-culture that is established through interactions is considered to be an emergent phenomenon through which meanings are continuously re-negotiated through the course of teacher and students’ interactions” (Becker, et al., 2013, p. 82). Examples of norms that happen in a classroom are that students cooperate to solve problems; that meaningful activities are paramount to correct answers; and that partners should reach consensus when working together (Hershkowitz & Schwarz, 1999). However, there are norms that come about because of a process by which both teachers and students contribute. These are called sociochemical norms, to designate the classroom’s social constructs specific to chemistry “that individuals negotiate in discussions to develop their personal understanding” (Hershkowitz & Schwarz, 1999, p. 150). These sociochemical norms are not predetermined ideas introduced into the classroom from the outside. Instead, these normative understandings are “continually regenerated and modified by the students and the teacher through their ongoing interactions” (Yackel & Cobb, 1996, p. 474). As teachers gain experience on how to instruct students with an inquiry approach to chemistry,
they begin to understand not only how students learn chemistry, they also see the importance of
social context of where the learning takes place.

In both the POGIL and the InDGIM approaches, learning appears to be a constructive
process that occurs while participating in and contributing to the classroom community. Through
the lens of the emergent perspective, students actively construct their understanding of chemistry
as they participate in the practices of the classroom community. According to Cobb and Yackel
(1995) the link “between collective and individual processes in this approach is therefore indirect
in that participation enables and constrains learning, but does not determine it” (p. 19). The
expected social norms that arise during inquiry activities in small group work, according to
Becker et al., (2013), are that students will explain reasoning, illustrate a clear understanding of
terminology and symbolic representations of chemical actions, and arrive at a consensus on
critical thinking questions set forth either by the instructor or the materials. In larger class
discussions students are expected to share their logical reasoning for an explanation, and that the
instructor would provide feedback in a way that would expand upon the ideas of the student and
allow for the complex relationships between the macroscopic, particulate, and symbolic-levels to
become more apparent (Becker et al., 2013).

Two sociochemical norms that are typically seen in inquiry base lessons, in both whole
class and small group activities, are that students will use particulate-level evidence to support
their arguments regarding chemical and physical properties and students will use physical and
mathematical models to justify their claims about chemical action. An example of this is when
students use the concept of particulate movement and spacing to illustrate how internal pressure
of a container might increase with the heating of the molecules inside. The increase in kinetic
energy, or particulate movement, will create more collisions inside the container, thus increasing the pressure. This direct correlation can best be illustrated mathematically as \( \frac{P_1}{T_1} = \frac{P_2}{T_2} \).

**Similarities and differences in InDGIM and POGIL approaches.**

Students in both POGIL and InDGIM appeared to establish sociochemical norms, but the question arises as to why. In the study by Becker et al., (2013), students established, and more importantly, kept coming back to, the sociochemical norm of using the particulate nature of matter to describe changes during a lesson on Thermodynamics. It was surmised by the researchers that the cyclic pattern of particulate-level evidence used by students was related to the structure of the POGIL lessons. According to Becker et al., (2013) the lesson “was consistently structured such that as new concepts such as entropy, or heat capacity were introduced, the workbooks modules would initially include a greater number of questions that asked for explanations or predictions related to various scenarios” (p. 91). Questions that focused on qualitative explanations were often used early, after the introduction of new content, and were designed to elicit prior knowledge rather than applications of new material. However, during the latter portions of POGIL modules, there was often a shift towards the use of mathematical expressions. Once this happens there is a shift towards the interpretation of mathematical expressions as justification for chemical processes as opposed to just particulate level ideas (Becker et al., 2013). This could be why the InDGIM approach was as, or in some cases, more successful than the POGIL method. InDGIM emphasizes a mathematical explanation for the behavior of matter once the basic model is understood. For example, once the Kinetic Theory can be understood using the data collected in the gas lab, for example, then the instructor can use math, such as that found in the combined gas law, \( \frac{P_1 V_1}{T_1} = \frac{P_2 V_2}{T_2} \) to explain how new
situations of pressure or temperature can have an effect on the volume of a container. Thus this movement away from concrete to abstract helps students understand chemical processes (Becker et al., 2013).

Another similarity between POGIL and InDGIM is the heavy reliance on collaboration and cooperative learning. Central to the goals of cooperative learning in science and mathematics education is the enhancement of achievement, problem solving skills, attitudes and values. According to Zakaria and Iksan (2007) cooperative learning is grounded in the belief that learning is most effective when students are actively involved in sharing ideas and work cooperatively to complete academic tasks. Zakaria, Chin, and Daud (2010) concluded that cooperative learning gives more space and opportunities for students to discuss, solve problems, create solutions, provide ideas and help each other. Since cooperative learning enhances scientific skills and increases scientific achievement (Zakaria & Iksan, 2007) then it could be concluded that the similarity in success for each methodology could be due to the fact that each incorporates cooperative learning into their structures.

A major difference between the two approaches is that the POGIL process is highly structured, whereas InDGIM is more dynamic, giving the instructor more flexibility. A study by Chase, Pahkira and Stains (2013) found that POGIL has some characteristics that might make it more challenging for students including the fact that the dominant mode of instruction is not teacher-centered. Students struggle with the role of the instructor as a facilitator of student learning rather than that of the dispenser of knowledge. Generally, students are first introduced to the topic or specific content and are not expected to have worked on any part of the activity prior to class meeting time. The working groups are expected to complete all of the Critical
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Thinking Questions during class (roughly 40 minutes of actual working time) and are not expected to work on any of the Exercises or Problems. With InDGIM the teacher may give direct instruction or be a facilitator; students usually are expected to come in with some background knowledge from previous readings or exercises; and the instructor can be more adaptive to the time frame and increase the time given for students to tackle higher level thinking questions. A possible reason for POGIL’s lack of success could be that it is too prescribed for the students at Falls Fallows High School.

The analysis does show that the comparison, or the InDGIM approach, was more successful than the POGIL method, when it came to helping students understand the concepts of chemistry. InDGIM has already met the definition of guided inquiry. This conformed to the findings of Minner, Levy and Century (2010), which claimed that students showed greater science achievement when involved in guided inquiry lessons when compared to traditional lectures. This would support the idea that the InDGIM approach is just as much of a guided inquiry approach as the POGIL approach.

**Gender differences.**

Although the differences between genders were not statistically significant, it would appear that female subjects outperformed the male subjects on both the ParNoMA2 and the GALT posttest. Female subjects showed a 10.4% increase from pretest to posttest, while male subjects decreased by 9.33% on the ParNoMA2. Both male and female subjects showed a decrease between their pre- and posttest GALT scores, however the males showed a larger decrease of 18.86%, whereas the females’ scores dropped by only 8.45%. Although the scores for both groups dropped on the GALT the final scores on the posttest were almost identical
between genders.

Regarding adjusted marginal means and gain scores, females outpaced the males as well. The females had higher marginal mean scores on both the ParNoMA2 and the GALT. Males had lower gain scores than females on the ParNoMA2, and the same pattern occurred with the GALT. In addition, the estimated marginal mean scores for females were higher than the males for both tests in both the comparison (InDGIM) and treatment (POGIL) groups, with the highest scores in the comparison group. The males scored slightly higher on the ParNoMA2 in the comparison group and on the GALT in the treatment group.

Looking at the demographics for Falls Fallow High School, the results of this study mirrors how females outperform males on other standardized tests as well. On the latest Keystone English Literature exam, 87% of the females scored proficient or advanced, compared to only 82% of the males (Springfield Township High School, 2018). The numbers were lower, but the trend held true, on the Keystone Algebra test, with 77% of the females scoring at least proficient, compared to only 72% of the males reaching proficiency (Springfield Township High School, 2018). At the state level, females at high schools in Pennsylvania outperformed males on the Keystone English Literature assessment 76% to 65%, but performed about equally as well as males on the Keystone Algebra test, 61% to 58% (Springfield Township High School, 2018).

Gibbs, Fergusson, and Horwood (2008) examined the educational achievement of 1265 individuals from birth to age 25 and found that there was a tendency for females to outperform males on standardized tests. The differences could not be explained by differences in cognitive ability as the students in the study were tested at similar IQ levels (Gibbs, Fergusson, & Horwood, 2008). Instead, the researchers determined that males were prone to being more
inattentive, restless, distracted, and antisocial than their female counterparts. When the associations between gender and measures of educational achievement were adjusted for teacher ratings of classroom behavior, gender differences were substantially reduced (Gibbs, Fergusson, & Horwood, 2008). Downey and Vogt Yuan (2005) reached similar conclusions in that the major reason for gender differences in high school performance was due to poor classroom behavior of males. Classroom behavior was not a variable collected by the instructor or the researcher of this study; however, it might be one variable used to explain why females outperformed the males on the standardized assessments.

Another reason for the slight differences might be explained in spatial ability between males and females. According to Yezierski and Birk (2006) in order to understand states of matter students must understand the spatial relationship between molecules and that the molecular movement seen in these models is such that the relative distance between them is appropriate for each phase of matter. Students with greater spatial visualization abilities might be able to create a stronger mental model of this particle movement and thus perform better on tasks requiring this ability (Yezierski & Birk, 2006). Baker and Tally (1972) found that success in chemistry is more strongly correlated to spatial visualization abilities than to general academic ability and Wu, Krajcik, and Soloway (2001) illustrated that constructing visual connections between abstract representations is important in making conceptual connections for chemistry success.

Chemistry educators can positively affect this spatial visualization ability by the types of interventions used in the classroom since most spatial ability differences between genders is attributed to experiential differences and not intellectual ones (Hamilton, 1998; Yezierski &
Birk, 2006). Even minor treatments, such as models and animations, according to Yezierski and Birk (2006), can have an impact since those treatments “are somehow providing background information previously possessed by males and not females” (p. 959). The exposure to both the InDGIM and the POGIL method in this study might have provided the background needed for the females to outperform the males on the two posttests. These results are consistent with Piburn, Reynolds, Leedy, McAuliffe, Birk, and Johnson (2002) in that females were able to close the gender gap regarding chemistry concepts following relatively minor treatments.

Linn and Hyde (1989) stated that educational “environments that encourage and reward the cooperative behavior that is often necessary in scientific investigations could be harnessed to minimize gender differences” (p. 25). Those environments that provide scaffolding so that participants can acquire new skills, encourage the sharing of ideas, and provide feedback and encouragement may increase persistence among those who are less confident (Linn & Hyde, 1989). The methodologies used to teach chemistry in both the comparison (InDGIM) and the treatment (POGIL) groups involve cooperative group work during the model development stages. Females may perform better than males with this cooperative approach.

Levine, Vailyeva, Lourenco, Newcombe and Huttenlocher (2005) determined that differences in gender spatial visualization was highly correlated to the socioeconomic status of students, with a greater disparity between genders with higher socioeconomic status and hardly any between genders with lower socioeconomic status (Levine, Vailyeva, Lourenco, Newcombe & Huttenlocher, 2005).

According to the National School Lunch Program, 20% of the students attending Falls Fallow High School are eligible for free and reduced lunches (Springfield Township High
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

School, 2018). Out of the current population, 16.8% of the students qualify for free lunch, with a family income under $15,171, and 3.2% of the current total student population qualifies for reduced lunch, with a family income below $21,590 (Springfield Township High School, 2018). Both of these incomes fall below the basic poverty line. If the 47 students in the study represent a cross-section of the student population at Falls Fallow High School, then 20% of them, or about nine, should fall into the category of low socioeconomic status.

Although this information was not collected directly from the students that participated in the study, the slight differences between genders on the standardized tests might be influenced by the socioeconomic status of the students. Those students with a higher socioeconomic status might have been able to afford the freedom to experience their environment more fully, which increases spatial visualization (Levine et al, 2005).

The decrease of the GALT scores.

Bunce (1993) stated that the GALT’s five formal operational reasoning modes have been demonstrated to predict critical thinking abilities and grades assigned by teachers in science and mathematics for students in grades 9-12. Bitner (1991) attached the prediction to the 8th grade students and Roadrangka (1982) established the validity of the GALT on a student sample that varied in ages from 6th grade through college.

The test includes questions related to mass and volume conservation, proportional reasoning, correlational reasoning, control of experimental variables, probabilistic reasoning and combinatorial reasoning (Roadrangka, 1982). The first of these skills (mass/volume conservation) is typically mastered at the concrete operational level, whereas all others correspond to the formal domain (Roadrangka, 1982). In terms of operational level most
introductory college students seem to be at the transitional or formal operational levels (Bunce, 1993). Does this mean that most high school students are functioning at the concrete operational level with only a few at a transitional level? Science educators have stated that an activity based science curriculum can enhance the development of logical thinking as measured by the GALT. Bunce (1993) attempted to test this idea by measuring the college students’ scores before and after a one-semester, non-science major’s course, but there was no significant correlation between the pre and posttest scores. Bunce (1993) concluded that change takes time and that a 13-week course may not be long enough to bring about measurable improvements in scores. If 13 weeks was not enough time to see improvement in logical thinking skills for college students, then 18 weeks may not be enough to see any significant improvement in scores for high school students.

Another connection has to do with modeling and formal operational development or abstract reasoning. POGIL’s premise is to create a concrete or conceptual model that helps students develop processing skills (POGIL Project, 2014). A concrete model is a simplified representation of a system or a working scale of a prototype, whereas a conceptual model focuses on an understanding of how a process works and can be expressed as visual or symbolic representations as well as through verbal descriptions or analogies (Bryce, Baliga, De Nesnera, Flack, Goetz, Tarjan, Wade, Yovovich, Baumgart, Bard, Ash, Parker, & Gilbert, 2016). Modeling of chemistry concepts can help students develop cognitively, however, the pace that this develops at is built upon the mental models that students bring to the classroom.

All students come to a class with a particular mental model. According to Bryce, Baliga,
A person’s conceptual understanding of a process or relationship (i.e., mental model) directly informs his or her creation of a model, whether that model is concrete, conceptual, or mathematical. Through testing and experience, these models can be updated to reflect reality more accurately. As students iteratively draft scientific models, they inevitably modify their underlying mental models through analysis. In a classroom context, students refine their own mental models as they observe, analyze, and discuss the modeling work of others (p. 36).

