University of Northern Iowa

UNI ScholarWorks

Honors Program Theses | Student Work

5-2020

Measuring meaningful learning in a chemistry laboratory

Kylie Engstrom
University of Northern Iowa

Let us know how access to this document benefits you

Copyright ©2020 Kylie Engstrom

Follow this and additional works at: https://scholarworks.uni.edu/hpt

Part of the Higher Education Commons, and the Science and Mathematics Education Commons

Recommended Citation

Engstrom, Kylie, "Measuring meaningful learning in a chemistry laboratory" (2020). *Honors Program Theses*. 416.
https://scholarworks.uni.edu/hpt/416

This Open Access Honors Program Thesis is brought to you for free and open access by the Student Work at UNI ScholarWorks. It has been accepted for inclusion in Honors Program Theses by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

Offensive Materials Statement: Materials located in UNI ScholarWorks come from a broad range of sources and time periods. Some of these materials may contain offensive stereotypes, ideas, visuals, or language.

MEASURING MEANINGFUL LEARNING IN A CHEMISTRY LABORATORY

A Thesis Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors

Kylie Engstrom

University of Northern Iowa

May 2020

This Study By: Kylie Engstrom

Entitled: Measuring Meaningful Learning in a Chemistry Laboratory

has been approved as meeting the thesis or project requirement for the Designation University Honors.

__________ ____

Date Dr. Dawn Del Carlo, Honors Thesis Advisor, Department of Chemistry, Biochemistry and Science Education

__________ ____

Date Dr. Jessica Moon, Director, University Honors Program

Abstract

In the past, students in general chemistry classes at the University of Northern Iowa have struggled with pre-labs due to cognitive overload. To better prepare students for the lab portion of the class and to reduce the amount of cognitive load, professors can provide students with strategies to scaffold information for students. The CLUE General Chemistry curriculum, which emphasizes the use of scientific practices, paired with a new SWH/ADI hybrid laboratory model was implemented in one of the sections of CHEM 1110 in the Fall 2019 semester. The goal of including these new features was to provide more scaffolding for increased student comprehension while also improving students' skills and experiences in the lab. In order to measure students' perceptions changed during the course of the semester, the Meaningful Learning in the Laboratory Instrument (MLLI) was used to measure the cognitive and affective domains of the students' laboratory experience. Through examining the cognitive and affective areas, it is possible through using the MLLI survey to determine how student perceptions and experiences have changed throughout the semester with the use of the hybrid ADI-SWH and the SWH lab approach.

INTRODUCTION

Work in the Lab

Laboratory courses are a common part of chemistry classes both for high school students and undergraduate students at universities. Laboratory work allows students to improve their observation and critical thinking skills, especially at the introductory level where students are first introduced to the material. Students often face difficulties in chemistry due to the overreliance on memorization. Instead of learning content, students are memorizing bits and pieces for an assessment, then forgetting it (Grove and Bretz, 2012). By doing this, it is clear that laboratory classes are not effective in allowing students to understand concepts at a deeper level (Bretz, 2019). The National Research Council (2012) states that laboratory practices that include realistic scientific practices and incorporate the use of technology are the most effective at helping students understand and know chemistry, not just simply memorize concepts. Including more effective practices in the lab would allow students to step away from memorization and truly learn chemistry concepts in the laboratory setting.

Prelabs and Cognitive Load

Most general chemistry laboratory curriculum includes prelabs, or work that is done outside of class by students in preparation before they come into the lab. The prelab allows students to gain background knowledge and recognize key ideas for the lab they will be doing. Ultimately, the goal of prelab activities is to reduce cognitive load for students (Hofstein & Lunetta, 2004; Llorens-Molina, 2008). Cognitive load theory focuses on “the development of instructional methods that efficiently use people’s limited cognitive processing capacity” (Paas et al., 2003, p. 64). Cognitive load theory is a “universal set of learning principles that are proven to result in efficient instructional environments as a consequence of leveraging human cognitive learning processes” (Clark, Nguyen, & Sweller, 2005, p. 7). Because the cognitive capacity for short term memory is about 7 +/- 2 chunks, the space cannot be wasted for efficient learning to happen. (Paas et al., 2003). Through the use of specific principles, the learning process occurs with more ease and in a shorter amount of time. Prelabs can be used in this way to introduce topics to students before they go into the lab, allowing students time to develop background knowledge about what they will be doing, as opposed to being overwhelmed. The most effective methods of instruction should be developed and used in the classroom to avoid overloading students’ brains with information (Clark et al., 2005).

There are three main types of cognitive load: intrinsic load, germane load, and extraneous load. Intrinsic load is defined as the mental work imposed by a specific topic or goal. The amount of intrinsic load can be managed through breaking down complex tasks into simpler ones and providing supporting information. Germane load is mental work that is put in to create knowledge related to goals that will stay in permanent storage. Organizing students into doing different tasks that all relate to the same activity and skill development is one way to increase germane load (Clark et al., 2005). Finally, extraneous (or irrelevant) load is a type of mental work that does not help learning and wastes limited storage in the brain. The limited storage space that is being used up could be used for germane load. Higher levels of extraneous load in activities results in more time needed to learn, poorer learning outcomes, or a combination of both (Clark et al., 2005). Intrinsic load and germane load are useful types of cognitive load, while extraneous load wastes mental resources. To minimize cognitive load in prelabs, extraneous load must be reduced, and intrinsic load decreased through scaffolding.

Current Prelabs in General Chemistry courses at UNI

Previous research showed that the prelabs assigned as part of the Science Writing Heuristic laboratory model used at the University of Northern Iowa were not effective in reducing cognitive load (Bonde & Del Carlo, 2018). Students were stressed and confused with what content they needed to understand and consequently unprepared coming into the lab. Due to cognitive overload, students were unable to decide which information they need, what the task of the lab is, or how to accomplish the lab. Students are simply overwhelmed by the amount of information they are expected to process, revealing there is too much extraneous load associated with this laboratory approach (Bonde & Del Carlo, 2018). Because of this, changes in the laboratory model and curriculum, including the prelab, were developed. This new format was created to scaffold information, so students have a better understanding of ideas and are much less confused when in the lab. Through this scaffolding, the intrinsic load was increased for students while the extraneous load was minimized. Since less mental workspace is being used for irrelevant information, more mental capacity can be used for germane load to commit ideas and concepts to long term storage.

The CLUE Curriculum

Advocating for change in the general chemistry curriculum comes with many challenges, such as going over too much material or not developing an understanding of basic chemistry concepts that must be built upon in future chemistry classes (Cooper, 2010). To address these problems, a new general chemistry curriculum called Chemistry, Life, the Universe and Everything (CLUE) was created. This model is based on five important but basic questions: (1) What should students know? (2) In what order should they learn it? (3) What do students bring with them to the course? (4) What materials are best suited for different purposes? (5) How can student understanding be assessed? These questions can directly help address the core concepts and ideas of the course which include molecular structure, macroscopic properties, energetics and thermodynamics. The curriculum has the ability to constantly evolve using student feedback (Cooper and Klymkowsky, 2013). This approach is not narrow and focused, and therefore can be used in many different laboratory settings.

The CLUE curriculum builds in strategies that benefit student learning. Learning progressions are conceptual and logical pathways that are more likely to lead to a deep understanding for students. These progressions also allow students to learn and develop ideas over time and not all at once (Corcoran et al., 2009). In addition to this, Novak found that the information students already know is extremely valuable to learning, if not the most valuable thing (Novak 2002). This idea is clearly reflected within the curriculum model through eliciting student thinking. The CLUE materials are flexible and the text is shorter than traditional chemistry textbooks, allowing for the use of online resources like videos or audio clips. Through these resources other than the textbook, problem solving descriptions or creating representations are easier (Mayer, 2009). Students are presented with real data they must analyze and struggle with. Through this struggle, more active learning occurs within the classroom (Cooper and Klymkowsky, 2013).

