Writing-to-Learn Physics: Helping Students Understand Energy Systems

by

Robert Dalka

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Bachelors of Science with Honors
(Physics)
in The University of Michigan
2019

Research Advisor:

Professor Timothy A. McKay

ACKNOWLEDGEMENTS

I would like to thank Professor Tim McKay for his guidance and encouragement throughout the duration of this project. I would not have had the opportunity to do this without his support. I am extremely and forever grateful for the exposure to Physics Education Research that he has provided for me.

Thank you to our research and instructional team that makes Physics 140 and the research we conduct in it possible. I would especially like to thank Nita Kedharnath, Dr. Tom Finzell and Juniar Lucien for the helpful discussions and notes. Thank you to the entire Foundational Course Initiative team and members at CRLT for their help in research and curriculum development. I would also like to thank Dr. Yuri Popov for working with me to make the implementation of the activities a possibility.

I am thankful for the M-Write team and all the work that goes into creating the program. Thank you to Professor Larissa Sano for her help in creating the assignments and all her work during the semester. A huge shoutout goes to the Writing Fellows. These activities could not have happened without them and working alongside each one of them has been an amazing experience.

Finally, I want to thank my family for always being supportive and helping me succeed during my undergraduate degree. Thank you to my mother for her endless support and being the strongest person I know. A final thank you to Maddie Barnes. You were there at every step and I can always count on you to talk me through any obstacle and help me find a solution.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

ABSTRACT

Writing-to-Learn Physics: Helping Students Understand Energy Systems

by

Robert Dalka

Advisor: Timothy A. McKay

In the Introduction to Mechanics course at the University of Michigan, writing-to-learn (WTL) activities have been implemented to engage students in writing about physics concepts related to real world scenarios. The instructional tool of WTL has been widely explored in various educational contexts with varying success, but rarely in a physics setting. However, there is significant evidence showing that WTL can facilitate student learning in a number of physics topics. Thus, we have brought three WTL activities into our Introduction to Mechanics course, Physics 140, that serves 600+ students each semester. In this thesis, I have investigated the success of our WTL activities through analyzing students' progression towards one of the learning goals associated with an activity. The learning goal that is the subject of this study is, "Understanding the importance of defining an energy system and how to choose the entities to include in it". Using both qualitative and quantitative analysis, we have found that clear progression was made by a significant number of students. This is supported by evidence from both the Fall 2018 and Winter 2019 terms. In both semesters, the activity also gave insight into student knowledge which was previously unavailable. The WTL activities that were used in the Introduction to Mechanics

course this past academic year have shown that they are useful for both students and instructors. These activities have the potential to be helpful for stimulating learning in a wide variety of physics topics. They should continue to be developed and the study of their implementation will be necessary.

CHAPTER I

Introduction

Writing-to-learn (WTL) activities have made their way into STEM courses at every level of college and secondary education *Gere et al.* (2019). These activities often feature students engaging deeply with writing about scientific concepts. As a part of a wider campus initiative called M-Write, our Introduction to Mechanics course, Physics 140 (P140), brought WTL activities into the work that we have students do.

These activities were first used in the Spring 2018 term as a part of a pilot in which the course was ran in a studio-style format. These activities then were adapted from the pilot course and made their way into all three sections in both the Fall 2018 and Winter 2018 terms of P140. This work was done in coordination by the Foundational Course Initiative (FCI). The FCI seeks to transform the courses that serve large populations of students into courses that allow for the opportunity of success for all students while maintaining their rigor. This is done through evidence-based course design to ensure that all changes to courses are improving the courses and create more equitable and inclusive environments. The work done on the WTL activities in P140 fits into this larger scale course transformation that is taking place. The work presented here focuses on our WTL activities effectiveness at helping students reach one of our learning goals, "Understanding the importance of defining an energy system and how to choose the entities to include in it".

The implementation of the WTL activities at this large scale is a large undertaking. In these type of course changes, there are multiple groups of people involved and a lot of time and effort is invested. Our WTL activities reached 600+ students in both the Fall 2018 and Winter 2019 semesters. Students spent a more than 7,200 hours total each semester engaged in WTL, creating a total of 18,000 words over both the First and Revised Drafts. Each semester, more than 200 hours of office hours were held specifically to help students with the WTL activities. 14 Writing Fellows graded and gave feedback on each of these drafts, spending upwards of 1,000 hours per term reading and responding to students submissions. These numbers showcase the scale of these activities and motivate the need to research and understand their impacts.