In practice, this approach reflects and is an extension of the scientific method (Bryce et al, 2016).

The use of models for K-12 students should progress from simple to complex applications as classroom activities transition from demonstrations by the instructor toward student-directed inquiry (Bryce et al, 2016). It is known that other science teachers at Falls Fallow High School do use conceptual and concrete models in Biology and Environmental Science (as well as mathematical) but it has not been determined if the same can be said for teachers at the lower grades.

According to Bryce et al, (2016) students in elementary school should progress from recognizing models as a tool that can be used to explain familiar structures or scientific processes to building or revising simple models to design solutions to problems or represent phenomena, describe processes, explain relationships and make predictions. As students advance to middle school the use of models should expand to predicting and testing more abstract phenomena and students should undertake increasingly more open-ended investigations stronger models (Bryce
et al, 2016). Finally, in high school, students should be able to construct and use models for more advanced prediction and to represent interactions between variables within a system with inquiry at this stage should be largely focused on critical evaluation and comparison of different models to improve predictions and explanatory power (Bryce et al, 2016). In this way models become constructs that are more than just ways to describe an object.

Students should proceed through these steps as they go through school, for revising models “provides students with metacognitive opportunities, they better understand their own thinking” (Bryce et al, 2016, p. 39). It is possible that this lack of experience by the subjects in the study may have delayed or diminished the onset of formal operational thinking at the secondary level (Lahti, 2013). Students need to be able to work through and see the connections between concrete, conceptual, and abstract models. Presenting ideas through different activities with more advanced modeling would seem to lend themselves well to developing formal thought.

Finally, the decrease in the GALT scores could be the result of testing fatigue. The posttest was given after the ParNoMA2 at the end of the course, roughly one week before the final exam. The students did receive some form of credit for completing the test online, however, the points awarded were for completion only, not for accuracy or improvement. There is a possibility that the students did not take the posttest seriously since they might not have seen a connection between the GALT posttest, the course, and grades.

Performance of the upper quartile sample.

The results of this study in which high achieving students performed better on the posttest compared to lower achieving students is similar to the results of a study by Kahn, Hussain, Ali,
Majoka, and Ramzan (2011). Their study found that inquiry-based instruction improved students’ achievement in the subject of chemistry at the secondary level, as measured with pre and posttest scores, with higher achievement gains for the group of high achievers, although they failed to explain why this was the case (Kahn, Hussain, Ali, Majoka, & Ramzan, 2011).

A cooperative learning study by Nattiv (1994) showed that students that fell into the higher ability quartile performed better statistically on their posttests than did the students that fell into the lowest quartile. The study indicated that high achieving students preferred a cooperative learning approach in that they were able to take advantage of their teams and get the needed support they required from their classmates as opposed to waiting for the teacher (Nattiv, 1994). The conclusion of Nattiv’s study was that high achieving students were “learning to share their expertise rather than to hoard it,” whereas lower-ability students would prefer to get help from the teacher (1994, p. 292). This could explain why in the current study those students in the highest quartile showed an increase in their posttest scores because both inquiry approaches, InDGIM and POGIL, played to their strengths as students and allowed them to interact with other students more than they could in a traditional lecture style course.

Lister and Ansalone (2006) illustrated that the modality of instruction, along with student attitude and involvement in their own learning, is an important factor in academic achievement of students. Their study (2006) revealed that “both achievement and attitude to learning are enhanced when tactual/kinesthetic strategies are employed” (p. 27) and that those students that are highly engaged in the learning process, those high achievers, usually perform better. Both instructional approaches in the current study, POGIL and InDGIM, use physical models and data collected from lab experiences, which may promote a more positive attitude among students
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

leading to a greater engagement with their education. These high achievers are more engaged with the course work and, according to Tamir (1993), they excel more with tasks that require greater cognitive abilities, having a greater preference for questioning than rote memorization. InDGIM and POGIL provide opportunities for students to discuss answers to questions that should lead them to consider the general ideas in question and to construct their own understanding of chemistry concepts. Ideas are formulated and shared by individuals and groups. This approach may have favored the higher achieving students more than their lower achieving counterparts.

In addition, spatial ability could come into play for high achieving students. According to Weckbacher and Okamoto (2012) there are many studies that show the relationship between high achieving students and their spatial ability, especially in the realm of mathematics. Weckbacher and Okamoto state that high achieving students have skills, “such as pattern recognition, visual transformation, and mental rotation (and) are spatially adept” (2012, p. 52). These high achievers tend to develop schematic rather than pictorial representations when solving problems using “spatial images that incorporate essential elements of a problem to achieve success” (Weckbacher & Okamoto, 2012, p. 52). These skills are essential when understanding “molecular geometry, kinetic molecular theory, stoichiometry represented at the particulate level, and crystal structure” (Harle & Towns, 2011, p. 355). High achievers are able to use a wide variety of methods to represent concepts in chemistry. Their spatial abilities certainly help with many specific chemistry ideas. Successful chemistry students need to be able to generate “representations that express their understanding of underlying entities and process; use these representations to explain chemical phenomena at the observable, physical levels in terms of
chemistry at the particulate level; and identify and analyze features of representations (such as a peak on a graph) and use them to explain, draw inferences, and make predictions about chemical phenomena or concepts” (Harle & Towns, 2011, p. 356). It might be that those high achieving students were able to use their spatial reasoning on the ParNoMA2 and the GALT and score better on the posttests than their lower achieving counterparts.

**Success with POGIL lessons.**

In this study there were two assessments where the POGIL methodology yielded better results than the InDGIM. The first of the assessments was on Light Waves & DeBroglie’s Wavelengths, and the second was on Empirical formulas. The question arises as to why students exposed to the POGIL lesson plans performed statistically better than their counterparts on these two assessments and none of the others. The lesson for wave behavior and empirical formulas can be found in Appendices K and L. These activities include a faculty provided-model with related content and a specific problem with a defined set of questions for small groups to solve and answer with little guidance from the instructor. Each activity begins with some very straightforward questions pertaining to the data presented about the subject (chemical formulas or light). The answers are obtained by interpreting a table of information. The second set of questions uses a slightly different model but requires the student to use the knowledge from the previous section to answer these questions. This is a pedagogical technique known as scaffolding. Scaffolding refers to a variety of instructional techniques used to move students progressively toward a stronger understanding, and, ultimately, greater independence in the learning process. In order to answer the questions in parts one and two of this activity, students must only use the models presented to them. Once a new model is presented, it can be
incorporated into the student’s learning and the discussions that occur within the group. This helps them develop group process skills while they are gaining content knowledge and begin processing the information at a higher educational level (P. Brown, 2010). According to the POGIL Project (2014) one of the most defining characteristics of this process is that students spend the majority of class time working in small groups on activities that require higher-order thinking such as synthesis, analysis, and the integration of ideas with previously learned concepts. The role of the instructor is to serve as a facilitator who listens to the discussion and intervenes at appropriate times to guide student learning and provide students with a framework to collaborate effectively as a team (POGIL Project, 2014). Without this framework, students may not know how to apply the collective knowledge of the group to a new situation involving similar data.

This is slightly different from InDGIM. For this particular subject matter, waves, the instructor began the lesson with a series of demonstrations on color and light behavior, which he followed up with by defining terms, administering notes, and laying out concepts before moving the students into a guided inquiry lab on light behavior. During the POGIL activity students had to actively work to master the material and formulate a deeper understanding of the content without having new terminology defined for them. Built into the experience is the support of a variety of important process skills, including communication, teamwork, and critical thinking, which, according to the POGIL Project (2014) translates to a more complete understanding of the entire concept, and a lasting understanding of the material. If this were the case then every POGIL activity should have resulted in higher scores on the assessments when compared with InDGIM.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Another key component to this variation was the teacher’s familiarity with this particular POGIL activity. In discussions with Mr. B. he stated that this was a very well thought out activity, one that he has done on numerous occasions over the past five years. Content and pedagogical content knowledge are critical factors in creating and sustaining instruction that promotes student discourse and understanding (Gagnon & Maccini, 2007), so familiarity with this POGIL activity might have led to the higher scores on the assessment on wave behavior. Also the number of years teaching affects a teacher’s use of a new instructional activity, where new teachers, those with five years or less of experience, were not as good at implementing new material as were their counterparts with more years of teaching (Gagnon & Maccini, 2007). Mr. B. has more than five years of experience teaching chemistry and using this POGIL activity in his classes. This could help explain why the POGIL group outperformed the InDGIM group on these two unit assessments.

Finally, it should be noted that Chase, Pakhira and Stains (2013) indicated that instructors “implementing POGIL should not presume to observe all expected outcomes after the first implementation (for) a delay may exist between the first implementation and positive outcomes for students” (p. 415). Their study (2013) indicated that using the POGIL methodology didn’t negatively affect students’ learning when compared to traditional teacher-centered learning environments, and had the ability to enhance it, but that it might not happen immediately. This delayed effect might help explain why the POGIL method didn’t produce the predicted outcomes.

Keystone biology exam as an indicator of chemistry success.

There are many chemistry concepts taught in biology. Biology students are subjected to
atomic structure and the Periodic Table as they study biochemistry; the chemistry of water as they study the cells; the concept of pH as they study enzyme activity; and basic chemical reactions as they study how macromolecules interact with each other. Table 5.1 illustrates some of the standards (the chemistry of life) that students must know for the Keystone Biology Exam (Pennsylvania Department of Education, 2015). The basics of chemistry are taught in Biology and that is why the results illustrate the importance of the Keystone Biology exam in predicting Academic success.

Table 5.1. Biology Keystone Assessment Anchors and Eligible Content as Developed by the Pennsylvania Department of Education

<table>
<thead>
<tr>
<th>Assessment Anchor Code</th>
<th>Eligible Content Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIO.A.2.1.1</td>
<td>Describe the unique properties of water and how these properties support life on Earth (e.g., freezing point, high specific heat, cohesion)</td>
</tr>
<tr>
<td>BIO.A.2.2.1</td>
<td>Explain how carbon is uniquely suited to form biological macromolecules.</td>
</tr>
<tr>
<td>BIO.A.2.2.2</td>
<td>Describe how biological macromolecules form from monomers.</td>
</tr>
<tr>
<td>BIO.A.2.2.3</td>
<td>Compare the structure and function of carbohydrates, lipids, proteins, and nucleic acids in organisms.</td>
</tr>
<tr>
<td>BIO.A.2.3.1</td>
<td>Describe the role of an enzyme as a catalyst in regulating a specific biochemical reaction.</td>
</tr>
<tr>
<td>BIO.A.2.3.2</td>
<td>Explain how factors such as pH, temperature, and concentration levels can affect enzyme function.</td>
</tr>
</tbody>
</table>

Along with chemistry concepts algebraic concepts are also applied when studying biology content. Direct variation (and thus the linear equation) appears while studying the heart and breathing. The relationship among the heart cardiac output (C), the heart stroke volume (V), and the heart rate can be expressed as $C = V \times R$ (Horak, 2005). Each variable in this relationship
represents a rate. The number of breaths a rescuer administers to a victim in a given number of minutes can be represented by the linear equation \( n = 2 + 10t \), in which \( t \) is the number of minutes during which rescue breathing is administered and \( n \) is the number of breaths administered after \( t \) minutes (Horak, 2005). According to Horak, 2005 there are other linear relationships that occur such as the one between the diameter of a twig and its length, the rate at which nerve impulses are conducted along a nerve fiber and the diameter of the nerve, and the caloric requirements of most animals and their size. Horak (2005), also points out non-linear relationships such as the resistance of blood flowing through a blood vessel. This can be expressed as \( R = 1/r^4 \), where \( R \) is the resistance of the blood flow and \( r \) is the radius of the radius of the blood vessel (an inversely proportional relationship). Another relationship is cell size to cell volume. If the cell is cube shaped, and all sides are equal so \( X = \text{length} = \text{width} = \text{height} \), then the surface area of the cube is the area of each side \( (X^2) \) times the number of sides (6) which equals \( 6X^2 \). The volume of the cube would be \( X^3 \). What happens to the surface to volume ratio as the cell grows? \( \text{SA:V} = 6X^2 / X^3 \). The student should see that as the cell size grows volume increases faster than the surface area. Considering the amount of chemistry and algebra skills required to understand biology, it is possible that scores on the Keystone Biology exam might be a better predictor, than the Keystone Algebra exam, of how students might perform in chemistry.

Teacher effect.

As mentioned earlier, sociochemical norms had a major impact on student learning. The analysis of these sociochemical norms indicates that the instructor plays an important role in establishing the tone of the classroom environment and establishing norms for chemical aspects of student activities. It further highlights the significance of the teacher’s own personal beliefs.
and values, and their own chemical knowledge and understanding (Yackel & Cobb, 1996). In this way, the critical and central role of the teacher as a representative of the chemistry community cannot be underscored. It must be stated that the instructor played a large role in the outcome of this study.

A conglomerate of POGIL users has developed and traced ways that teachers have adopted this pedagogy. Effective POGIL users typically have a strong pedagogical content knowledge (Shulman, 1986), particularly with an inquiry-based orientation toward science teaching (Magnusson, Krajcik, & Borko, 1999), a constructivist view of learning (Wheatley, 1991), a focus on aligned instructional strategies for addressing students’ learning needs (Magnusson, Krajcik, & Borko, 1999), and a practice of reflection (Loughran, 2002; and Sellars, 2012).

Through various adoption models and assessments of implementation, findings support that there are essential features to implementing POGIL instruction (Bauer & Cole, 2012; Bunce et al., 2010; Hanson, 2006), and research by Daubenmire et al., (2015) indicates how instructors implement POGIL has an impact on student outcomes. Their research indicates that the teacher still makes a difference, even when implementing a research- and reform-based instructional approach (Daubenmire et al., 2015).