While originally designed for the non-laboratory component of general chemistry, all of the strategies specific to the CLUE curriculum give students a better opportunity to deepen their knowledge in a laboratory setting, not just in the classroom. The CLUE curriculum emphasizes the use of scientific practices such as developing and using models, formulating arguments from evidence, and developing scientific explanations (Cooper & Klymkowsky, 2013). These science practices demonstrate how science is done, not just learned. Therefore, it is conceivable that students can perform higher level scientific skills in the laboratory due to how the CLUE curriculum prepares them in the non-laboratory component of their class.

The Science Writing Heuristic

Past research found that students learn laboratory techniques, but little else within a chemistry lab (Hofstein & Lunetta, 2004). While learning laboratory techniques is an important part of the lab, the real goals of laboratory courses are to experience and understand both the scientific phenomena and content that relates to the specific class a student is in. The Science Writing Heuristic (SWH) laboratory model allows students to meet these goals and is currently used in general chemistry laboratories at the University of Northern Iowa. There are two main areas emphasized in the SWH model. The first is collaborative learning, which has been shown to benefit students sociologically, personally and academically (Bowen, 2000; Smith et al., 1991). Collaborative learning has been proven to benefit students more than traditional techniques and allows students to learn tough topics together in a non-threatening and non-judgmental environment (National Research Council, 2000). The second area emphasized is writing, which is important to learn for using scientific language. The writing component of the SWH model, specifically writing evidence based claims, is based on the idea that the process of writing leads to deeper understanding of developing concepts (Graham & Hebert, 2011). Through these ideas, the SWH model is an effective laboratory model that differs from traditional laboratory practices but improves understanding for students.

Features and drawbacks of SWH

The Science Writing Heuristic strategy has many positive effects on learning in the lab. The SWH includes exercises that allow students to engage with the ideas of science (Keys, Hand, Prain, & Collins, 1999) and teaches students to recognize what exactly it means to construct scientific knowledge (Burke, Greenbowe, & Hand, 2006). The model also leads to improved performances on lecture exams, which students attributed to better conceptual understanding of topics (Burke et al., 2006). However, there are some drawbacks with the SWH model. The University of Northern Iowa has been using the Science Writing Heuristic in their general chemistry curriculum. However, the pre-lab activities that go along with the model have not been effective in reducing cognitive load for students (Bonde & Del Carlo, 2018). The pre-labs do not appear to provide students enough information to understand the concepts of the lab. In order to improve student learning, there must be changes made to the pre-labs.

Argument-Driven Inquiry

Another relatively new laboratory instructional model is called Argument-Driven Inquiry or ADI (Sampson et al., 2011). It was developed specifically for undergraduate college chemistry lab courses and “helps students to learn important content, practices, and habits of mind during a laboratory course” (Sampson et al., 2011, p. 225). What makes the ADI model different from other traditional general chemistry models is how it provides students with the opportunity for scientific social interaction. A large emphasis is placed on students talking to their peers about their ideas while exploring what others’ thoughts are. Two steps that are specific to ADI are the class claims discussion and peer review step. Students present their claims to their classmates for comparison and critique. Following this, students craft a lab report and submit them for peer review that is done during class time. Through these steps, students develop ideas, and get them “validated” by their peers and professors and demonstrate to the students how true knowledge is created (Driver et al., 1994).

The ADI model incorporates 7 steps, starting from introducing the task/research questions, then developing a plan to gather the data and carrying it out. Afterward, groups create arguments that are presented to the other groups in the class in a discussion called the "argumentation session" where arguments are shared and critiqued (Walker, Sampson, & Zimmerman, 2011). Students then create their reports and subsequently get their report reviewed by a set of peers. Unlike other models which incorporate one lab a week, the ADI method maintains a theme over several weeks that starts with pre-lab exercises specific to the lab. The next week, the investigation is performed and followed immediately by the argumentation session, all within the same class period. The lab is finally wrapped up with the peer review the following class period, followed by the pre-lab activities for the next lab that will be done. (Walker, Sampson, & Zimmerman, 2011).

Features and drawbacks of ADI

Throughout the course of the semester using ADI, Sampson & Walker (2012) found that student scores on the investigation report improved significantly, with the most improvement occurring in the quality of students' arguments. Further studies indicate that not only do scores improve, but students' ability to write scientifically improves too. Of the parts of the ADI model, students found the peer review and opportunity for revision to be the most helpful (Walker & Sampson, 2013). The ADI model is a highly structured model with all parts of the model being interrelated (Walker et al., 2011). The ADI model provides students with a great amount of scaffolding by including devices and strategies that support student learning, such as providing background information and beginning questions in a lab setting (van Merrienboer et al., 2003).

The best of both worlds

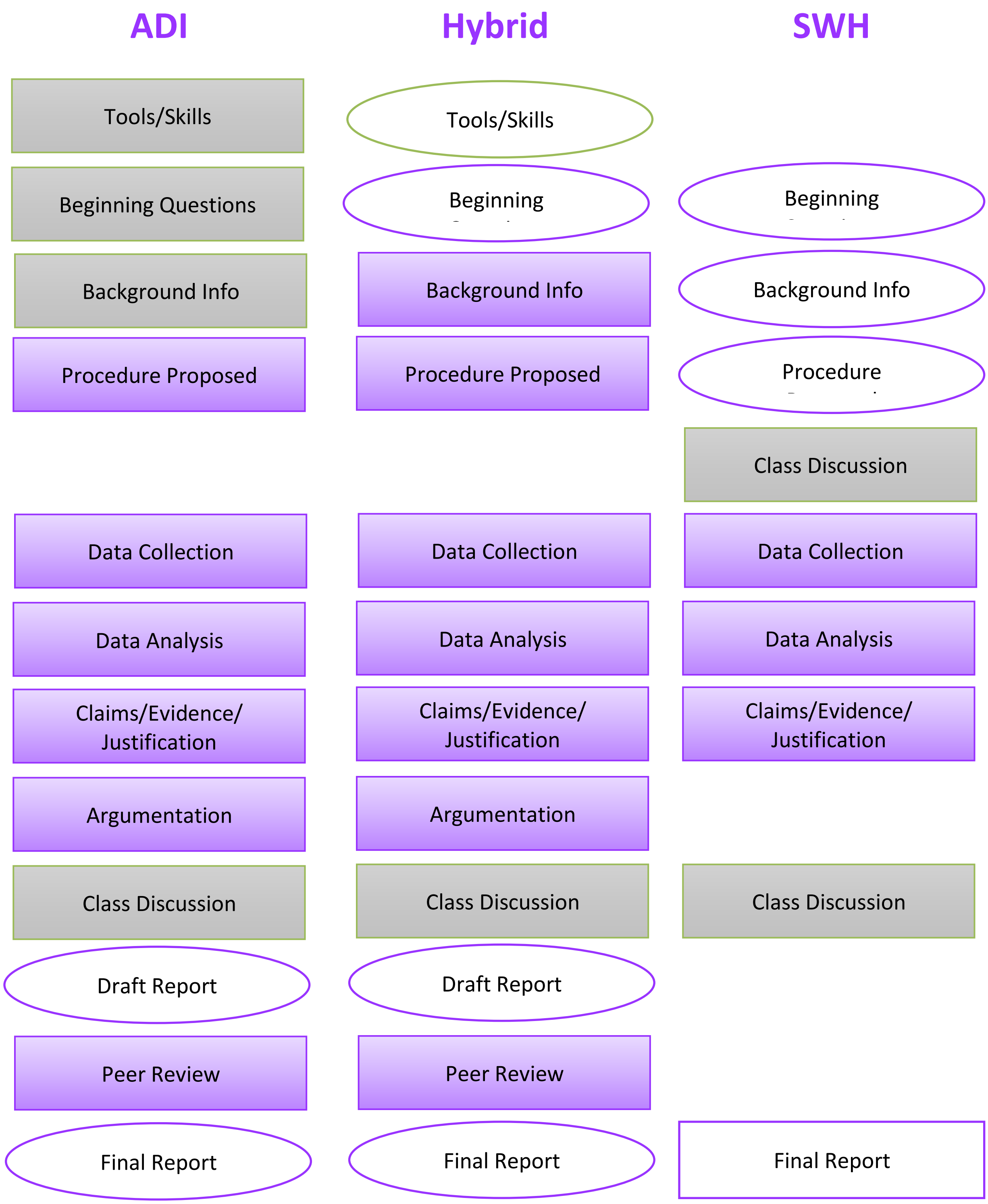
The ADI model pre-lab gives students the majority of the information they need to start the lab, while the SWH provides much less and has the students work independently to figure out the logistics of the lab (Walker et al., 2011). Because of this, a middle ground would be the most suitable for students. An intermediate amount of scaffolding will provide students with the tools they need to reduce their cognitive load, while omitting some parts of the process for students to figure out. Because an intermediate amount of scaffolding would be ideal, a hybrid model between ADI and SWH was created. The hybrid model aims to reduce cognitive load for students while still allowing them the freedom to determine the details of their own experiment. When the hybrid model was created, there were parts selected from both the ADI and the SWH models in order to improve the scaffolding for students.