To investigate how well we were able to engage students in their learning process with WTL, we looked at our activities as serving two purposes. One, how well are we able to evaluate students knowledge and what information can we gain that is not present in other forms? And two, how well did the WTL activity facilitate students progression towards the learning goal?

1.1 Outline

This thesis begins by giving context to the course in which these WTL activities were conducted. I then present a literature review of both writing-to-learn and the current body of Physics Education Research (PER) in student's conceptual understanding of energy within systems. The methods used in this research are then explained, including the specific form that our WTL activities took along with the tools used in the analysis of student responses. Following the methods, the development process of the specific WTL activity and the prompt that is given to students is discussed. After giving the context of the prompt, the analysis section then explores the student submissions and how our students progressed towards the learning goal.

Finally, we discuss the implications of our findings and give recommendations for future implementations of WTL activities and the research projects that can pair with them.

1.2 Physics 140: Introduction to Mechanics

Each semester, over 600 students enroll in P140. This course is a pre-requisite for many courses on campus and has become known as a Foundational Course. While investigating WTL in this course, it is important to keep both the students who take the course and the expectations of the course in mind. Both of these factors impact how the students engage in the WTL activities.

1.2.1 Course Demographics

According to the FCI course report, over the academic years 2013 - 2017, women are underrepresented in our course by about 20% when compared to the student body as a whole. The population of other traditionally underrepresented groups matches with the University demographics.

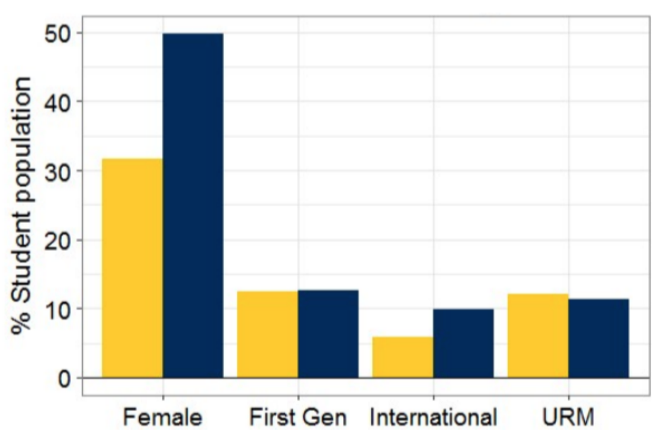

Figure 1.1: The overall course demographics, compared to U-M as a whole over the academic years 2013 - 2017. In maize is the percent student population of P140. In blue is the percent of student population for the entire University. *Figure created by Heather Rypkema, FCI course report.*

Over this same time period, about 75% of our students have come from the College of Engineering (CoE), with the second largest group of students enrolling through the College of Language, Science and Arts (LSA). The majority of students coming from both CoE and LSA identified as male.

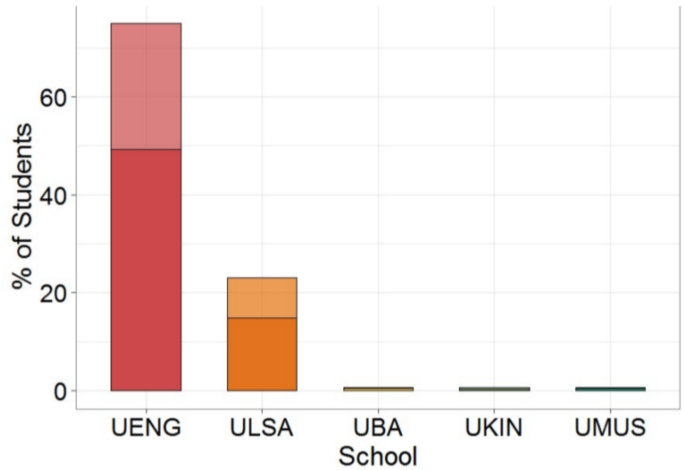

Figure 1.2: School/College affiliation of students enrolled in P140 over the academic years 2013 - 2017. The dark shade for each population category represents the percent of students who identify as male, while the lighter shade represents those students who identify as female. *Figure created by Heather Rypkema, FCI course report.*

P140 predominantly serves first year and sophomore students, with a slightly higher population of second year students (see figure below). This first and second year label is tied to the number of credits a student has. It is a possibility that some students labeled as sophomores are simply coming to college with credits from high school.