When comparing students experiencing POGIL to those that were experiencing only lecture, the POGIL students outperformed their non-POGIL peers in college chemistry. There were, however, differences between the different POGIL groups, as measured by the American Chemical Society Conceptual scores. Daubenmire et al., (2015) determined that these differences were the results of the differences in how the instructors interacted with their students. Asking
more questions as a way to guide students, afforded students more practice using reasoning skills, and more practice with these skills lead to greater conceptual understanding of chemistry (Daubenmire et al., 2015). As a result, even in reform-based, student-centered instruction, the style of approach of the instructor remains critical to shaping what, how, and how much students learn.

This dialogue between teacher and student is very important. Warfa, Roehrig, Schneider, and Nyachwaya (2014) illustrated the importance of teacher-student discourse in the development of fluency in chemistry. They stated that there are two kinds of teacher-initiated discourse: dialogic and monologic (Warfa, Roehrig, Schneider & Nyachwaya, 2014). Dialogic discourse occurred “as a result of direct interaction with students in their small group settings and often required the course instructor to make several scaffolding moves in response to students’ acts” (Warfa, Roehrig, Scheider & Nyachwaya, 2014, p. 787) and is student-centered. Monologic discourse “occurred during whole class discussions in which the course instructor first presented data or macroscopic demonstrations of a concept and then prompted a whole-class discussion on that same concept” and is teacher-centered (Warfa, Roehrig, Scheider & Nyachwaya, 2014, p. 787). As was witnessed during the implementation of both the POGIL and InDGIM methodologies, the instructor, Mr. B., often initiated dialogic discourse when interacting with individual student groups to check on their progress and solutions to the analysis questions in different activities. His scaffolding techniques during these interactions included ways of communicating to the students what counts as acceptable particulate representation of chemicals ideas. There are certain patterns of instructor facilitation, such as questioning, re-voicing/repeating which leads to re-voicing/expanding that are particularly useful in generating a
more explicit discussion of relationships among levels of chemical representation (Becker, Stanford, Towns, & Coles, 2015). This approach, encouraged through the inquiry process, could help explain why there was no statistical difference between the treatment and comparison groups in this study, in that a teacher that engaged them in dialogic discourse taught both groups. Warfa et al., (2014) determined that students exposed to a higher degree of dialogic discourse had a higher degree of chemistry fluency and retention of information.

Lee (2011) created a developmental rubric on the use of inquiry guided learning that reflects the varied conceptualizations of teaching that instructors hold. The rubric “recognizes that how instructors implement inquiry guided learning depends as much on their own frequently unexamined assumptions as their instructional response to the developmental level of their students” (Lee, 2012, p. 10). According to Lee (2012), an instructor might find herself in one of three stages: experimenting, developing, or committed. Table 5.2 illustrates the differences of each teaching stage. An instructor, who sees herself as a presenter of knowledge and trusts primarily, her own control over knowledge delivery, will implement inquiry guided learning quite differently from an instructor who sees herself as a collaborator with students in the process of inquiry and trusts the process of inquiry itself as a force in learning regardless of the level of students. Although it was determined where each activity fell on the spectrum of inquiry, this was not done so for the instructor, Mr. B.

<table>
<thead>
<tr>
<th>Experimenting Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Some inquiry as a mode of assessment but only after explicit preparation of students using traditional instructional methods.</td>
</tr>
<tr>
<td>• Acquisition of knowledge through explicit instruction with some experimentation engaging students in the skills of inquiry through isolated learning activities.</td>
</tr>
<tr>
<td>• Primary source of trust is in instructor control over knowledge delivery.</td>
</tr>
<tr>
<td>• Instructor exhibits little tolerance for uncertainty beyond isolated and carefully controlled opportunities for student engagement.</td>
</tr>
<tr>
<td>• Instructor functions chiefly as an organizer and presenter of knowledge.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Developing Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Inquiry as a mode of learning but often after explicit preparation of students using more traditional instructional methods.</td>
</tr>
<tr>
<td>• Separate development of the skills of inquiry and the acquisition of knowledge through explicit instruction.</td>
</tr>
<tr>
<td>• Balance of challenge and support in ways appropriate to the developmental level of students; mechanisms of support are visible.</td>
</tr>
<tr>
<td>• Primary source of trust is in the guidance of the instructor with guidance taking a variety of forms.</td>
</tr>
<tr>
<td>• Instructor exhibits some tolerance for uncertainty within anticipated boundaries of student performance. Instructor functions chiefly as a guide to students during the process of inquiry</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Committed Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Inquiry is the dominant mode of learning and the primary stimulus for knowledge acquisition.</td>
</tr>
<tr>
<td>• Seamless development of the skills of inquiry and the acquisition of knowledge through the process of inquiry itself.</td>
</tr>
<tr>
<td>• Skillful, and often invisible, balance of challenge and support in ways appropriate to the developmental level of students; enables students to function with a high degree of independence.</td>
</tr>
<tr>
<td>• Primary source of trust is in the process of inquiry as a mode of learning and the outcomes and products of inquiry as credible or valid assessment.</td>
</tr>
<tr>
<td>• Instructor exhibits a tolerance for uncertainty in the inquiry process and openness to unexpected directions set by students.</td>
</tr>
<tr>
<td>• Instructor functions chiefly as a collaborator with students in the process of inquiry.</td>
</tr>
</tbody>
</table>
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

According to Wenglinsky (2001) it appears that there are various teacher characteristics that are related to student achievement when class size and socioeconomic status are taken into account. In particular, a teacher with a chemistry major; one that receives professional development in higher-order thinking skills; one that receives professional development in the aspects and importance of diversity; one that uses hands-on learning; and one that encourages higher-order thinking skills in class is the type of teacher that produces strong results in the classroom (Wenglinsky, 2001). Although he has a degree in Biology, which requires minimal training in Chemistry, Mr. B. doesn’t have a degree in chemistry. Instead his undergraduate majors were Religious Studies and Biology and his graduate majors were in Sustainable Systems and Secondary Science Education. Although this investigation didn’t look at the impact of teacher differences in training and background, this could be a factor may have influenced the results of this study.

Limitations

This study utilized a nonrandomized pretest-posttest design. The lack of randomization is a limitation of this study. Lack of randomization was controlled for by selecting students who were randomly placed by computer into one of the four Academic Chemistry sections offered at Falls Fallows High School. Participation in the study was voluntary, and almost all students volunteered. Pretest effects on posttest scores were controlled by the use of ANCOVA in which the pretest for both ParNoMA2 and GALT were used as the covariant.

The number of participants for the study was low. Although the threat to internal validity based on attrition was not an issue because only four students, out of the original 52 students, did not complete the study and one didn’t complete all aspects of the study (a dropout rate of
9.80%), other potential sources of internal invalidity are still present. In order to have confidence that the study results are representative of a larger population, it is critically important to have a sufficient and representative number of randomly selected participants. For a 95% confidence level, a good estimate of the margin of error is given by \( \frac{1}{\sqrt{N}} \), where \( N \) is the number of participants or sample size (Niles, 2017). Using this formula, the margin of error for the number of participants in this study is estimated to be 14.5%. A sample that is too small may prevent the findings from being extrapolated to another educational setting or to an educational setting of a different size; therefore, increasing the potential effect on population and ecological external validity.

The full impact of POGIL methods would be better measured if the students in the experimental group had only used the POGIL activities for the duration of the study. The POGIL activities available for this study did not cover all the concepts covered in the Academic Chemistry course nor did they expand across a unit of study. Many of the activities are mostly stand-alone activities that require the instructor to infuse other activities to help solidify a concept. For example, there is only one activity involving gas variables with POGIL whereas the InDGIM has lab activities that cover Boyle’s Law, Charles’ Law, and the Combined Gas Law, multiple simulations, and mathematical problems assigned to students for homework. The POGIL activity was used in the experimental group as an introductory lesson to help lay the foundation of the concepts, however, non-POGIL activities were used for the remainder of the unit of study. A POGIL approach can be implemented in a variety of ways, from completely converting the classroom to a POGIL learning environment to scheduling POGIL experiences on a relatively infrequent but regular basis throughout the course (Moog & Spencer, 2008). For this
study POGIL was used to introduce as many concepts as possible. At the time of the study there were no POGIL activities developed to introduce nuclear chemistry, a concept covered in the Academic Chemistry course. This may have had an impact on the study.

Although trained in the POGIL approach Mr. B. has been teaching Academic Chemistry for the last five years using InDGIM and this might have biased his approach in the implementation of POGIL methods. During short debriefings Mr. B. shared his thoughts that in some cases the POGIL activities were overly complex (such as the Polyatomic Ions activity), whereas other concepts were better covered by the InDGIM approach (the activities pertaining to the Gas Laws). The perceived POGIL shortcomings might have influenced how Mr. B. approached the activity when he taught it again in a separate Academic Chemistry class. Having an independent observer and Mr. B. rate the lessons was an attempt to reduce this teaching bias.

There was also the chance of scorer bias on all the assessments that were administered in the chemistry classes. Although the standard rubric awards points for the correct mathematical answer, students may not earn all points possible unless they set the problem up correctly; showed that they were able to solve the problem correctly; wrote the answer in proper scientific notation; and wrote the answer with the proper number of significant figures. Although this has been the general approach to scoring these assessments, there is a definite possibility for scorer bias since not all mathematical problems in chemistry need to be solved using the same number of steps. This could lead to more points being awarded on one assessment than the other. In an attempt to eliminate this problem, Mr. B. would grade the same question on each assessment for both the POGIL and the InDGIM group before moving on to the next question (i.e., question one on all tests, then question two on all tests, etc.).
Further research

**Attitudes.**

According to one of the theoretical concepts behind cooperative and collaborative learning (a principle found in both POGIL and InDGIM), an expected outcome of this process is an increase in students’ attitude or thoughts and feelings toward chemistry (Chase, et al., 2013). Previous studies by Farrell, Moog, and Spencer (1999) indicated that students had positive attitudes towards the POGIL method and that cooperative learning increased their achievement. Eberlien, et al. (2008) reported that instructors and students liked the POGIL environment more and Schroeder and Greenbowe (2008) pointed out that the POGIL process increased students’ self-confidence. P. Brown (2010) reported that almost all of the students in the study said that POGIL was an effective and useful way to study chemistry, and, according to Conway (2014), it is the POGIL approach that changed students’ attitudes towards chemistry. The current study did not investigate this variable, thus the extent of the impact of either methodology on students’ attitude toward chemistry is uncertain. However, this could be done in future studies at Falls Fallow High School.

Chase, et al., (2013) did conduct a study with college students that measured the impact of the POGIL process on students’ grades; retention; attitudes towards chemistry; attitudes toward their learning environment; and self-efficacy by using the Attitude toward the Subject of Chemistry Inventory (ASCI) and the Chemistry Attitude and Experience Questionnaire (CAEQ). These two surveys can be quickly used by an instructor to help measure students’ attitudes regarding chemistry.

The 20 item ASCI evaluates students’ attitude towards chemistry along five variable:
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

fear, interest and utility, intellectual accessibility, anxiety, and emotional satisfaction and can be used when studying longitudinal change or when making comparisons between groups (Bauer, 2008). From a diagnostic perspective using the ASCI as a pretest and posttest allows an instructor to understand the role the course has in changing students’ attitudes towards the subject of chemistry, with positive changes being the preferred outcome and negative changes being the cause for introspection (Bauer, 2008). Research by Brandriet, Xu, Bretz, and Lewis (2011) concluded, with the assistance of the ASCI, that “A and B students’ attitudes significantly increased from the beginning to end of a general chemistry semester, while D and F students’ attitudes significantly decreases, and C students’ attitudes did not change” (p. 277). Although correlation between attitudes and grades does not illustrate causation, it does seem to illustrate that attitudes and success in chemistry do interact with one another. The ASCI could also be useful for detecting differences in attitudes between genders. Generally, females seemed to have less favorable attitudes towards the subject of chemistry than males (Brandriet, Xu, Bretz, & Lewis, 2011) which was not evident through observation by the researchers. The use of ASCI could quickly detect differences in the attitudes held by a student population thus allowing modification to the course as needed, or supplying evidence that a particular population or gender’s attitudes are/are not linked to their attitudes regarding the course.

The CAEQ surveys attitudes towards chemists, skills of chemists, attitudes towards chemistry in society, leisure interest in chemistry, and career interests in chemistry from a societal perspective by using roughly 70 questions that focus on the “antecedents of attitude towards enrolling in chemistry: namely, their learning experiences, attitude-towards-chemistry and chemistry self-efficacy” (Coll, Dalgety, & Salter, 2002, p. 24). When first administered by
Coll, Dalgety, and Salter (2002) the survey revealed that their college level students preferred a more structured learning environment and the conclusion reached was that the earlier school experiences, and the “associated mode of assessment, may influence their expectation of appropriate pedagogy” (p. 26). When it was used by Chase, et al., (2013) it was concluded that students exposed to the POGIL methodology had more positive attitudes regarding the time spent in lecture, discussions and laboratory settings than those in the non-POGIL comparison group.

Using the ASCI and the CAEQ with students at Falls Fallow High School in the future, something that has not been done before, might help the Chemistry instructors determine the proper educational environment for students, whether that is through the use of POGIL activities or the InDGIM. By taking stock of the attitudes of the students, the instructors may be able to better facilitate classroom practices that lead to greater student success.

**Keystone Biology Quartiles and Methodology.**

Another area for research is looking at the relationship between these two pedagogical approaches (InGDIM and POGIL) and how students rank in different quartiles based on their Keystone Biology scores. The impact of this exam on the performance of these students in chemistry came as quite a shock to this researcher. Knowing that students with higher Keystone Biology scores perform better in an Academic Chemistry class, is there a correlation between that score and a particular methodology? This is one thing that was not looked at during this study. If students were grouped based on their Keystone Biology performance would one approach meet their needs better than another?

Was there some form of gender bias taking place for these students? All of the students
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

at Falls Fallow High School take Biology before they take Chemistry (whether it is Academic or Honors levels courses). Does the gender of the teacher a student has in Biology affect how he/she will perform on the Keystone Biology test, and thus, how they perform in Academic Chemistry?

Conclusion

Both the POGIL approach and the InDGIM resulted in the same relative gains for the students in all of the Academic Chemistry courses. The similarities of these two programs far exceed any perceived differences.