Specifics of the hybrid model

The first few steps in both models focus on tools and skills needed for the lab, beginning questions, background information, and the procedure. The ADI model provides all of that information for students (Walker et al., 2011), while the SWH model has students do all of those parts on their own (Burke et al., 2006). The hybrid model starts with students doing tutorials on tool use and lab skills outside of class. Instead of having students create their own procedure individually like in the SWH model, students work with their group in class to come up with their own questions for the experiment and unique procedure after deciding on the beginning questions the group wishes to explore. The instructor-led class discussion in the SWH model was replaced with instructor guided small group interactions. The data collection, analysis and claims/justification sections were included in all of the models with students working with their peers in class. The argumentation session that is unique to the ADI model was also included in the hybrid, allowing students to state, support and justify their claims and see others' claims from the experiment (Walker et al., 2011). A class discussion after the argumentation session is included in the hybrid model, pulled from both models to discuss the results and ideas generated from the lab. Students then individually create a lab report and submit it for peer review in class, This idea is based on the ADI model (Sampson & Walker, 2012). Through the hybrid model, it is hoped that students are provided enough scaffolding to avoid cognitive overload but are also able to have more choice and ownership when it comes to the procedure. In addition to this, students are able to utilize their instructor during the pre-lab and their peers as tools to create the best final report possible and truly understand the purpose of the lab.

Figure 1 shows the ADI, SWH, and new hybrid model. The rectangles represent group work, the circles represent individual work, the shading represents activities done in class vs outside of class, the purple represents full student work, and the gray represents professor involvement in the lab. Figure 1 is found on the following page.

Figure 1. Lab model comparison



RESEARCH QUESTIONS

Because of the development of the new hybrid model, we were interested in seeing the students' thoughts about their general chemistry experience. One way to do this was to see if there were significant differences between the pre-survey and post-survey results As such, the research question that guided this project was:

> How are the cognitive and affective expectations of students met through the use of the hybrid Science Writing Heuristic/ADI model and a general Science Writing Heuristic model in an undergraduate general chemistry course?

METHODS

The MLLI Tool

Throughout the chemistry education community, there is a lack of evidence that laboratory aspects of classes greatly impact student learning. Because of the unavailability of effective assessment tools in the past, the Meaningful Learning in the Laboratory Instrument (MLLI) was created based off of Joseph Novak's theory of Meaningful Learning and Human Constructivism. Novak's theory states that, "Meaningful learning underlies the constructive integration of thinking, feeling, and acting leading to human empowerment for commitment and responsibility" (Novak, 1998, p. 18). The experience a person has depends on how the person thinks, feels, and acts in that moment. According to Galloway and Bretz (2015), meaningful learning occurs when the cognitive (thinking), affective (feeling) and psychomotor (doing) domains are all active and engaged. Students must use cognitive and affective strategies to perform their laboratory work (psychomotor), meaning these ideas can easily be applied to that setting.

The MLLI was created to measure meaningful learning in the cognitive and affective domains, as well as measure students' expectations of the lab course. Being set up in a survey design, the MLLI contains some questions targeted toward the cognitive domain, others related to the affective domain, and additional questions focused both on the cognitive and affective domains. By going beyond if a student simply "likes" the course, the MLLI allows professors and researchers to apply improved teaching styles and curriculum to their classes (Galloway & Bretz, 2015). In order to examine general chemistry students' thoughts on their laboratory experiences in the Fall 2019 semester, the MLLI was sent out at the beginning and end of the semester (Appendix A). A detailed timeline of events can be found below (Table 1). This project was done in collaboration with Tabitha Alitz, a fellow undergraduate student. However, we each collected different data to answer different research questions. All procedures for this study were approved by the UNI Institutional Review Board (HP Protocol # HP 19-0180).

Table 1. Timeline of planned events

Events	Date of Events
Recruited students from general chemistry classrooms & sent email inviting them to take the MLLI pre-survey	Week 1 of Fall 2019 Semester 08/26 - 08/30
Pre-survey portion closed	09/05
Analyzed the Pre-survey	09/05-11/05
Went into general chemistry classes again, thanked them for participation and sent email inviting them to take the post-survey	Last Week of Classes for Fall 2019 Semester 12/9 - 12/ 13
Post-survey portion closed	12/20
Looked at the results of the post-survey, compared with the results of the pre-survey, analyzed and made conclusions	12/20 - 4/30/20

DATA COLLECTION

The main focus of the project dealt with the laboratory models and curriculum for general chemistry classes at UNI. One “experimental” section of general chemistry (out of six offered) used the hybrid ADI/SWH laboratory model with experiments whose content paired with the CLUE lecture curriculum. The existing SWH laboratory model was used in all other sections. The intent of the study was to compare the experimental hybrid section to the sections that did not use the hybrid and instead used the SWH model as a guide. This way, we could identify what the differences are between the two sections. However, within the formation and logistics of the survey, we did not collect any information to discern the two sections apart from one another. Because of this, the plan for the research had to change. Instead of comparing the sections, the focus of the project changed to see if the expectations for students are met in the cognitive and affective aspects in general. The hybrid section and regular sections no longer were compared. While this is unfortunate, the hope was to still gather meaningful data that can help inform the curriculum and lab format for future years.

The primary and most important step in the process was gathering participants. During the first week of classes, the researchers visited all of the general chemistry lab classrooms and invited the students to participate. That same day the mandatory consent form, general guidelines, and survey link in email form was sent out to students formally inviting them to take the survey. The participants took the survey within a week of the researchers inviting them to participate. To incentivize students to participate in the study, $25 Amazon gift cards were given out to three random participants who have taken any part of the survey. We were successful in recruiting participants, with 38 filling out the pre-survey and 32 filling out the post-survey.