One of the similarities that each approach promotes is the idea of working on problems in small cooperative groups. There is evidence that small group-centered learning leads to improved student performance as well as quality and efficiency of instruction (Springer, Stanne, & Donovan, 1999). Students placed in small groups actively discuss and process information related to the task. The instructor facilitates meaningful learning by motivating students and providing them with opportunities to participate in the learning process (Jones, 2006) whether assigned roles (POGIL) or not (InDGIM). Previous studies involving cooperative learning (Bilgin & Geban, 2006; Köse, Şahin & Gezer, 2010) have found that the particular cooperative learning approaches, found in POGIL and InDGIM, have a more positive impact on student achievement than the traditional lecture approach. The extent to which the cooperative learning aspect of both methods was responsible for student gain is difficult to determine.

These small groups, in both pedagogical approaches, might help students to acquire related process skills such as communication, reading, writing, leadership, time management, and problem solving. Individual or solitary learning, as emphasized by the traditional educational
approach, with the teacher lecturing the class, seldom offers students these much needed collaborative skills found in many present-day jobs that are increasingly team-centered and demand the application of process skills already acquired while a student (Qureshi, Vishnumolakala, Southam, & Treagust, 2016).

Both pedagogical approaches encourage specific sociochemical norms. Within inquiry-oriented classrooms, students explain their reasoning, students listen to and try to make sense of other students’ reasoning, and students indicate agreement or disagreement with the reasoning of others (Becker, Rasmussen, Sweeney, Wawro, Towns, & Cole, 2013). These sociochemical norms are critical parts of the classroom environment, for both POGIL and InDGIM, in that they shape student views of what counts as “appropriate justification in chemistry, how different types of representations should be interpreted, and what counts as ‘good’ explanations in chemistry” (Becker, et al., 2013, p. 83). These sociochemical norms are the same for each approach.

The POGIL and the InDGIM approaches use various strategies for helping students understand chemistry including the idea of modeling. When students were given an opportunity to learn chemistry using multiple levels of representation, including symbolic representation, students display a much better understanding of chemistry at the macroscopic and microscopic levels (Mocerino, Chanrasegaran & Treagust, 2009). One form of symbolic representation used in this study is modeling. The aim of modeling is to make a particular part of the science content easier to understand, define, quantify, visualize, or simulate by referencing it to existing and accepted knowledge (Grosslight, Unger, Jay & Smith, 1991). When students have limited exposure to models, they think of them as physical copies of reality that embody different spatial perspectives rather than potential different theoretical perspectives, but with more exposure, they
increasingly understand how models are designed for a particular purpose, and that is to help communicate an idea (Grosslight, Unger, Jay & Smith, 1991). When using models it was determined, by Harrison and Treagust (1996) that students’ conception of models that were used to describe chemistry concepts revealed much about the difficulties that students struggled with as they tried to incorporate new scientific principles and terminology into their own way of thinking. Models, whether developed for POGIL, InDGIM, or another form of guided inquiry lesson, can only be truly understood when the instructor has helped students develop modeling skills and students are able to understand the significance of this process (Harrison & Treagust, 1996). Using models is essential in science education, however, understanding how students interpret those models may be even more critical. The use of modeling can be a useful tool for identifying ideological changes that might happen during the learning cycle.

Modeling is more than just a visual representation of an idea. It is often classified as an approach to teaching as well. The InDGIM approach is organized around scientific models as coherent units of structured knowledge, which in turn engages students collaboratively in making and using models to describe, explain, predict, design, and control physical phenomena in the environment (Jackson, Dukerich, & Hestenes, 2008). A unit of study typically begins with a demonstration and class discussion. This establishes a common understanding of a question to be asked of nature. Then, in small groups, students collaborate in planning and conducting experiments to answer or clarify the question. Students present and justify their conclusions in oral and written form, including the formulation of a model for the phenomena in question and an evaluation of the model by comparison with data. The POGIL approach is similar with the approach of group work, but typically the students are collaborating on learning concepts
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

through use of pencil and paper activities as opposed to conducting experiments to collect their own data. The role of the teacher, in both POGIL and the InDGIM, is to introduce technical terms and representation tools as need to sharpen the model, to facilitate the activity and to improve the quality of discourse. The teacher is prepared with an agenda for student progress and guides student inquiry and discussion in a particular direction.

During the second phase of modeling students employ their newly discovered model to new situations to refine and deepen their understanding. In both POGIL and InDGIM students work on challenging worksheet problems in small groups, and then present their findings to the whole class. Students are assessed on their knowledge through a quiz or a test. This modeling approach, incorporated by both POGIL and InDGIM, stresses developing a sound conceptual understanding through graphical and diagrammatic representations before moving on to an algebraic treatment of problem solving (Jackson, Dukerich, & Hestenes, 2008).

A key component of this approach, found in both POGIL activities and InDGIM, is that it moves the teacher from the role of authority figure who provides the knowledge to that of a coach/facilitator who helps the students construct their own understanding. Since students systematically misunderstand most of what we tell them the emphasis is placed on student articulation of the concepts. According to Jackson, Dukerich, and Hestenes (2008), data from some 20,000 students who have been through science programs that follow this extensive modeling regime achieve twice the gains on standard tests of conceptual understanding as students who are taught conventionally.

Each approach attempts to develop students’ reasoning skills. According to Osborne
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

(2010), inquiry-based programs help students to

“identify patterns in data, such as co-variation, and to make inferences; to coordinate theory with evidence and to discriminate between evidence that supports (inclusive) or does not support (exclusive) or that is simply indeterminate; to construct evidence-based, explanatory hypotheses or models of scientific phenomena and persuasive arguments that justify their validity; and to resolve uncertainty, which requires a body of knowledge about concepts of evidence such as the role of statistical techniques, the measurement of error, and the appropriate use of experimental designs, such as randomized double-blind trials (p. 465).

Osborne (2010) acknowledges the importance of inquiry, hand-on programs, which allow student collaboration and discourse, to the development of these reasoning skills.

The strength of both of these approaches is that they incorporate all of these essential components: cooperative learning in small groups, guided inquiry with the teacher being a facilitator, the use of models, the establishment of sociochemical norms, and the development of reasoning skills through collaboration, to support student learning.

Implications

One of the most important findings of the work described herein is the explication and dissemination of tools chemistry education researchers can use to investigate classroom level dynamics. Social factors and attitudes play an important role in framing the classroom learning environment and their impact on student learning could certainly be more fully explored. The POGIL and InDGIM approaches rely heavily on socialization between students, so more work is
needed in order to coordinate individual and social views of learning.

From this perspective scientific knowledge and understanding are constructed when individuals engage socially while working together on shared problems or tasks and making meaning is thus a “dialogic process involving persons-in-conversation” (Driver, Askoko, Leach, Mortimer, & Scott, 1994, p. 7). Learning science involves both personal and social processes. Learners must be given access not only to physical experiences, through the guided inquiry method, and models of conventional science, but more importantly, the time to cement these ideas through socialization. Instructors must be able to provide the support and guidance for students to make sense of the ideas presented in different science courses, and they must listen and diagnose the ways in which these instructional activities are being interpreted (Driver, et al., 1994), but they must be willing to step back and let the students carry on dialogues with each other to allow the development of a stronger conceptual framework.

Another thing that the instructors must do is to be more sensitive to whatever gender bias they bring to the teaching of courses in chemistry and to be aware of how it is manifested in their own teaching. The results support that females will perform better in the class, especially those students that performed well on the Keystone Biology test. Instructors should look at variables that might help them predict which students might be successful and which ones might need more assistance. Using data to help establish a particular approach would be beneficial to all students.

This study would indicate that both the pedagogical approaches (POGIL and InDGIM) worked equally in the four academic chemistry classes at Falls Fallow High School since the data illustrated that there were no significant differences in either approach. It is recommended
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

that teachers continue to explore both approaches in future classes.

As the instructors at Falls Fallow High School evaluate their courses of study in the fields of Environmental Science, Biology, Chemistry and Physics, they need to establish dialogic learning environments that incorporate different levels of inquiry and models and include opportunities for socialization that their courses provide students. The current study would indicate that guided inquiry along with a high degree of socialization would benefit all students. Students participating in sequenced inquiry lessons with explicit goals showed improved learning compared to students who simply performed an investigation (Dushl, 2008). Understanding the purposes of science inquiry makes a difference.
References


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


physical chemistry class. *Chemistry Education Research and Practice, 16*, 769-785.

Retrieved from http://pubs.rsc.org/en/content/getauthorversionpdf/c5rp00064e

doi: 10.1002/tea.3660300607


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher* 45(1), 18-35. doi: 10.3102/0013189X16633182


National Center for Educational Statistics. Retrieved from http://nces.ed.gov/ccd/districtsearch/district_detail.asp?Search=1&InstName=Springfield+Township&City=Oreland&State=42&Zip=19075&DistrictType=1&DistrictType=2&DistrictType=3&DistrictType=4&DistrictType=5&DistrictType=6&DistrictType=7&NumOfStudentsRange=more&NumOfSchoolsRange=more&ID2=4222620&details=1

THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


POGIL Project. Retrieved from https://pogil.org


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


Tang, H. & Abraham, M. R. (2016). Effect of computer simulations at the particulate and macroscopic levels on students’ understanding of the particulate nature of matter. *Journal of Chemical Education, 93*(1), 31-38. doi 10.1021/acs.jchemed.5b00599

THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES


THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Appendix A: Copy of Consent Form

Informed Consent Form
(Electronic GALT and ParNoMA2 Consent)

Principal Investigator:
Scott M. Zgraggen, MA Ed.

Purpose:
The goal of this proposed research is to compare how students perform using one of two different pedagogical approaches in Academic Chemistry: Process Oriented Guided Inquiry Learning (POGIL) developed by the POGIL project, or the Independently Developed Guided Inquiry Method (InDGIM) developed by the instructors at your child’s high school. This study will examine the achievement of chemistry standards using data from standardized online tests and independently developed chemistry tests and quizzes.

Results:
The results of this study may be presented at professional meetings. The results may also be published in a professional journal. All data will remain anonymous. If you wish to see a copy of the final work, send a request to zgrags86@arcadia.edu.

Duration/Location:
Your child’s class will be selected to be part of the treatment (POGIL) group or the comparison (InDGIM) group for the 2016-17 school year. The same chemistry topics, unit tests, and final exam will be used in each course. In order to understand the significance of using either method, I am requesting the following data from your child:

1) Responses to a standardized test administered on the computer. Your child will be answering two self-administered, electronic tests within the next five school days. The first test is labeled the Group Assessment of Logical Thinking (GALT) and the second test is labeled the Particulate Nature of Matter Assessment (ParNoMA2). Each test, developed independently by researchers, should take no longer than 20 minutes to complete. Your child will be asked to take the same two tests over again at the end of the semester. If allowed, the results of these tests will be statistically analyzed to determine if there was a greater gain with one pedagogical approach or the other.

2) Demographic data including gender to see if this variable has an impact on learning outcomes.

3) Performance as measured by grades on all unit tests and quizzes.

4) Proficiency scores on the Keystone Algebra I and Biology tests.

Inclusion/Exclusion Criteria:
This document shall serve as informed consent. If you decide to not have your child participate, you must contact the primary researcher to have your child excluded from this study. The primary researcher is available by email and phone: Zgrags86@arcadia.edu or 215-233-6030.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

Protection of Subjects:
Names will not be collected in this study. All data collected will be by student number so the principal investigator will not know the names of the subjects involved. The final published work will include only grouped results, and therefore it will not be possible to identify individual participants. The results of all data will be collected and kept on personal, password protected computer to which only the principal investigator has access. This information will not be used in any way to evaluate your child with regards to his/her grade in the course. All data will be destroyed upon completion of the study.

Potential Risks or Discomforts:
The questions on the two electronic tests are of a non-sensitive nature. However, your child can skip any questions he/she does not want to answer. Your child will not be penalized in any way for leaving questions blank.

Compensation:
There is no financial compensation for participating in this study.

Voluntary Participation and Withdrawal:
Your child’s participation is completely voluntary. Your child may withdraw at any time which means his/her data will no longer be used for the study.

Consent:
The Arcadia University Institutional Review Board (IRB) approved this study’s protocol. To ensure that this research continues to protect your child’s rights and minimize your child’s risk, the IRB reserves the right to examine and evaluate the data and research protocols involved in this project. If you wish additional information regarding your child’s rights in this study, you may contact the Office of Research Subject Protection at 267-620-4111. If you have additional questions or comments, please contact the principal investigator or faculty advisor:

Mr. Scott M. Zgraggen, MA Ed. at Zgrags86@arcadia.edu
(215) 233-6030
or
Dr. Steve Gulkus at gulkuss@arcadia.edu
(215) 572-2120

“I have read the consent form. I agree that my child meets all the inclusion criteria and I agree to have the information collected from the electronic tests, unit tests, and proficiency scores on the Keystone Algebra and Biology tests to be used anonymously in this study. I understand that my child can choose to leave a question blank on the electronic tests, if he/she would rather not answer it and that my child can exit the electronic tests AND/OR the research at any time. I have been given a copy of this consent form.”
Appendix B: Copy of Summary Letter Sent Home to Parents

Parent/Guardian Information Letter

September, 2016

Dear Parent/Guardian,

My name is Scott M. Zgraggen and I am a doctoral student at Arcadia University, in the Department of Education. Your child is invited to participate in a study I am conducting for my doctoral dissertation on comparing two approaches to teaching chemistry: Process Oriented Guided Inquiry Learning (POGIL) and the Independently Developed Guided Inquiry Method (InDGIM) produced by the instructors at Springfield Township High School.

POGIL is a research-based, student centered philosophy and science pedagogy in which students continuously work in small groups to engage in guided inquiry using carefully designed materials that direct and guide students to build and rebuild their content knowledge, similar, but not identical to the process of InDGIM. Both approaches teach content and key process skills of science, and they focus on core concepts and processes of science as they encourage and foster a deeper understanding of the course material while developing higher-order thinking skills. Although similar there are enough subtle differences that make comparing these two methods important.