DATA ANALYSIS

The results of the preliminary MLLI survey allowed us to see what expectations students had for the course, regardless of the curriculum and model types. By breaking down the results into descriptive statistics, it was possible to look individually at the cognitive, affective and cognitive/affective expectations of students. The mean and standard deviation for the cognitive, affective, and a mix of the cognitive/affective revealed a general idea of what the students expected before starting the lab. Cronbach’s alpha was also included to illustrate the reliability of the survey.

Along with general descriptive statistics, t-tests were performed on each of the 3 domains that were present in the MLLI. These t-tests were run to see if there was a significant difference between the pre to post survey scores. T-tests were also run for each individual category within the cognitive domain after that domain was the only one found to be statistically significant.

Reliability and Validity

Table 2 shows Cronbach's alpha results for all three of the categories, both pre and post. The results all were above the 0.7 threshold, reaching in some categories up to 0.9. Cronbach’s alpha is a measure of the average correlation within students’ answers (Galloway and Bretz, 2015). Consistency between the affective results were excellent, while the two other categories for both pre and post survey were at a “good” level. The values also appeared to be consistent pre to post as well as in each category. These results demonstrated the survey was highly reliable. We figured this would be the case because the survey was copied from Galloway and Bretz’s work.

Because we based our study off of the MLLI survey used by Galloway and Bretz in 2015 that was deemed valid, the assumption was made that the survey we used was also valid. The only differences between the original use of the MLLI and the use of the MLLI survey in this project was the addition of the questions dealing with the prelab, which was based off of the original questions developed on the MLLI.

RESULTS

Table 2 shows the descriptive statistics for both the pre and post MLLI survey. The mean and standard deviation were calculated for each category both pre and post survey.

Table 2. Descriptive statistics for MLLI pre- and post-test administrations

		Pre-survey				Post-survey	
	Cognitive	Cogn/Aff	Affective		Cognitive	Cogn/Aff	Affective
Mean (SD)	69.36 (24.61)	58.21 (29.53)	56.55 (32.31)		58.43 (28.74)	50.53 (30.94)	56.44 (33.62)
□	0.774 Acceptable	0.750 Acceptable	0.904 Excellent		0.872 Acceptable	0.776 Acceptable	0.922 Excellent

All scores from pre to post decreased, meaning that students' expectations of the lab section of the general chemistry courses were not met (Table 2). The internal reliability within each category was found to be reliable. Cronbach's alpha was included in these statistics. The criteria used to define Cronbach's alpha are those recommended by Mohsen and Dennick (2011) where $\alpha \geq 0.7$ is "Acceptable", $\alpha \geq 0.8$ is "Good", and $\alpha \geq 0.9$ is "Excellent". Cronbach's alpha was chosen to be part of the data analysis because it was in the original analysis of the first use of the MLLI survey (Galloway & Bretz, 2015) and told us about the internal reliability of the survey.

The results of the t-tests run on the three domains of the MLLI survey were also calculated (Table 3). These t-tests were run to measure the probability that the differences in the survey from pre and post are due to chance. The number of questions in each domain is listed below due to the number of questions possibly having an effect on the results of the t-test. Effect size was also listed to provide a measure of size weighted according to the relative size of each sample. Hedges' *g* was chosen to measure effect size due to the different sample sizes in each category (Table 3).

Table 3. T-tests for each surveyed domain with effect size

Domain	Number of Questions within Domain	P-Value	Results	Effect Size (Hedges' *g*)
Cognitive (Pre to Post)	22	0.0001	Statistically significant	0.41 Small effect
C/A (Pre to Post)	8	0.26	Not statistically significant	N/A
Affective (Pre to Post)	12	0.98	Not statistically significant	N/A

Results from the statistical analysis indicated that the cognitive domain scores were the only ones to change from pre- to post (Table 3). Both the affective and cognitive/affective domains were not statistically significant, meaning the differences in scores was purely due to chance. We must consider that the cognitive domain had the largest amount of questions, while the other two non-significant categories had about half of the amount of questions in comparison or even less. Having less questions could have led to a lesser chance of the data in those groups being statistically significant. The effect size was only applicable for the cognitive group because it was the only category found to be statistically significant. The cognitive effect size was found to be minimal.

Because the cognitive domain was the only domain found to be statistically significant, we chose to further analyze groups of individualized categories within that domain. Groups were chosen by the researchers by looking at individual questions within the survey and matching up the ideas within what the questions are asking. The categories and their assigned questions are organized in Table 4.

Table 4. Cognitive domain categories

Category Name	Survey Questions within Category
Prelab Category	- In the prelab, to experience moments of insight - In the prelab, to be confused about how the instruments and laboratory equipment work - In the prelab, to think about what the molecules are doing - In the prelab, to be confused about the underlying concepts - In the prelab, to think about chemistry I already know - In the prelab, to learn troubleshooting and problem recognition skills
Concept Related Category	- To think about chemistry I already know - To think about what the molecules are doing - To be confused about the underlying concepts - To use my observations to understand the behavior of atoms and molecules
Data Category	- To make decisions about what data to collect - To consider if my data makes any sense - To interpret my data beyond only doing calculations - To be confused about what my data mean
Effort Level Category	- To experience moments of insight - To "get stuck" but keep trying - To make mistakes and try again
Procedure Related Category	- To be confused about how the instruments work - To find the procedures simple to do - To focus on procedures, not concepts
Skills Category	- To learn critical thinking skills - To learn problem solving skills

Table 5 shows the p-values, with the effect size and number of questions in each category. The number of questions in each category was included because it likely influences the p-value found for each category. Hedges' *g* was chosen to measure effect size due to the varying number of questions that were placed in each of the different categories.

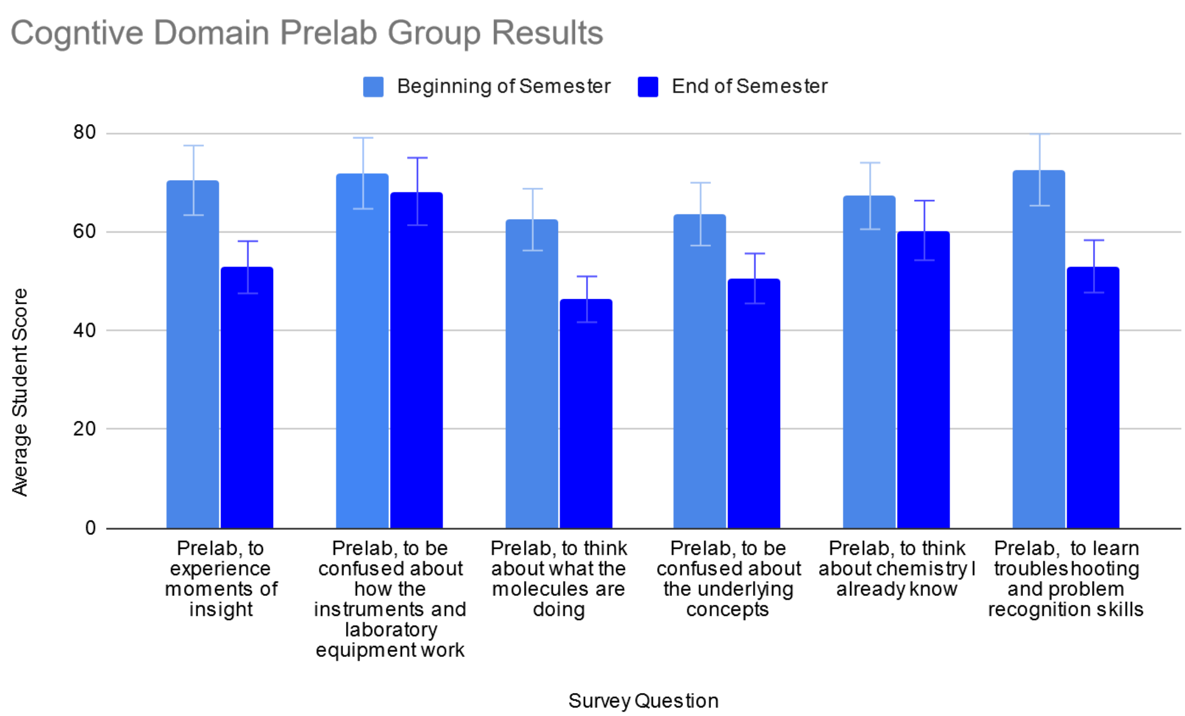
Table 5. T-test results for cognitive domain