I am interested in learning how students respond to either form of science inquiry learning. The title of my study is “Comparing the Process Oriented Guided Inquiry Learning (POGIL) Method to an Independently Developed Guided Inquiry Method (InDGiM) in a High School Academic Chemistry Course.” This study is important because it will make a contribution to the cumulative data on the POGIL approach and it will be one of the few studies that compares two inquiry approaches at the high school level.

The attached form is asking for your permission to include your child in this study because s/he can provide valuable insight into which approach should be used for future chemistry courses at Springfield Township High School. It is my goal to have 100% student participation in this project.

Thank you for your consideration,

Sincerely,
Scott M. Zgraggen, MA Ed.
Appendix C: Copy of Student Assent Letter and Classroom Assent Script.

Student Assent Form
Dear Student,
My name is Scott M. Zgraggen and I am a graduate student at Arcadia University. I am completing a study at Springfield Township High School and I am trying to understand which teaching approach is better for students studying college preparatory chemistry: Process Oriented Guided Inquiry Learning (POGIL) developed by the POGIL project, or the Independently Developed Guided Inquiry Method (InDGIM) developed by the instructors at your high school.

I am asking for your permission to use the results of two tests Mr. Britton and I will be asking you to take on the computer. Everyone will be taking the tests; I am asking that you allow me to keep the results of your test to use in my research. The first test everyone will take is labeled the Group Assessment of Logical Thinking (GALT) and the second is labeled the Particulate Nature of Matter Assessment (ParNoMA2). I will also be asking for your permission to be given access and to use your proficiency rating on the Keystone Algebra I and Biology tests as well as final performance grades earned on unit tests and quizzes in this Chemistry class for the purposes of my research. This information, along with your gender, will be analyzed to determine if there is a greater gain with one instructional approach or the other, and to see what variables, if any, have an effect on the results.

Participation is strictly voluntary, which means you do not have to take part if you don’t want to. Please remember that when I say “participating,” it means that you allow me to use the results of the assessments I discussed, for my research. Everyone will be participating in the assessments. Again, participating in this study is strictly voluntary, which means you do not have to take part if you don't want to. Nothing will happen to you if you decide not to participate. Participating in this study will not affect your grade or how your teacher, or your school treats you. You can ask questions about this study at any time.

If you agree to participate you will first need to obtain your parents’ or guardians’ consent. Everyone is asked to take a letter explaining the research home to your parents or guardians. They should have already received one in the mail. You are not allowed to participate unless your parents or guardians grant their consent. This letter should be signed, brought back, and given to Mr. Britton before the start of class within the next two days.

Once your parents or guardians grant their consent, your scores from the two computerized tests, Keystone Algebra I and Biology scores, and grades on in-class tests and quizzes will be used in my research. All scores will be reported to me by Mr. Britton using your school ID number so that confidentiality may be maintained.
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

PARTICIPATING:

Please read the following and sign below if you agree to participate. I understand that:

- if I don’t want to take the computerized tests that’s ok and I won’t get into trouble
- anytime that I want to stop participating in this study that’s ok
- my name will not be known and my answers will be completely private
- by signing this I WILL participate in the study

Signature: _______________________________________________________________________
Name: (please print) _______________________________________________________________________
Date: ___________________________________________________________________________

There are two copies of this letter. After signing them, keep one copy for your records and return the other one. Thank you in advance for your cooperation and support.

For further information regarding this research please contact:

Dr. Steve Gulkus at gulkuss@arcadia.edu or at (215) 572-2120

Mr. Scott M. Zgraggen at zgrags@arcadia.edu or at (215) 233-6030

If you have any questions about your rights as a participant you may contact the Arcadia University Institutional Review Board at (215) 572-2900.

NOT PARTICIPATING:

Please sign below if you WILL NOT participate. I understand that:

- by signing this I WILL NOT be participating in the study

Signature: _______________________________________________________________________
Name: (please print) _______________________________________________________________________
Date: ___________________________________________________________________________
Hello, my name is Mr. Scott M. Zgraggen and I am a graduate student at Arcadia University. As part of my studies in Educational Leadership, I am completing a study at Springfield Township High School and I am trying to understand which teaching approach is better for students studying college preparatory chemistry: Process Oriented Guided Inquiry Learning (POGIL) developed by the POGIL project, or the Independently Developed Guided Inquiry Method (IndGIM) developed by the instructors here at your high school.

Please take this Informed Consent form home and review it with your parents or guardians. If your parents or guardians do not want you to participate in this study, then do not return the consent form. If they give permission for you to participate, they must sign the form and you can return it to Mr. Britton at the beginning of class within the next two days. If your parent/guardian does not give permission for you to participate, then you cannot participate.

This form is the Student Assent Form. This form asks for your own permission to participate in the study. If your parent or guardian has not given permission for you to participate, then you will not be able to participate. Please remember that when I say “participating,” it means that you allow me to use the results of the assessments I discussed, for my research. Everyone will be participating in the assessments. Again, participating in this study is strictly voluntary, which means you do not have to take part if you don't want to. Nothing will happen to you if you decide not to participate. Participating in this study will not affect your grade or how your teacher, or your school treats you. You can ask questions about this study at any time.

If you agree to participate you will first need to obtain your parents’ or guardians’ consent. Everyone is asked to take a letter explaining the research home to your parents or guardians. They should have already received one in the mail. You are not allowed to participate unless your parents or guardians grant their consent. This letter should be signed, brought back, and given to Mr. Britton before the start of class within the next two days.

Once your parents or guardians grant their consent, your scores from the two computerized tests, Keystone Algebra I and Biology scores, and grades on in-class tests and quizzes will be used in my research. All scores will be reported to me by Mr. Britton using your school ID number so that confidentiality may be maintained.

Please fill out the Student Assent Form now.

Thank you for your time and willingness to participate. Please review the Informed Consent Form with your parents tonight. If you have any questions later on you may reach me by email at zgrags86@arcadia.edu or by phone at 215-233-6030. I will see you again in a few days.
Appendix D: The Group Assessment of Logical Thinking (GALT)

The Group Assessment of Logical Thinking (GALT) Test

Instructions:
This is an assessment that will be used as part of the research being conducted by Mr. Scott M. Zgraggen. This research is being conducted to determine if there are any differences that exist in the outcomes of two approaches to presenting chemistry in high school.

Please remember that if you have agreed to participate, I will use the results from your assessment in my research. If you have not agreed to participate, I will not use the results of this assessment in my research.

Carefully read each question. Choose the best answer for each one and bubble in your response.

Situation One:
Tom has two balls of clay. They are the same size and shape, as shown in the picture below. When Tom places the balls on the balance, they weigh the same.

The balls of clay are removed from the balance pans. The orange ball is flattened like a pancake, as shown below:
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

1) Which of these statements is true?
   A. The pancake-shaped clay weighs more.
   B. The two pieces weigh the same.
   C. The ball-shaped clay weighs more.

2) What is the reason for your response above?
   A. You did not add or take away clay.
   B. When the orange ball was flattened like a pancake, it had a greater area.
   C. When something is flattened, it loses weight.
   D. Because of its density, the round ball had more clay in it.

Situation Two:
   Linn has two jars. They are the same size and shape. Each is filled with the same amount of water, as shown below.

   ![Jar 1 and Jar 2](image)

   She also has two metal weights of the same volume. One weight is light. The other is heavy.

   ![Light Metal Weight and Heavy Metal Weight](image)

   She lowers the light weight into jar 1. The water level in the jar rises, as shown below.
3) 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.

4) What is the reason for your answer to the last question?
   A. The weights are the same size so they will take up equal amounts of space.
   B. The heavier the metal weight, the higher the water will rise.
   C. The heavy metal weight has more pressure, therefore the water will not rise as much.
   D. The heavier the metal weight, the lower the water will rise.

Situation Three:
The drawing shows two glasses, a small one and a large one. It takes 15 small glasses of water of 9 large glasses of water to fill the large jar. It takes 10 small glasses of water to fill the small jar.
5) How many large glasses of water does it take to fill the small jar?
   A. 4  
   B. 5  
   C. 6  
   D. Some other amount

6) What is your reason for your answer above?
   A. It takes five less small glasses of water to fill the small jar, so it will take five less large glasses of water to fill the same jar.
   B. The ratio of small to large glasses will always be 5 to 3.
   C. The small glass is half the size of the large glass. So it will take about half the number of small glasses of water to fill up the same small jar.
   D. There is no way of predicting.
Situation Four:
Joe has a scale like the one below.

When he hangs a 10-unit weight at point D, the scale looks like this:

7) Where would he hang a 5-unit weight to make the scale balance again?
   A. At point J.
   B. Between K and L.
   C. At point L.
   D. Between L and M.
   E. At point M.

8) What is your reason for your answer above?
   A. It is half the weight so it should be put at twice the distance.
   B. The same distance as 10-unit weight, but in the opposite direction.
   C. Hang the 5-unit weight further out, to make up for it being smaller.
   D. All the way at the end gives more power to make the scale balance.
   E. The lighter the weight, the further out it should be hung.
Situation Five:

Three strings are hung from a bar. String #1 and #3 are of equal length. String #2 is longer. Charlene attached a 5-unit weight at the end of string #3 and #2. A 10-unit weight is attached at the end of string #1. Each string with a weight can be swung. Charlene wants to find out if the length of the string has an effect on the amount of time it takes the string to swing back and forth.

9) Which string and weight would she use for her experiment?
   A. String #1 and #2.
   B. String #1 and #3.
   C. String #2 and #3.
   D. String #1, #2 and #3.
   E. String #2 only.

10) What is your reason for the answer you gave in the last question?
   A. The length of the strings should be the same. The weights should be different.
   B. Different lengths with different weights should be tested.
   C. All strings and their weights should be tested against all others.
   D. Only the longest string should be tested. The experiment is concerned about length, not weight.
   E. Everything needs to be the same except the length, so you can tell if the length makes a difference.
Situation Six:

Eddie has a curved ramp. At the bottom of the ramp, there is one ball called the target ball. There are two other balls, a heavy and a light one. He can roll one ball down and hit the target ball. This causes the target ball to move up the other side of the ramp. He can roll the balls from two different points, a low point and a high point.

Eddie released the light ball from the low point. It rolled down the ramp. It hit and pushed the target ball up the other side of the ramp.

He wants to find out if the point a ball is released from makes a difference in how far the target ball goes.
11) To test this, which ball would he now release from the high point?
   A. The heavy ball.
   B. The light ball.

12) What is your reason for the answer you gave above?
   A. He started with the light ball so he should finish with it.
   B. He used the light ball the first time. The next time he should use the heavy ball.
   C. The heavy ball would have more force to hit the target ball farther.
   D. The light ball would have to be released from the high point in order to make a fair comparison.
   E. The same ball must be used as the weight of the ball does not count.

Situation Seven:

In a cloth sack, there are 3 brown wooden squares, 4 black wooden squares, 5 white wooden squares, 4 brown wooden diamonds, 2 black wooden diamonds, and 3 white wooden diamonds. All of the square pieces are the same size and shape. The diamond pieces are also the same size and shape. One piece is pulled out of the sack.
13) What are the chances that it is a brown piece?
   A. 1 out of 3
   B. 1 out of 4
   C. 1 out of 7
   D. 1 out of 21
   E. Other, not listed

14) What is your reason for the answer you gave?
   A. There are twenty one pieces in the cloth sack. One brown piece must be chosen from these.
   B. One brown piece needs to be selected from a total of seven brown pieces.
   C. Seven of the twenty one pieces are brown pieces.
   D. There are three sets in the cloth sack. One of them is brown.
   E. 1/4 of the square pieces and 4/9 of the diamond pieces are brown.
Situation Eight:

In a cloth sack, there are 3 brown wooden squares, 4 black wooden squares, 5 white wooden squares, 4 brown wooden diamonds, 2 black wooden diamonds, and 3 white wooden diamonds. All of the square pieces are the same size and shape. All of the diamond pieces are the same size and shape. Reach in and take the first piece you touch.

15) What are the chances of pulling out a white or brown diamond?
   A. 1 out of 3
   B. 1 out of 9
   C. 1 out of 21
   D. 9 out of 21
   E. Other, not listed.
16) What is your reason for the answer you gave?
   A. Seven of the twenty one pieces are brown or white diamonds.
   B. 4/7 of the brown and 3/8 of the white are diamonds.
   C. Nine of the twenty one pieces are diamonds.
   D. One diamond piece needs to be selected from a total of twenty one pieces in the cloth sack.
   E. There are 9 diamond pieces in the cloth sack. One piece must be chosen from these.

Situation Nine:
A farmer observed the mice that live in his field. He found that the mice were either fat or thin. Also, the mice had either brown tails or pink tails. This made him wonder if there might be a relation between the size of the mouse and the color of its tail. So he decided to capture all the mice in one part of his field and observe them. The mice that he captured are shown below.

17) Do you think that there is a relation between the size of the mice and the color of their tails (that is, is one size of a mouse more likely to have a certain color tail and vice versa?)
   A. Yes
   B. No
18) What is your reason for the answer you gave?
   A. 8/11 of the fat mice have pink tails and 3/4 of the thin mice have brown tails.
   B. Fat and thin mice can have either a brown tail or a pink tail.
   C. Not all fat mice have pink tails. Not all thin mice have brown tails.
   D. 18 mice have pink tails and 12 have brown tails.
   E. 22 mice are fat and 8 mice are thin.

Situation Ten:
Some of the fish blow are big and some are small. Also, some of the fish have wide stripes on their sides. Others have narrow stripes.

19) Is there a relationship between the size of the fish and the kind of stripes it has (that is, is one size of fish more likely to have a certain type of stripes and vice versa)?
   A. Yes
   B. No

20) What is the reason for your answer above?
   A. Big and small fish can have either wide or narrow stripes.
   B. 3/7 of big fish and 9/21 of small fish have wide stripes.
   C. 7 fish are big and 21 are small.
   D. Not all big fish have wide stripes and not all small fish have narrow stripes.
   E. 12/28 of fish have wide stripes and 16/28 of fish have narrow stripes.
Situation Eleven:

After supper, some students decide to go dancing. There are three boys: Albert (A), Bob (B), and Charles (C), and three girls: Louise (L), Mary (M), and Nancy (N). One possible pair of dance partners is A-L, which means Albert and Louise.

21) List all other possible couples of dancers. For sake of this argument, boys do not dance with boys and girls do not dance with girls.
Situation Twelve: In a new Shopping center, 4 stores are going to be placed on the ground floor. A Barber Shop (B), a Discount Store (D), a Grocery Store (G), and a Coffee Shop (C) want to locate there. One possible way the stores could be arranged is BDGC, Which means the Barber Shop first, the Discount Store next, then the Grocery Store and the Coffee Shop last.

22) List all other possible ways the stores can be lined up in the four locations.
Appendix E: The Particular Nature of Matter Assessment (ParNoMA2)

Particulate Nature of Matter Assessment (ParNoMA2)

Instructions:
This is an assessment that will be used as part of the research being conducted by Mr. Scott M. Zgraggen. This research is being conducted to determine if there are any differences that exist in the outcomes of two approaches to presenting chemistry in high school.

Please remember that if you have agreed to participate, I will use the results from your assessment in my research. If you have not agreed to participate, I will not use the results of this assessment in my research.

Carefully read each question. Choose the best answer for each one and bubble in your response.

1. A diagram representing water molecules in the solid phase (ice) is shown below.

Which of these diagrams best shows what water would look like after it melts (changes to a liquid)? Which of these diagrams best shows what water would look like after it melts (changes to a liquid)? Which of these diagrams best shows what water would look like after it melts (changes to a liquid)?

A. B. C. D. E.
2. Consider three samples of water in three phases. The first is solid water (ice) at 0°C, the second is liquid water at 24°C, and the third is gaseous water at 100°C. The water molecules in the liquid phase __________ the water molecules in the gaseous phase.
   A. move faster than
   B. move slower than
   C. move at the same speed as
   D. move more randomly than
   E. travel in the same direction as

3. Which of the following processes will make water molecules larger?
   A. freezing
   B. melting
   C. evaporation
   D. condensation
   E. none of the above

4. A sample of liquid ammonia (NH₃) is completely evaporated (changed to a gas) in a closed container as shown:

Which of the following diagrams best represents what you would “see” in the same area of the magnified view of the vapor?
5. When water changes from a liquid to a gas through evaporation or vaporization, energy is required to
   A. break the bonds between the hydrogen atoms.
   B. form new bonds between the atoms.
   C. break the bonds between the oxygen and hydrogen atoms in the molecules.
   D. break the water molecules away from other water molecules.
   E. form new bonds between the molecules.

6. A water molecule in the gas phase is ______ a water molecule in the solid phase.
   A. smaller than
   B. lighter than
   C. heavier than
   D. larger than
   E. the same weight as

7. When water is vaporized, it is changed to
   A. hydrogen and oxygen
   B. hydrogen only
   C. gaseous water
   D. air, hydrogen, and oxygen
   E. oxygen only

8. A pot of water is placed on a hot stove. Small bubbles begin to appear at the bottom of the pot. The bubbles rise to the surface of the water and seem to pop or disappear. What are the bubbles made of?
   A. heat
   B. air
   C. gaseous oxygen and hydrogen
   D. gaseous water
   E. none of the above

9. A pot of water on a hot stove begins to boil rapidly. A glass lid is placed on the pot and water droplets begin forming on the inside of the lid. What happened?
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

A. The lid became sweaty.
B. Steam cools and water molecules moved closer together.
C. Water from outside leaked into the pot.
D. Hydrogen and oxygen combined to form water.
E. Steam combined with the air to wet the inside of the lid.

10. Consider three samples of water in three phases. The first is solid water (ice) at 0°C, the second is liquid water at 24°C, and the third is gaseous water at 100°C. The water molecules in the liquid phase __________ the water molecules in the solid phase.
   A. move faster than
   B. move slower than
   C. move at the same speed as
   D. move less randomly than
   E. travel in the same direction as

11. A wet dinner plate is left on the counter after it has been washed. After awhile it is dry. What happened to the water that didn’t drip onto the counter?
   A. It changes to carbon dioxide.
   B. It just dries up and no longer exists as anything.
   C. It goes into the air as molecules of water.
   D. It goes into the plate.
   E. It changes to oxygen and hydrogen in the air.

12. Which of the following processes does NOT require heat energy?
   A. evaporating water
   B. melting ice
   C. boiling water
   D. vaporizing water
   E. condensing water

13. When water molecules in the gas phase are heated, the molecules themselves
   A. expand.
   B. move faster.
   C. become less massive.
   D. change to a liquid.
   E. release air.

14. Which of the following processes will make molecules smaller?
   A. freezing
   B. melting
   C. evaporation
   D. condensation
   E. none of the above
15. Oxygen and hydrogen gases may be formed from liquid water through the process of 
   A. vaporization. 
   B. evaporation. 
   C. decomposition. 
   D. freezing. 
   E. boiling.

16. A diagram representing carbon dioxide molecules in the solid phase, also known as dry ice, is shown below. 

   ![Diagram of carbon dioxide molecules in the solid phase]

   Which of these molecular diagrams best shows what dry ice would look like after it melts (changes to a liquid)?

   A. ![Diagram A]  
   B. ![Diagram B]  
   C. ![Diagram C]  
   D. ![Diagram D]  
   E. ![Diagram E]

17. When water at 25°C is heated and changes to a gas at 110°C, the water molecules 
   A. become more organized. 
   B. move farther apart. 
   C. stop moving. 
   D. move closer together. 
   E. move more slowly.

18. Which of the following processes requires heat energy? 
   A. condensation 
   B. freezing 
   C. evaporation 
   D. cooling 
   E. none of the above
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

19. A water molecule in the liquid phase is ______ a water molecule in the solid phase.
   A. smaller than
   B. lighter than
   C. heavier than
   D. larger than
   E. the same weight as

20. When water at 24°C is cooled to 0°C and freezes, the water molecules
   A. become less organized.
   B. move much faster.
   C. stop moving.
   D. break apart.
   E. move much more slowly.
Appendix F: Operational Definitions

**Group Assessment of Logical Thinking (GALT).** The GALT is a 15-20 minute Piagetian test of logical thinking, focusing on six modes of reasoning, including concrete and formal operational. This 21 question test, developed by Roadrangka, Yeany and Padila (1983), had an indicated coefficient Cronbach alpha reliability of 0.85 will provide a means to assess cognitive development of a large number of students within a single class period.

**Guided Inquiry.** An inquiry approach to teaching and learning in which teachers provide scaffolding for students as they explore natural phenomena. Teachers serve as facilitators of learning in this pedagogy. Often models and written documents are used to guide students to discover scientific phenomena.

**Inquiry/Inquiry-Based Instruction.** Inquiry-based instruction is a learning process in which students are engaged, or are active in the learning process. It is “something that students do, not something that is done to them” (Anderson, 2002, p. 3) and it is created through “a classroom where students are engaged in essentially open-ended, student-centered, hands-on activities” (Colburn, 2000, p. 42).

**Independently Developed Guided Inquiry Method (InDGIM).** An instructional approach that uses a wide variety of methods including proper questioning strategies (designed to stimulate thought and action), science process skills, discrepant events, inductive activities, information gathering, and problem solving.

**Learning cycle.** The theory that states that learning occurs in three stages: exploration, concept invention, and application.
Kinetic Theory of Matter. Also known as the particle theory of matter, states that all matter is composed of particles (i.e. atoms, ions, molecules, subatomic particles) that are in constant motion. The amount of motion of the particles is determined by the energy they possess. The state of matter (solid, liquid, gas, plasma) is determined by the energy of the particles.

Particle Theory of Matter. Also known as the kinetic theory of matter, or particulate nature of matter, states that all matter is composed of very tiny particles that are in constant motion (see kinetic theory of matter).

Particulate Nature of Matter (ParNoMA2) Test. The 20-item multiple-choice Particulate Nature of Matter Assessment, with a Cronbach reliability of 0.83, targets misconceptions surrounding phases of matter (Yezierski & Birk, 2006).

Process Oriented Guided Inquiry Learning (POGIL). A student-centered instructional approach that simultaneously develops discipline content mastery and key process skills such as critical thinking, effective communication, and teamwork (Moog & Spencer, 2008). POGIL activities guide students through an exploration to construct, deepen, refine, and/or integrate understanding of relevant disciplinary content.

Directions: Reflect on the science lesson that you taught today. In your reflection, consider each of the following categories, and the six statements on the left, written in bold. After looking at each bold statement, assess today’s science instruction based on the categories delineated for the statement. Place an “X” in the corresponding cell for each bold-faced statement. If there is no evidence on one of the statements in today’s lesson, place a slash through the bold-faced statement. When you are finished, you should have six total responses.

<table>
<thead>
<tr>
<th>Learners are engaged by scientifically oriented questions.</th>
<th>Teacher Centered</th>
<th>Learner Centered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher provides an opportunity for learners to engage with a scientifically oriented question.</td>
<td>Learner is prompted to formulate own questions or hypothesis to be tested.</td>
<td>Teacher suggests topic areas or provides samples to help learners formulate own questions or hypothesis.</td>
</tr>
<tr>
<td>Learners give priority to evidence, which allows them to develop and evaluate explanations that address scientifically oriented questions.</td>
<td>Teacher engages learners in planning investigations to gather evidence in response to questions.</td>
<td>Learners develop procedures and protocols to conduct a full investigation, providing support and scaffolding with making decisions.</td>
</tr>
<tr>
<td>Teacher helps learners give priority to evidence which allows them to draw conclusions and/or develop and evaluate explanations that address scientifically oriented questions.</td>
<td>Learners determine what constitutes evidence and develop procedures and protocols for gathering and analyzing relevant data (as appropriate).</td>
<td>Learners direct learners to collect certain data or only provides portion of needed data. Often provides protocols for data collection.</td>
</tr>
</tbody>
</table>

| Learners formulate explanations and conclusions from evidence to address scientifically oriented questions. | Learners formulate conclusions and/or explanations from evidence to address scientifically oriented questions. Learners are prompted to analyze evidence (often in the form of data) and formulate their own conclusions/explanations. | Teacher prompts learners to think about how analyzed evidence leads to conclusions/explanations, but does not cite specific evidence. | Teacher directs learners’ attention (often through questions) to specific pieces of analyzed evidence (often in the form of data) to draw conclusions and/or formulate explanations. | Teacher directs learners’ attention (often through questions) to specific pieces of analyzed evidence (often in the form of data) to lead learners to predetermine correct conclusions/explanations (verification). | No evidence observed. |

| Learners evaluate the explanations in light of alternative explanations, particularly those reflecting scientific understanding. | Learners evaluate their conclusions and/or explanations in light of alternative conclusions/explanations, particularly those reflecting scientific understanding. Learners are prompted to examine other resources and make connections and/or explanations independently. | Teacher provides resources to relevant scientific knowledge that may help identify alternative conclusions and/or explanations. Teacher may or may not direct learners to examine these resources, however. | Teacher does not provide resources to relevant scientific knowledge to help learners formulate alternative conclusions and/or explanations. Instead, the teacher identifies related scientific knowledge that could lead to such alternatives, or suggests possible connections to such alternatives. | Teacher explicitly states specific connections to alternative conclusions and/or explanations, but does not provide resources. | No evidence observed. |

| Learners communicate and justify their proposed conclusions and/or explanations. | Learners communicate and justify their proposed conclusions and/or explanations. Learners specify content and layout to be used to communicate and justify their conclusions and explanations. | Teacher talks about how to improve communication, but does not suggest content or layout. | Teacher provides possible content to include and/or layout that might be used. | Teacher specifies content and/or layout to be used. | No evidence observed. |
Appendix H: POGIL Guideline Checklist

Directions: Reflect on the science lesson that you taught today. In your reflection, consider each of the following categories, and the ten statements on the left, written in bold. After looking at each bold statement, assess today’s science instruction based on the categories delineated for the statement. Place an “X” in the corresponding cell for each bold-faced statement. If there is no evidence on one of the statements in today’s lesson, place a slash through the bold-faced statement. When you are finished you should have ten total responses.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Yes</th>
<th>No</th>
<th>No Evidence Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Students are working collaboratively.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Students are working in groups of 3 or 4.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Students have assigned roles within their groups.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Students are working on a POGIL specific activity.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. The activity is the first introduction to the topic or specific content.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Students working on the activity during class time while teacher is present.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Teacher serves predominately as a facilitator of student learning (lesson is not lecture-based or instructor-centered).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Students have not worked on any part of the activity prior to class meeting time.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Groups complete all of the Critical Thinking Questions during about 40 minutes of class time.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Groups are expected to work on the Exercises or Problems at home.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix I: Sample of Unit Test in Chemistry

Dealing with the Kinetic Theory: Quiz A

The following is a quiz that has been designed to assess student understanding of the material that has been covered over the last three weeks in Chemistry. Notes are allowed, however, only 30.6 minutes has been allocated for this quiz, so please work quickly. Calculators are allowed, however they may not be shared between students. Please make sure that all answers are written in **SCIENTIFIC NOTATION** as well as with the appropriate **SIGNIFICANT FIGURES**. If there are any questions, then please ask the instructor. You may use the space provided or a separate sheet of paper. Remember to **SHOW ALL WORK** and **NO TALKING UNTIL EVERYONE IS FINISHED!**

1) Atmospheric pressure on the peak of Mt. Everest can be as low as 0.197368 atm, which is why climbers need to bring oxygen tanks for the last part of the climb. If the climbers carry 10.0 liter tanks with an internal pressure of 40.00 atm, what will be the volume of the gas when it is released from the tanks at 0.197368 atm? (10.00 points)

2) Some students believe that teachers are full of hot air. If a well-known chemistry teacher inhales 2.200 liters of gas at a temperature of 18.0 °C and then blows out a volume of 2.351 liters, what is the final temperature (in Kelvin) of this air? Extra credit for determining the final temperature in Celsius. (10.00 points)