Categories Being Compared	Number of Questions within Category	P-Value	Results	Effect Size (Hedges' *g*)
Prelab Category (Pre to Post)	6	0.008	Statistically significant	2.08 Very large effect
Concept Related Category (Pre to Post)	4	0.052	Not statistically significant	N/A
Data Category (Pre to Post)	4	0.405	Not statistically significant	N/A
Effort Level Category (Pre to Post)	3	0.006	Statistically significant	4.30 Very large Effect
Procedure Related Category (Pre to Post)	3	0.384	Not statistically significant	N/A
Skills Category (Pre to Post)	2	0.007	Statistically significant	1.41 Very large effect

Out of all of the groups, the prelab, effort level, and skills group were found to be statistically significant (Table 5). This means there was a real difference in these scores from pre-survey to post-survey and that difference is not attributed solely to chance. The effect sizes for the significant groups were all very large. The prelab group had the largest amount of questions with 6, and the rest of the groups had 4 or less. We therefore attributed at least some of the large effect sizes to the small number of questions in each group.

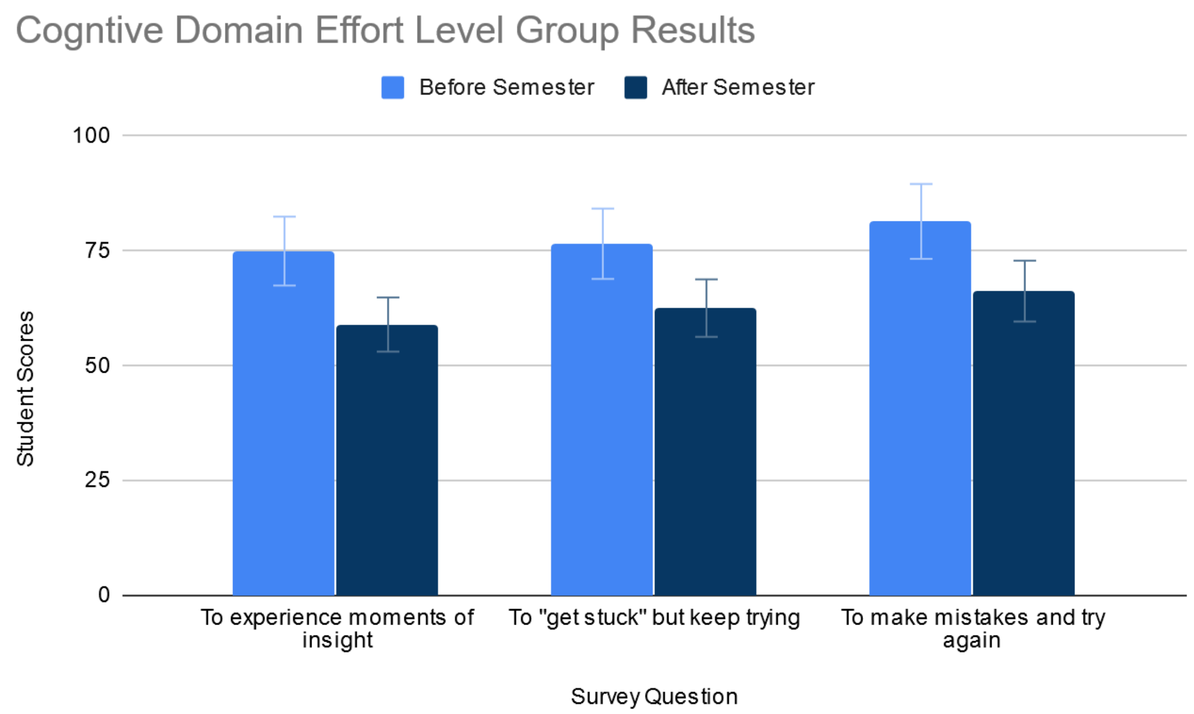
Individual question averages within the prelab group were compared from pre to post (Figure 2). All of the questions were broken down individually. The scores from pre to post decreased for all questions. This means that students' perceptions while working on and thinking about their prelabs at the beginning of the semester, decreased as they progressed in the course. There was the smallest difference between pre and post for the question regarding being confused about using laboratory equipment. This might have been due to the fact that the prelabs given tended to explain how to use the equipment in an effective way, so students understood what was going on and the variance from pre to post was less than other categories.

Figure 2. Cognitive domain prelab category comparison



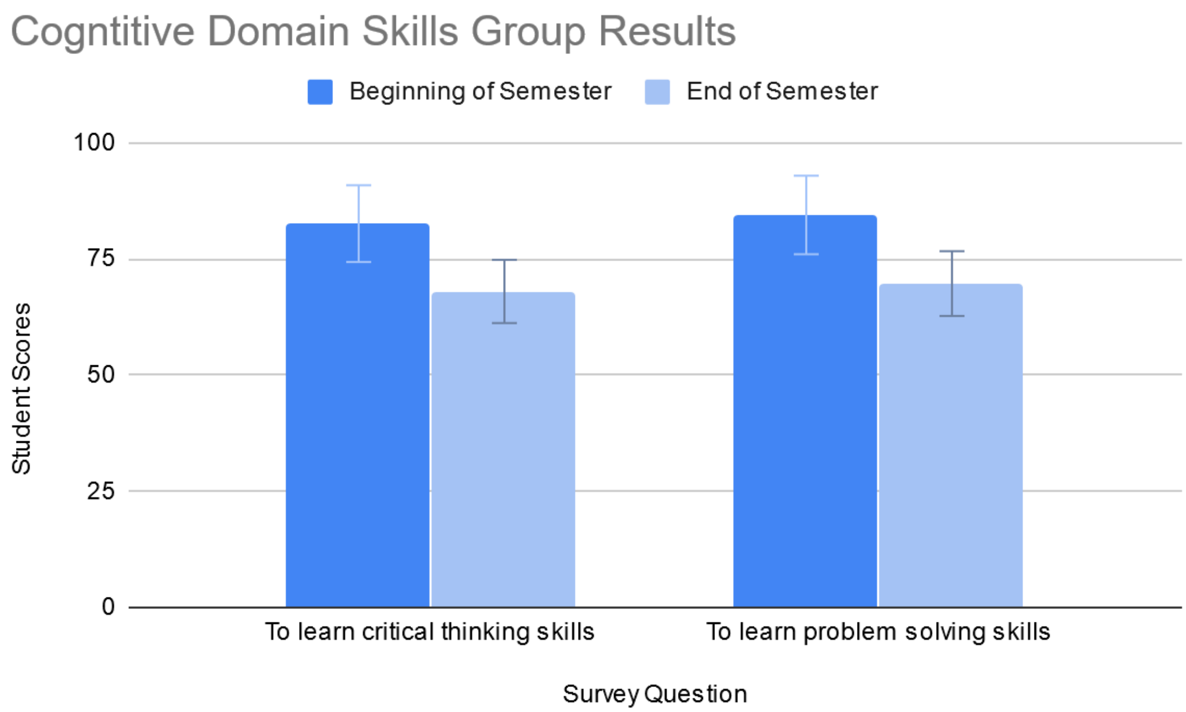
A similar analysis was performed on the individual questions in the effort level category within the cognitive domain (Figure 3). All of the scores started out after the pre-survey in approximately the same range, and the numbers for the post survey were all approximately within the same range. Similar to the prelab group, all of the numbers did decrease from pre to post survey. This means that within this area, students' expectations were not met through this class.