3) A balloonist fills a balloon with 30000.00 L of helium gas. The temperature is 25.0 °C at ground level and the barometric pressure is 1.000 atm. At a height of 3500 m, the temperature has dropped to a very chilly 4.0 °C and the barometric pressure is 0.800 atm. What is the volume of the balloon’s flexible nylon envelope (gasbag) at this astonishing height? Show all work to receive full credit. (15.00 points)

4) Popcorn, also known as popping corn, is a type of corn that expands from the kernel and puffs up when heated. It is able to pop because, like amaranth grain, sorghum, quinoa, and millet, its kernels have a hard moisture-sealed hull (traps water) and a dense starchy interior. Something interesting happens when the kernel is heated. Use the Kinetic Theory **and** the different gas laws to explain why popcorn pops. In your answer make sure to explain some of the basic components of the Kinetic Theory. You may use pictures/diagrams to help explain your answer. (20.00 points.)
### Appendix J: Final Exam for Academic Chemistry

**Part One: General Chemistry.**

*Multiple Choice (1 point each): Please use the Scantron sheet to answer TWENTY of the following TWENTY-FIVE (25) questions:*

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer Options</th>
</tr>
</thead>
</table>
| 1) Why did Rutherford conclude from his gold foil experiment that an atom is mostly comprised of empty space? | A) The positively charged particles shot into the foil were deflected by the nuclei of the gold atoms.  
B) The positively charged particles shot into the foil were attracted to the electrons of the gold atoms.  
C) Most of the alpha particles shot through the gold foil passed straight through the material. Only a few bounced back, supporting the idea of a dense nucleus.  
D) The radioactive particles shot into the gold foil which caused the gold atoms to give off their own radiation. |
| 2) Which of the following exists as a diatomic element?                  | A) zinc  
B) iodine  
C) helium  
D) magnesium  
E) sodium |
| 5) Which method can be used to separate mixtures based on *boiling points*? | A) distillation  
B) filtration  
C) crystallization  
D) chromatography  
E) centrifugation |
| 6) Which of the following is a good example of a *chemical change*?      | A) Rotting of wood  
B) Melting of ice  
C) Boiling of alcohol  
D) Putting on makeup  
E) Evaporation of gasoline |
| 7) Which of the following is a good example of a *physical change*?      | A) Digesting of lunch |
3) Which method can be used to separate pigments in a plant leaf, pigments in ink or pigments in lipstick?
   A) distillation
   B) filtration
   C) crystallization
   D) chromatography
   E) centrifugation

4) What is the correct molecular formula of hydrazine, also known as dinitrogen tetrahydride?
   A) N₄H₂
   B) N₂H₄
   C) N₂(OH)₄
   D) N₄(OH)₂

B) Curdling of milk
C) Cooking a hamburger or hot dog
D) Cutting up lettuce for a salad
E) Oxidizing of metal

8) Copper (II) sulfate has the chemical formula
   A) CuSO₄
   B) Cu₂SO₄
   C) Cu₂(SO₄)₂
   D) CuS₂O₈

9) Isotopes differ in
   A) the number of electrons in their atoms.
   B) the number of neutrons in their atoms.
   C) the number of protons in their atoms.
   D) the amount of empty space in their atoms.
10) Which of the following statements about the chemistry of the human body is true?

   A) Healthy body temperature is 37°C; it’s mostly made up of Oxygen (65%); and it goes through a series of physical & chemical changes in a lifetime.

   B) Healthy body temperature is 273 K; it’s mostly made up of Carbon (57%); and it only goes through a series of physical changes in a lifetime.

   C) Healthy body temperature is 98.6°F; it’s mostly made up of Water (90%); and it only goes through a series of chemical changes in a lifetime.

   D) Healthy body temperature is 37°F; it’s mostly made up of NaCl (47%); and its metabolism maintains homeostasis.

11) Which of the following is the best definition for chemistry?

   A) Chemistry is the study of matter.

   B) Chemistry is the study of matter and energy.

   C) Chemistry is the study of how matter and energy interact with each other.

   D) Chemistry is the study of physical changes of matter.

   E) Chemistry is the study of chemical changes of matter.

14) What did Newlands, Dobereiner, Mendeleev and Moseley contribute to the development of chemistry?

   A) The law of conservation of matter and energy.

   B) E = mc².

   C) The arrangement of elements into groups or families based on similar properties.

   D) The law of definite proportions.

   E) The discovery of the transuranium elements.

15) Democritus and Dalton are key chemical heroes because:

   A) they helped develop the idea that matter is composed of four elements, earth, wind, water and fire.

   B) they helped develop the idea that matter is made up of indestructible atoms.

   C) they helped develop the wave-particle duality theory.

   D) they discovered that air was not a substance but rather a mixture of nitrogen and oxygen.

   E) they discovered the concept of radiation.
12) The basic components of an atom are:
   A) Protons, Electrons and Neutrons.
   B) Protons, Electrons and Croutons.
   C) Proteins, Lipids and Carbohydrates.
   D) s, p and d orbitals.
   E) Protoplasm, Electrolytes and Neurons.

13) Which key chemical concept was developed by Lavoisier and Joules?
   A) The law of conservation of matter.
   B) The law of conservation of energy.
   C) The law of conservation of matter and energy.
   D) The law of definite proportions.
   E) The law of wave-particle duality.

16) How was the discovery made by Bequerel and the Curies (Marie & Pierre) used to further chemistry?
   A) It was used to discover the nucleus of an atom.
   B) It was used to discover new elements such as Polonium.
   C) It has been used to treat cancer.
   D) It has been used to generate energy (from atomic weapons to power plants).
   E) All of the above.

17) What will determine the distance between an orbiting electron and the nucleus of an atom?
   A) The amount of energy in the electron.
   B) The mass of the electron.
   C) The energy level holding the electron.
   D) The electromagnetic frequency of the electron.
18) To achieve stability sodium atoms will do which of the following?

A) Share electrons to form covalent compounds.

B) Transfer electrons to form ionic compounds.

C) Take seven electrons from chlorine to achieve the valence electron structure of a noble gas.

D) Share seven electrons with chlorine to achieve the valence electron structure of a noble gas.

E) They will become more reactive and chemically bond with noble gases.

19) The Periodic law states that elements show a

A) repetition of their physical properties when arranged by increasing atomic radius.

B) repetition of their chemical properties when arranged by increasing atomic mass.

C) periodic repetition of their properties when arranged by their increasing atomic number.

D) periodic repetition of their properties when arranged by their increasing atomic mass.

20) An atom has net neutral electric charge because

22) What is the complete electron configuration of a scandium (Sc) atom?

A) 1s²2s²2p⁶3s²3p⁶4s²3d¹

B) 1s²2s²2p⁷3s²3p⁷4s²3d¹

C) 1s²2s²3p⁵3s²3p⁵4s²3d¹

D) 1s²2s¹2p⁷3s¹3p⁷4s²3d¹

23) Why is the size of an aluminum atom larger than an atom of silicon?

A) The atomic radii of metal atoms are larger than the atomic radii of nonmetal atoms.

B) The positive charges in an aluminum atom’s nucleus have a greater attraction on the atom’s electron cloud.

C) The positive charges in a silicon atom’s nucleus have a greater attraction on the atom’s electron cloud.

D) The silicon atom contains more electrons and protons that pull on those electrons.

E) None of the above are valid reasons.

24) Which **CANNOT** be used to speed up a chemical reaction?

A) An increase in temperature.

B) An increase in concentration.

C) Add a catalyst.
21) Which element has the electron configuration of [Xe]6s^24f^{14}5d^6?
   A) La
   B) Titanium
   C) W
   D) Osmium
   E) Al

25) What did Boyle, Charles and Gay-Lussac contribute to the field of chemistry?
   A) Their work showed that the behavior of gas particles could be measured and studied.
   B) Their work showed that the size of gas particles could be determined.
   C) Their work determined the number of gas particles commonly found in 1.00 mole.
   D) Their work determined the charge on the proton, electron and neutron.

Part Two: Calculating, Writing and Naming Formulas.

26) You were asked to analyze 50.00 grams of a compound that contained 28.2 grams of silver, 9.27 grams of chlorine; and 12.535 grams of oxygen. Use this data and determine:

- The Empirical Formula
- The Name of the Formula
- The Molecular Mass of the Formula.

Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit. Ten (10) points.

A) AgClO_2; Silver monochlorine dioxide; 176 amu.
B) AgClO_3; Silver chlorate; 192 amu.
C) AgClO_4; Silver chlorate tetraoxide; 208 amu.
D) AgCLO_2; Silver chlorite; 176 amu.
Part Three: Balancing Reactions.

Please use your lined paper to answer the following questions regarding chemical reactions. Each question is worth five (5) points.

For TWO (2) (others may be done for extra credit) of the following reactions (#27 through #30):

A) Write the correct names for the reactants and the products.
B) Write the correct balanced equation with the proper coefficients.
C) Determine the type of chemical reaction taking place (synthesis, decomposition, single displacement, double displacement or combustion).

27) Sb₂S₃ (aq) + HCl (aq) → SbCl₃ (aq) + H₂S (aq)
28) KClO₃ (s) → KCl (s) + O₂ (g)
29) Al (s) + Cl₂ (g) → AlCl₃ (aq)
30) Mg (s) + TiCl₄ (aq) → MgCl₂ (aq) + Ti (s).

For TWO (2) (others may be done for extra credit) of the following reactions (#31 through #34):

A) Write the correct formulas for the reactants and the products.
B) Write the correct balanced equation with the proper coefficients.
C) Determine the type of chemical reaction taking place (synthesis, decomposition, single displacement, double displacement or combustion).

31) Copper (II) metal plus liquid bromine yields liquid copper (II) bromide.
32) Solid lead (II) plus an aqueous solution of silver nitrate yields solid silver and an aqueous solution of lead (II) nitrate.
33) Gaseous tricarbon octahydride (propane) plus oxygen gas yields gaseous carbon dioxide, water vapor and energy.
34) Solid nitrogen tribromide when heated produces nitrogen gas and liquid bromine.

Part Four: Quantifying Reactions and Atomic Behavior.

Please use your lined paper to answer the following questions regarding chemical reactions. Use the data below to answer the following questions (#35 through #42).

- Reactants: Solid Copper (II) plus Aqueous Silver Nitrate
- Mass of white silver nitrate crystals used: 1.20 grams
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

- Mass of copper before experiment: 12.90 grams
- Mass of copper after the experiment: 7.40 grams
- Mass of copper converted in the reaction: 5.50 grams
- Volume of water used: 100.00 mL
- Mass of filter paper and dried precipitate/metal: 19.48 grams
- Mass of filter paper: 1.49 grams
- Mass of dried silver metal: X grams

Answer THREE (3) of the following questions (#35 through #39) the others may be done for extra credit. Each question is worth three (3) points.

35) How many **GRAMS** of the **SILVER** were **ACTUALLY PRODUCED**? Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

   A) 5.50 x 10^0 grams
   B) 1.49 x 10^0 grams
   C) 1.799 x 10^1 grams
   D) 1.79 x 10^2 grams

36) How many **MOLES** of the **SILVER** were **ACTUALLY PRODUCED**? Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

   A) 5.0926 x 10^-2 moles
   B) 2.833 x 10^-1 moles
   C) 1.666 x 10^-1 moles
   D) 1.94 x 10^-3 moles

37) The water evaporated in the drying oven. How many **LITERS** of **WATER VAPOR** evaporated (assuming STP)? Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

   A) 2.24 x 10^3 liters
   B) 1.2444 x 10^2 liters
   C) 1.25 x 10^3 liters
   D) 6.844 x 10^0 liters
38) Determine the **molarity** of the silver nitrate solution used. Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

A) $1.20 \times 10^0 \text{ M}$  
B) $7.06 \times 10^{-2} \text{ M}$  
C) $1.11 \times 10^{-4} \text{ M}$  
D) $7.06 \times 10^{-3} \text{ M}$

39) In another experiment 50.00 grams of silver and displaces 4.54 mL of water. What is the calculated **density and percent error** for this experiment? (3 points)? Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

A) $1.100 \times 10^1 \text{ g/mL with a percent error of 4.18%}$  
B) $1.10 \times 10^1 \text{ g/mL with a percent error of 4.36%}$  
C) $9.08 \times 10^{-2} \text{ g/mL with a percent error of 16.10%}$  
D) $2.27 \times 10^2 \text{ g/mL with a percent error of 78.46%}$

Please answer **TWO (2) of the following questions (#40 through #42)** for credit, the other may be done for extra credit. Each question is worth ten (10) points.

40) **Theoretically** (using stoichiometry...the mole to mole relationship in the balanced reaction, and starting with the mass of Copper converted from data table), how many grams of the **silver metal should have been produced**? Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

A) $1.86 \times 10^1 \text{ grams}$  
B) $1.188 \times 10^3 \text{ grams}$  
C) $1.71875 \times 10^{-1} \text{ grams}$  
D) $3.52 \times 10^2 \text{ grams}$

41) In order for silver to be used by industries it must be extracted from its ores and then melted down. When this metal is placed in a flame it will emit a particular amount of energy, and therefore, a distinctive color. Use the following information to determine the **frequency and color** emitted by this hot metal. The mass of an electron is $9.11 \times 10^{-31} \text{ Kg}$, and the velocity of the electron is $1.26 \times 10^3 \text{ m/s}$. Use the chart of colors located on the formula sheet for verification. Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

A) $5.77 \times 10^{-7} \text{ Hz; color green}$  
B) $5.20 \times 10^{14} \text{ Hz; color yellow-green}$  
C) $3.33 \times 10^{14} \text{ Hz; color red}$  
D) $7.27 \times 10^{-4} \text{ Hz; color blue}$
42) Carbon dioxide is dissolved in a sweetened, flavored liquid solvent to make a carbonated soft drink. Lowering the temperature and raising the pressure are ways to keep the gas dissolved in the liquid. There are, roughly, 400.0 mL of carbon dioxide in a can of soda at a temperature of 22.0 °C and 1.50 atm of pressure. What is the new volume if the can is opened up at room temperature, 25.0 °C and 1.00 atm of pressure? Remember the temperature needs to be in Kelvin! Place your answer in the appropriate space on the Scantron form, however, you must show your work in order to receive full credit.