Figure 3. Cognitive domain effort level category comparison



The data below shows the questions within the skill related category in the cognitive domain (Figure 4). There were two questions within this group and their scores for the pre and post survey were very similar. Like Figures 2 and Figure 3 show, their scores also decreased from the pre survey to the post survey. This means students' thoughts and expectations were not met through their lab experience in the skill area.

Figure 4. Cognitive domain skill category comparison



DISCUSSION

Claims

When comparing the pre-survey and post-survey scores, there were differences found in the cognitive domain demonstrating that there was a decrease in pre and post survey scores, while there was found to be no statistical real difference in scores in the cognitive / affective and affective domains. Within the cognitive domain questions, differences were found in students' expectations of the prelab, effort levels, and skill question groups.

Reviewing the Results

All of the scores in each of the three domains decreased their overall scores from the survey at the beginning of the semester to the survey at the end of the semester (Table 2). It can be concluded from these results that meaningful learning did not occur in all domains of meaningful learning. Performing t-tests on the three domains (Table 3) revealed that the cognitive domain was the only domain found to be statistically significant. This means that the pre- and post- changes in the cognitive/affective and affective domains happened purely by chance, but the changes in the cognitive domain were due to the students' laboratory experiences. Because the cognitive domain was the only statistically significant domain of results, the researchers chose to look deeper into the specific questions within it.

All of the categories within the cognitive domain were broken down (Table 4) and t-tests were performed on each. Out of all of the categories, only three were found to be statistically significant (Table 5). The prelab, effort level, and skill categories were all found to have a substantial difference from the survey at the beginning of the semester to the survey at the end of the semester. Each of these categories were graphed to visualize the differences between the two surveys (Figures 2-4). Within these categories, all of the questions demonstrated that scores went down from the beginning to the end of the semester.

Patterns Observed

Students in the general chemistry sections did not demonstrate that their expectations were met or that truly meaningful learning occurred in the lab. This matched the results that were gathered by Galloway and Bretz in their original use of the MLLI survey. This showed that students' expectations at UNI were and currently are going unmet in general chemistry. However, there were categories that increased. Within the cognitive domain, they included:

- This semester when performing experiments in my chemistry lab course I expected to or experienced being confused by the underlying concepts (reverse scored)
- This semester when performing experiments in my chemistry lab course I expected to or experienced being confused by what the data mean (reverse scored)

The categories that increased in the cognitive/affective domain included:

- This semester when performing experiments in my chemistry lab course I expected to or experienced feeling disorganized (reverse scored)
- This semester when performing experiments in my chemistry lab course I expected to or experienced being worried about the quality of my data (reverse scored)

The categories that increased in the affective domain included:

- This semester when performing experiments in my chemistry lab course I expected to or experienced being frustrated (reverse scored)
- This semester when performing experiments in my chemistry lab course I expected to or experienced being nervous about working with chemicals (reverse scored)
- This semester when performing experiments in my chemistry lab course I expected to or experienced being nervous about making mistakes (reverse scored)
- This semester when performing experiments in my chemistry lab course I expected to or experienced being worried about finishing on time (reverse scored)
- Both in-class and out-of-class, while working on anything relating to the prelab I expected to or experienced feeling intimidated (reverse scored)

Interestingly, all of these questions, regardless of category, were reverse scored, meaning that they originated as "negative" questions on the MLLI survey. It was odd that only negative questions increased within the survey. Upon further investigation, we looked at the verbs used in these questions. Although these increases were found in all domains (cognitive, cognitive/affective, and affective), we found that the verbs used in all of these questions tended to be affective in nature and therefore associated with feelings or emotion. For example, "worried", "intimidated" and "confused" were all words used in the questions that illustrated an increased score. These words all trigger deep emotion regardless of their classification. Because of this, we speculate that while taking the survey, students respond differently to these questions since emotions were likely involved. All of these questions increased while the rest decreased. This was attributed to the emotional words used in the question that impacted the students' answers.

Comparing to the Original

When comparing this to Galloway and Bretz's study in 2015 first using the MLLI study, the results were found to be similar. The results showed in both instances that there was the largest difference pre to post survey in the cognitive domain. In the original full study, Galloway and Bretz also found that the scores in all three domains decreased (Galloway & Bretz, 2015). The standard deviations calculated for each category in Galloway and Bretz' paper was between 10 and 20, while this study had them at about 20 to 30. This could be attributed to the overall lower number of participants that completed this survey compared to Galloway and Bretz's survey. Overall, the results that were found in this study were a good reflection of the original study with the majority of the categories and ideas being unfulfilled by chemistry laboratory courses.

The MLLI survey tool created by Galloway and Bretz was designed to measure meaningful learning within the chemistry laboratory. The MLLI survey broke down the aspects of the laboratory by asking questions in the cognitive, affective, and cognitive/affective domains. When looking at the chemistry sections at the University of Northern Iowa, the MLLI was able to discern if student expectations and experiences were net within each of the three domains. The results were averaged, broken down by category, then grouped for similarities and further examined through t-tests and averages. From the results that were collected at the beginning and again at the end of the semester, it was determined that generally students' expectations were not fulfilled through their chemistry laboratory experience.

Reasoning for Decreasing Scores

When looking at overall patterns of the data, all three domain averages went down from the beginning of the semester to the end of the semester. One idea that must be taken into consideration is the age and experience levels of the students in these general chemistry classes. The general chemistry students surveyed were predominantly freshman and sophomores in college who likely do not have much exposure to college classes. Most students probably had some exposure to chemistry in high school, but the format of college labs was found to be quite different. Because of this, students had little exposure to college classes, much less college chemistry. Students could have had higher scores to begin with because they had no real idea of what college chemistry classes would be like. Their expectations for the class could have been skewed higher than expected because of the positive thoughts students might have about the class. When the lab was not what they thought, their reflective survey at the end of the semester could have then led to lower scores. These set of conditions most likely heavily influenced the results.

Not Meeting Expectations: What Does This Mean?

The common theme found throughout the results was that scores from the beginning to the end of the semester decreased. This indicated that students had expectations that were left unmet in their lab experience. However, just because student expectations were not met, and scores did not increase does not mean that the lab portion of the class is currently failing students or that the lab approaches going on have not been effective. From looking at the data, it is possible that some students could have had extremely high expectations coming into the lab and thought that things would be extremely difficult. This might have been the reason for the decrease of scores from the beginning of the semester to the end of the semester. However, because the majority of scores for the questions decreased, it was evident that the class was less confusing and less difficult than students thought. On the other end of the spectrum, students may have assumed that the chemistry lab portion would be easy and low-effort, when that ended up not being the case. With these students, it was possible that their scores stayed the same for the end of the year or decreased even more, leading to overall decreasing scores. Both of these situations could have happened with this survey and does not necessarily mean that the current approaches are not contributing to student learning.