A) 1.276 x 10¹ mL  
B) 2.34 x 10² mL  
C) 6.06 x 10² mL  
D) 6.81 x 10² mL

Part Five: Nuclear Behavior.

Please use your lined sheet of paper to answer the next set of questions (#42 through 50).

Please answer TWO (2) of the following questions (#43 through #46) for credit, the others may be done for extra credit. Each question is worth five (5) points.

43) Illustrate, using a nuclear equation, what happens when Iodine-131 absorbs an alpha particle.

44) Illustrate, using a nuclear equation, what happens when Thorium-232 absorbs a beta particle.

45) Illustrate, using a nuclear equation, what happens when Uranium-239 absorbs a positron.

46) Illustrate, using a nuclear equation, what happens when Radon-222 absorbs a neutron.

Please answer TWO (2) of the following questions (#47 through #50) for credit, the others may be done for extra credit. Each question is worth five (5) points.

47) Illustrate, using a nuclear equation, what happens when Iodine-131 emits an alpha particle.

48) Illustrate, using a nuclear equation, what happens when Thorium-232 emits a beta particle.

49) Illustrate, using a nuclear equation, what happens when Uranium-239 emits a positron.

50) Illustrate, using a nuclear equation, what happens when Radon-222 emits a neutron.

Part Six: What is Chemistry?

Please answer ONE (1) of the following questions (questions #51 through #53). The others may be done for extra credit. Make sure your written response is legible and coherent. Write your answer directly on your lined sheet of paper. Each question is worth ten (10) points.
51) Chemistry is a broad science that touches nearly every aspect of human life. Take soda for example. Over 15 billion gallons of soda are sold every year! That’s a lot of chemistry! If a can of soda is shaken up and then opened, a small disaster will occur. Using the ideas of the Gas Laws and the Atomic/Kinetic Theory, explain why this would happen.

52) Throughout this course there have been many chemical reactions illustrated and studied. Explain how a balanced chemical reaction supports the Law of Conservation of Matter and Energy. Why is this law so important to chemistry? Use examples to illustrate your point.

53) Although Element #117 (Ununseptium) was supposedly synthesized at the Berkeley Lab in 2001, the results could not be verified until 2010. Although it hasn’t been studied extensively, we can nevertheless make some reasonable predictions of its chemical and physical properties. What can be said about this element with regards to its chemical and physical properties as compared to another element in the same family? Include the following:

- Electron configuration for the element.
- Electron configuration for the ion.
- The state of matter at room temperature.
- The atomic radius.
- The density.
- The ionization energy.
- The electro-negativity.
- The reactivity.
- Three other properties (for a total of ten) that you can include.
Appendix K: POGIL Lab on Electrons and Wave Behavior

Electron Energy and Light

How does light reveal the behavior of electrons in an atom?

Why?

From fireworks to stars, the color of light is useful in finding out what's in matter. The emission of light by hydrogen and other atoms has played a key role in understanding the electronic structure of atoms. Trace materials, such as evidence from a crime scene, lead in paint or mercury in drinking water, can be identified by heating or burning the materials and examining the color(s) of light given off in the form of bright-line spectra.

Model 1 – White Light

<table>
<thead>
<tr>
<th>Color</th>
<th>Photon Energy (\times 10^{-31}) (J)</th>
<th>Wavelength Range (nm)</th>
<th>Speed (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reds</td>
<td>269–318</td>
<td>625–740</td>
<td>3.00 (\times 10^8)</td>
</tr>
<tr>
<td>Oranges</td>
<td>318–337</td>
<td>590–625</td>
<td>3.00 (\times 10^8)</td>
</tr>
<tr>
<td>Yellows</td>
<td>337–352</td>
<td>565–590</td>
<td>3.00 (\times 10^8)</td>
</tr>
<tr>
<td>Greens</td>
<td>352–382</td>
<td>520–565</td>
<td>3.00 (\times 10^8)</td>
</tr>
<tr>
<td>Blues</td>
<td>382–452</td>
<td>440–520</td>
<td>3.00 (\times 10^8)</td>
</tr>
<tr>
<td>Violets</td>
<td>452–523</td>
<td>380–440</td>
<td>3.00 (\times 10^8)</td>
</tr>
</tbody>
</table>

1. Trace the arrows in Model 1 and shade in the table with colored pencils where appropriate.

2. What happens to white light when it passes through a prism?

3. Why are the color labels in the table in Model 1 plural (i.e., “Reds” rather than “Red”)?

4. Do all colors of light travel at the same speed?

5. Do all colors of light have the same energy? If no, which colors have the highest energy and the least energy, respectively?

6. Consider the light illustrated in Model 1.
   a. Which color corresponds to the longest wavelengths?

   b. Which color corresponds to the shortest wavelengths?

   c. Write a sentence that describes the relationship between wavelength and energy of light.

---

Electron Energy and Light

© 2012, Flex Scientific, Inc., and HSGI. The POGIL Project. All Rights Reserved. Reproduced for one-time use with permission from Flex Scientific, Inc., Batavia, Illinois, U.S.A. No part of this material may be reproduced or transmitted in any form or by any means, electronic or mechanical, including, but not limited to, photocopying, recording, or any information storage and retrieval system, without permission in writing from Flex Scientific, Inc.
Model 2 – Emission Spectra for Hydrogen and Boron Atoms

7. Use colored pencils to color the hydrogen and boron spectral lines within their respective spectra in Model 2.

8. List the spectral lines for hydrogen gas by color and corresponding wavelength.

9. The spectral lines for boron were produced using the same method as hydrogen. List three of the colors and corresponding wavelengths for boron’s spectral lines as its light passes through a prism.

10. Consider the hydrogen spectrum in Model 2.
   a. Which color of light corresponds to the shortest wavelength?
   b. Which color of light corresponds to the longest wavelength?
THE COMPARISON OF TWO DIFFERENT GUIDED INQUIRY METHODS IN HIGH SCHOOL CHEMISTRY CLASSES

11. Consider the hydrogen spectrum in Model 2.
   
   \(a\). Which color of light has the most energy?
   
   \(b\). Which color of light has the least energy?

12. Does a gas discharge tube filled with boron emit the same wavelengths of light as a tube filled with hydrogen? Use evidence from Model 2 to support your answer.

13. “The spectral lines for atoms are like fingerprints for humans.” How do the spectral lines for hydrogen and boron support this statement?

Circle the appropriate word to complete each statement in Questions 14–17.

14. Electrons and protons (attract/repel) each other.

15. As an electron gets closer to the nucleus the (attraction/repulsion) to the nucleus gets (stronger/weaker).

16. For an electron to move from an energy level close to the nucleus to an energy level far from the nucleus it would need to (gain/lose) energy.

17. For an electron to move from an energy level far from the nucleus to an energy level close to the nucleus it would need to (gain/lose) energy.

---

Read This!

Niels Bohr modified Rutherford’s Nuclear Atom model to explain how light interacted with the electrons in an atom to produce spectral lines. His model included electrons orbiting the nucleus at specific energy levels. Electrons absorb energy from various sources (electricity) when they move from lower energy levels (ground state) to higher energy levels (excited states). Energy is released as electrons return to their lower energy levels.

18. Is energy absorbed or released for the electron transition shown in the diagram to the right? Explain.

---

Electron Energy and Light
19. Identify the drawing in Model 3 that depicts a hydrogen atom with an electron moving from energy level 5 to energy level 2. Refer to Models 1 and 2 for the following questions.
   a. Label the picture with “n=5 to n=2” and list the corresponding color of light emitted.
   b. This electron transition (absorbs/releases) energy.
   c. This electron moves from a (lower/higher) energy state to a (lower/higher) energy state.
   d. Is light absorbed or released in the electron transition?
20. Label the remaining drawings in Model 3 with the electron transitions that are occurring \((n=? \text{ to } n=?),\) the wavelengths and corresponding colors as given in example A in Model 3. See Model 2 in order to identify the color of spectral lines produced in each of the hydrogen atom electron transitions shown in Model 3. Use colored pencils to trace the light wave in each of the four pictures with the appropriate color.

21. Consider the electron transitions in Model 3.
   
   \(a.\) Which of the electron transitions involves the most energy?

   
   \(b.\) Explain why this transition involves the most energy based on your understanding of the attractive forces between the electrons and protons in the atom.

22. Explain why a single atom of hydrogen cannot produce all four hydrogen spectral lines simultaneously.

23. If Question 22 is true, how can we see all four colors from a hydrogen gas discharge tube simultaneously?
Extension Questions

24. The hydrogen spectral lines in Model 2 are only the wavelengths of light that are in the visible range and therefore “seen” by the naked eye. However, many other wavelengths can be detected with special equipment.

   a. Propose a hydrogen electron transition that involves light with a wavelength in the ultraviolet (UV) range (10–400 nm).

   b. Propose a hydrogen electron transition that involves light with a wavelength in the infrared (IR) range (1000–106 nm).

25. Below are diagrams for the bright line spectra of four elements and the spectrum of a mixture of unknown gases.

   Li
   H
   He
   Na
   Unknown

   a. Which element(s) are not present in the Unknown?

   b. Which element(s) are in the Unknown?

26. Model 2 shows the emission spectra for hydrogen and boron. Scientists can also record the absorption spectra for elements. Propose how this might be done, and what the absorption spectra of hydrogen and boron would look like.
Appendix M: POGIL Activity on Empirical Formulas

Empirical Formula and Molecular Formula
Which formula is more informative?

Model 1: Comparison of Percent Composition and Empirical Formula

<table>
<thead>
<tr>
<th>Substance</th>
<th>Compound Name</th>
<th>Molecular Formula</th>
<th>Empirical formula</th>
<th>Percent Composition: Carbon</th>
<th>Percent Composition: Hydrogen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ethane</td>
<td>C₂H₆</td>
<td>CH₂</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>propene</td>
<td>C₃H₈</td>
<td>CH₃</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ethyne (acetylene)</td>
<td>C₂H₂</td>
<td>CH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>benzene</td>
<td>C₆H₆</td>
<td>CH</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Divide up the work within your team and calculate the percent composition for substances in the table in Model 1. Put the values into the table. Show your calculation(s) below.

2. Identify the substances in Model 1 that have the same empirical formula.

3. Identify the substances in Model 1 that have the same percent composition.

4. Using the examples in Model 1, suggest an explanation as to why substances with different molecular formulas (but the same empirical formula) can have the same percent composition.

5. For each substance in Model 1 you will notice that the molecular formula and empirical formula are not the same. What mathematical operation would enable you to determine the empirical formula of a compound if you are given the compound’s molecular formula?
6. Using the operation that you suggested in question 5, determine the empirical formula for the compounds with the following molecular formulas:
   a. C_6H_8
   b. C_6H_{12}

7. Which substance(s) in Model 1 have the same empirical formula(s) as the substances in Question 6?

8. Predict the percent composition(s) for C_6H_8 and C_6H_{12}. Check with your teacher to see if you have the correct percent composition.

9. As a chemist, if you were given the choice of knowing the molecular formula or empirical formula for a substance, which formula would give you more information so that you could uniquely identify the substance? Explain your choice using examples from Model 1.

---

Model 2: Determining Molecular Formulas

<table>
<thead>
<tr>
<th>Substance Name</th>
<th>Molecular Mass (g) (MF mass)</th>
<th>Empirical formula</th>
<th>A Empirical formula mass (g) (EF mass)</th>
<th>B Ratio of (MF mass) (EF mass)</th>
<th>C Representing # of EF units in each MF</th>
<th>D Molecular Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane</td>
<td>16.0</td>
<td>CH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethyne</td>
<td>26.0</td>
<td>CH</td>
<td>13.0</td>
<td>26.0/13.0 = 2 units</td>
<td>(CH)_2</td>
<td>C_2H_2</td>
</tr>
<tr>
<td>Benzene</td>
<td>78.0</td>
<td>CH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethene</td>
<td>28.0</td>
<td>CH_2</td>
<td>14.0</td>
<td>28.0/14.0 = 2 units</td>
<td>(CH)_2</td>
<td>C_2H_4</td>
</tr>
<tr>
<td>Propene</td>
<td>42.0</td>
<td>CH_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclohexane</td>
<td>84.0</td>
<td>CH_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethane</td>
<td>30.0</td>
<td>CH_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
10. Determine the empirical formula mass for each substance in Model 2 and enter it into Column A. Be sure that your entire group agrees on these values.

   a. Do the substances with the same empirical formula have the same empirical formula mass?

   b. Do these substances have the same molecular mass?

11. Determine the (MF mass)/(EF mass) ratio for each substance and enter it into Column B of the table in Model 2. Be sure that your entire group agrees on the values.

12. Look at the information in Column C and Column D that is given for ethyne and ethene and write a complete sentence to describe the relationship between the ratio in Column B and the molecular formula for each of the compounds.

13. Using your answer to Question 12 as a guide, complete Columns C and D for all of the substances in Model 2. Again, be sure that all members of the group agree on the information that is in the table.

14. If compounds have the same empirical formula, what must be true about the molecular formulas of these compounds?

15. One of the compounds in Model 2 has the same empirical formula and molecular formula. Name the compound and indicate what information this conveys about the compound.

16. Determine the molecular formula for the following compounds:

   a. Empirical formula is NaO, mass is 78 g/mole. ________________

   b. Empirical formula is CH₂Cl, mass is 99.0 g/mole. ________________

   c. Empirical formula is C₆H₆, mass is 121 g/mole. ________________
Extension Questions:

17. An oxide of nitrogen is found to contain 69.6% oxygen and has a molar mass of 92.0 g/mole.
   a. What is the % nitrogen in this compound?

   b. Find the empirical formula and molecular formula for this compound.

18. Calculate the molecular formula for caffeine, a compound with a molar mass of 195 g and the following percent composition: 49.5% C, 5.15% H, 28.9% N, 16.5% O.