In addition to students having varying expectations before taking this class, the idea of meaningful learning needs to be discussed. The MLLI was set out to measure meaningful learning. The results all showed a decrease, but this does not necessarily mean that there was not meaningful learning that occurred. The researchers never directly looked at lab reports or exam scores from students to see that there were indeed improvements in learning from the beginning to the end of the course. There were questions that did increase in scores, which were all mentioned in the above "Discussion" session. In that regard, learning most likely happened in these areas. However, we need to also consider the other questions and categories that did decrease. These scores decreased, but it does not mean that there was no learning taking place. There could have been great learning taking place, but it was just not the most effective or was not a great experience for the students, leading to the scores decreasing. The scores did decrease, but it does not mean that the current lab approaches are completely failing students.

CONCLUSIONS

Impacts on Teaching

The MLLI survey can be used by teachers to see if the learning and experiences in chemistry lab are being met for students. By breaking down the survey question by question, the MLLI allows instructors to focus on what aspects of their teaching, curriculum, and lab practices allow these insightful moments to occur for students. Collecting and analyzing the results from students about their lab experiences, instructors can make necessary changes in their curriculum or preparation to create a better and more meaningful laboratory experience. By breaking down specifics in this study in general chemistry, it was evident that the curriculum may need to change to be more engaging and impactful on students. For example, within the cognitive domain skills category, students indicated that they expected more problem solving and critical thinking skills to be occurring. With this in mind, a better lab approach can be created to try and improve upon the aspects that were indicated as lacking in the MLLI survey.

Comparing the MLLI to Lab Approaches

From the results gathered, it was clear that student expectations were not being fulfilled. By diving deeper into the individual aspects of the course, these changes can be identified and implemented. The SWH model is unique in that it has students complete an extensive pre lab portion for students to complete before they go into the lab. The prelab category would be best to help draw connections between the SWH model and the MLLI survey. The prelab category was one of the categories that was found to be statistically significant. This signified that the prelab portions of the chemistry labs were effective. However, this contradicts previous research done on the prelabs specifically at the University of Northern Iowa that said that prelabs were not effective due to the cognitive overload that occurs for students (Bonde & Del Carlo, 2018). Both lab formats were included within the survey, so it is possible that the results with prelabs were different because the new format of prelabs with the hybrid model is significant.

The ADI lab model provides students with the opportunity for scientific social interaction. A large emphasis is placed on students talking to their peers about their ideas while exploring what others' thoughts are. Two ways students demonstrated this during the lab was through making a procedure and determining claims as a group based on the data collected. The best way to see if this model were effective would be to look into some of the specific questions in the MLLI survey. There were multiple questions that addressed the procedure and claims portion of the lab within the cognitive domain. The data category and the procedure related categories can both reveal if the parts of the ADI model were effective within the current lab models.

Limitations

This study was done only at one mid-sized university in Iowa. In order to get more generalizable results, there would have to have been more semesters of results compiled together from the University of Northern Iowa to be able to make solid conclusions. In addition, the inclusion of more results from more universities may have been helpful. It can be assumed that professors at different universities would teach chemistry labs in different ways with different approaches. Including a variety of lab formats would help identify the best practices for chemistry labs with the MLLI survey in the future. To do this, the best strategy would be to find professors across the country who teach general chemistry lab approaches and have them provide the study to their students. In addition, as stated previously, we were not able to separate the results from the hybrid section and the normal SWH approach sections. This means that there could have been improvements in components of the lab approach with the hybrid model, but we were unable to directly see these results. If continued, the next step would be to look at the results from those distinct sections and determine if the hybrid model for the lab is more effective through incorporating meaningful learning.

The difference in scores that was found from pre to post may have happened in part due to the number of questions in each domain. The cognitive domain had the largest number of questions within it by far compared to the other two domains. The number of questions could have an effect on the results of the t-test and the p-value that resulted. If the number of questions within the survey was changed to me more even across the domains and categories, the p-values might have changed for some of the results.

Next Steps

Using the results from the survey, we compared the results gathered from the MLLI and were able to compare the lab models, but not to the full extent. Although some comparisons and recommendations were made, there still is more investigating that needs to be done. Due to the error that was made by the researchers and the inability to compare the hybrid SWH/ADI model and the normal SWH model, a good next step would be to determine if the hybrid model provided a better learning experience for students. If there is a difference, the hybrid model can act as a template to build on for the general chemistry curriculum. If there is not a difference, further research can be done using the SWH and ADI curriculum ideas as a base to experiment what curriculum parts work the best for students in the chemistry laboratory.

In addition to looking more deeply at the comparison of the hybrid and traditional models directly, it would be helpful to actually measure if meaningful learning is occurring for future studies. The MLLI told us parts of the lab approach that may need changes to be more effective for students, but the survey was not able to tell us if students are actually learning in meaningful ways within the lab. One way to do this would be through pairing the MLLI survey with measuring meaningful learning by looking at prelabs, lab reports, and exams. Tracking the progress of students to see if they are retaining information and seeing their expectations for the class through the MLLI survey would be a great way to truly capture how effective the chemistry laboratory is in regard to meaningful learning for students in the future.

References

Bowen, C. W. (2000). A Quantitative Literature Review of Cooperative Learning Effects on High School and College Chemistry Achievement. *Journal of Chemical Education, 77*(1), 116–119.

Bretz, S. L. (2019). Evidence for the Importance of Laboratory Courses. *Journal of Chemical Education*, *96*(2), 193–195. https://doi.org/10.1021/acs.jchemed.8b00874

Burke, K. A., Greenbowe, T. J., & Hand, B. M. (2006). Implementing the science writing heuristic in the chemistry laboratory. *Journal of Chemical Education*, *83*(7), 1032–1038.

Clark, R. C., Nguyen, F., & Sweller, J. (2011). *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. Somerset: Wiley.

Cooper, M. (2010). The case for reform of the undergraduate general chemistry curriculum. *Journal of Chemical Education*, *87*(3), 231-232.

Cooper, M., & Klymkowsky, M. (2013). Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *Journal of Chemical Education*, *90*(9), 1116–1122. https://doi.org/10.1021/ed300456y

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning Progressions in Science: An Evidence-Based Approach to Reform. CPRE Research Report # RR-63*. Consortium for Policy Research in Education. https://eric.ed.gov/?id=ED506730

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing Scientific Knowledge in the Classroom. *Educational Researcher*, *23*(7), 5–12. https://doi.org/10.2307/1176933

Galloway, K.R, & Bretz, S.L. (2015). Measuring Meaningful Learning in the Undergraduate General Chemistry and Organic Chemistry Laboratories: A Longitudinal Study

Journal of Chemical Education, *92*(12), 2019-2030. doi:10.1021/acs.jchemed.5b00754

Graham, S. & Hebert, M. (2011), Writing to Read: A Meta-Analysis of the Impact of Writing and Writing Instruction on Reading. *Harvard Educational Review, 81*, 710-744.

Grove, N. P., & Bretz, S. L. (2012). A continuum of learning: from rote memorization to meaningful learning in organic chemistry. *Royal Society of Chemistry*, *13*, 201–208. https://doi.org/10.1039/C1RP90069B

Hofstein, A. & Lunetta, V.N. (2004), The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*, 28-54. https://doi.org/10.1002/sce.10106

Keys, C.W., Hand, B., Prain, V. & Collins, S. (1999), Using the Science Writing Heuristic as a Tool for Learning from Laboratory Investigations in Secondary Science. *Journal of Research in Science Teaching, 36*, 1065-1084. https://doi.org/10.1002/(SICI)1098-2736(199912)36:10<1065::AID-TEA2>3.0.CO;2-I

Llorens-Molina, Juan-Antonio. (2008). Design and Assessment of an Online Pre Lab Model in General Chemistry: A Case Study. *Journal of the Research Center for Educational Technology*, *4*(2), 15-31.

Mayer, R. E. (2009). *Multimedia learning, 2nd ed*. Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mohsen, T., & Dennick, R. (2011). Making Sense of Cronbach's Alpha. *International Journal of Medical Education, 2*, 53-55.

National Research Council. 2000. *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9596.

Novak, J. D. (1998). Learning, Creating, and Using Knowledge: Concept Maps as Facilitative

Tools in Schools and Corporations. *Journal of E-Learning and Knowledge Society, (6)*3, 21-30.

Novak, J.D. (2002), Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education, 86*, 548-571. https://doi.org/10.1002/sce.10032

Paas, F., Tuovineni, J.E., Tabbers, H., & Van Gerven, P.W. (2003). *Educational Psychologist, 38*(1), 63–71.

Sampson, V., Grooms, J., amd Walker, J. P. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education*, *95*(2), 217–257. https://doi.org/10.1002/sce.20421

Sampson, V.,and Walker, J. P. (2012). Argument-Driven Inquiry as a Way to Help Undergraduate Students Write to Learn by Learning to Write in Chemistry. *International Journal of Science Education*, *34*(10), 1443–1485. https://doi.org/10.1080/09500693.2012.667581

Smith, M. E., Hinckley, C. C., & Volk, G. L. (1991). Cooperative learning in the undergraduate laboratory. *Journal of Chemical Education*, *68*(5), 413–415. https://doi.org/10.1021/ed068p413

van Merrienboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the Load Off a Learner's Mind: Instructional Design for Complex Learning. *Educational Psychologist*, *38*(1),513.https://doi.org/10.1207/S15326985EP3801_2

Walker, J. P., and Sampson, V. (2013). Argument-Driven Inquiry: Using the Laboratory To Improve Undergraduates' Science Writing Skills through Meaningful Science Writing, Peer-Review, and Revision. *Journal of Chemical Education*, *90*(10), 1269–1274. https://doi.org/10.1021/ed300656p

Walker, J. P., Sampson, V., & Zimmerman, C. O. (2011). Argument-Driven Inquiry: An Introduction to a New Instructional Model for Use in Undergraduate Chemistry Labs. *Journal of Chemical Education*, *88*(8), 1048–1056. https://doi.org/10.1021/ed100622h

Appendix A: MLLI Pre-Survey

UNIVERSITY OF NORTHERN IOWA

HUMAN PARTICIPANTS REVIEW

Project Title: Development and Evaluation of an ADI/SWH Hybrid Model Laboratory Curriculum

Name of Investigator(s): Kylie Engstrom, Tabitha Alitz, Dawn Del Carlo

You are invited to participate in a chemistry education research project conducted through the University of Northern Iowa. The following information is provided to help you make an informed decision about whether or not to participate in this research project.

Nature and Purpose: The purpose of this study is to look at meaningful learning in the chemistry laboratory when comparing the ADI/SWH hybrid class format compared to the Science Writing Heuristic class format

Explanation of Procedures: This email will have the link to the survey and a consent form. The questions on the survey will ask about your experiences in the lab and what you expect from the class. The results will be shared with the research team only. Taking the survey will not impact your grade or class participation in any way.

Privacy and Confidentiality: Information obtained during this study which could identify you will be kept confidential. The email you enter at the end is simply for the purposes of awarding the gift cards, and your answers will not be attached to that. The summarized findings with no identifying information may be published in an academic journal or presented at a scholarly conference. No guarantees can be made regarding the interception of data transmitted electronically. None of the results on an individual's responses will be shared with anyone outside of the research team. None of the instructors for the course will be informed on who takes the survey or any specific answers on the survey.

Discomforts, Risks, and Costs: Risks to participation are minimal. Questions on the survey focus on expectations and overall feeling about the laboratory class. If the questions make you feel uncomfortable, it is the participants' choice to stop taking the survey.

Benefits and Compensation: No direct benefits to participants are expected, but this research may generate important information for the chemistry education community about the impact of different class formats on meaningful learning for students in chemistry laboratory. At the end of the survey, you will be directed to a separate site where you can enter your email address and be entered to win one of three $50 Amazon gift cards.

Right to Refuse or Withdraw: Your participation is completely voluntary. You are free to withdraw from participation at any time or to choose not to participate at all simply by closing your browser. By doing so, you will not be penalized or lose benefits to which you are otherwise entitled.

If you have questions regarding your participation in this study or about the study generally, please contact Kylie Engstrom at engstkac@uni.edu or the project investigator's faculty advisor, Dawn Del Carlo, in the Department of Chemistry and Biochemistry, University of Northern Iowa, at 319-273-3296 or dawn.delcarlo@uni.edu. For answers to questions about the rights of research participants and the research review process at UNI, you may contact the office of the IRB Administrator at 319-273-6148. If you do not wish to consent to taking this survey, simply exit the browser.

○ If you are interested in continuing, click here to acknowledge that you are fully aware of the nature and extent of your participation in this project as stated above and the possible risks arising from it. You hereby agree to participate in this project. You acknowledge that you have received a copy of this consent statement. You are 18 years of age or older.

Q1: When completing both the in-class and out-of-class work before performing my experiment (i.e. "pre-lab"), I expect…

	Completely Disagree (0%) … Completely Agree (100%) 0 10 20 30 40 50 60 70 80 90 100
to feel unsure about the purpose of the activities.	
to experience moments of insight.	
to be confused about how the instruments and laboratory equipment work.	
to be excited to do chemistry.	
to think about what the molecules are doing.	
to feel disorganized.	
to feel confident about planning and executing my own procedure.	
to be confused about the underlying concepts.	
to think about chemistry I already know.	
to be frustrated.	
to be intrigued by the instruments and laboratory equipment.	
to feel intimidated.	
to learn troubleshooting and problem recognition skills.	

Q2 When performing experiments in my chemistry lab course this semester, I expect…

	Completely Disagree (0%) … Completely Disagree (100%) 0 10 20 30 40 50 60 70 80 90 100
to learn chemistry that will be useful in my life.	
to worry about finishing on time.	
to make decisions about what data to collect.	
to feel unsure about the purpose of procedures.	
to experience moments of insight.	
to be confused about how the instruments work.	
to learn critical thinking skills.	
to be excited to do chemistry.	
to be nervous about making mistakes.	
to consider if my data makes any sense.	
to think about what the molecules are doing.	
to feel disorganized.	
to develop confidence in the laboratory.	
to worry about getting good data.	
the procedures are simple to do.	
to be confused about the underlying concepts.	
to "get stuck" but keep trying.	
to be nervous when handling chemicals.	
to think about chemistry I already know.	
to worry about the quality of my data.	
to be frustrated.	
to interpret my data beyond only doing calculations.	
We use this statement to discard the survey of people who are not reading the questions. Please select forty percent for this question.	
to focus on procedures, not concepts.	
to use my observations to understand the behavior of atoms and molecules.	
to make mistakes and try again.	
to be intrigued by the instruments.	
to feel intimidated.	
to be confused about what my data mean.	
to be confident when using equipment.	
to learn problem solving skills.	