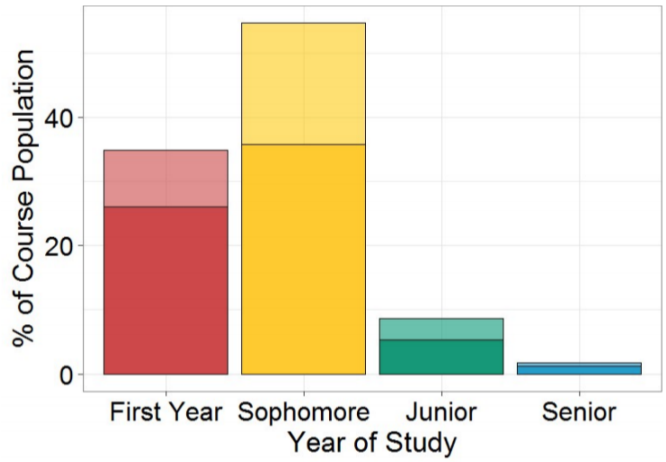

Figure 1.3: Year of study for students enrolled in P140 during the academic years 2013 - 2017. The dark shade for each population category represents the percent of students who identify as male, while the lighter shade represents those students who identify as female. *Figure created by Heather Rypkema, FCI course report.*

The racial and ethnic diversity within P140 is consistent with the diversity profile of the university as a whole within 1%. Students who identify as white make up the majority of the students who take our course.

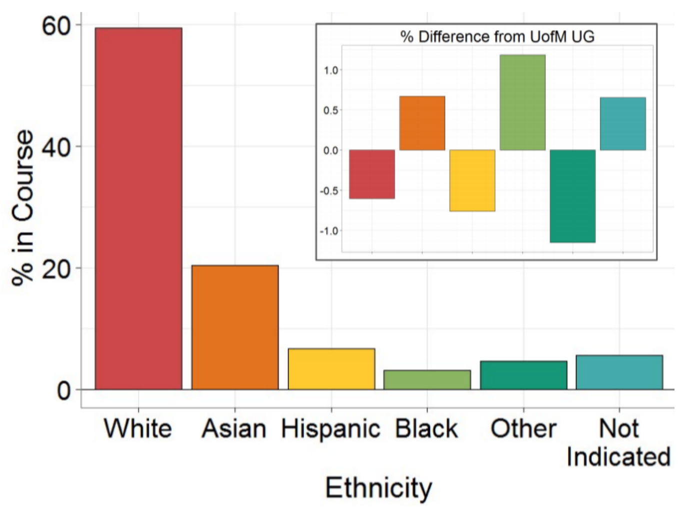

Figure 1.4: Ethnicity distribution of students enrolled in P140 during the academic years 2013 - 2017. *Figure created by Heather Rypkema, FCI course report.*

It is important to keep our students identities in mind when implementing WTL activities. A diverse course, such as ours, must be aware of everyone's background and ensure that all students have the tools that they need to succeed.

1.2.2 Course Expectations

Traditionally, this course has adhered to the following grade rubric; 8% Class participation (through i-clicker responses), 4% Prelectures and checkpoint questions (through Flip It Physics), 16% Homework (conducted through Mastering Physics), 20% labs, 36% Three midterm exams (at 12% each), and 16% Final exam.

Within these grade guidelines, about half (52%) of the points are attributed to the exams. These exams are multiple choice tests in which we do not look through the students work and we do not see the students thought process. The lab grades are assessed separately in a Lab Course that is not directly under the control of P140. The next most points was awarded through homework problems assigned through Pearsons Mastering Physics web-homework system. This system gives students the opportunity to learn through doing problems with step-by-step methods. These are not incorrect modes of learning and evaluating physics, however each is missing a crucial part of what science is. Through the introduction of WTL activities into the course, the instructional team sought to bring a new type of evaluation to compliment exams in the form of written explanations and a new learning tool to work alongside the traditional homework problems.

With these new WTL activities, the course grading guidelines shifted slightly. The course for the Fall 2018 and Winter 2019 terms adhered to the following grade rubric; 5% Class participation (through i-clicker responses), 5% Prelectures and checkpoint questions (through Flip It Physics), 14% Homework (conducted through Mastering Physics), 5% Three WTL activities (M-Write), 20% labs, 36% Three midterm exams (at 12% each), and 15% Final exam. The percentage of the course grade determined by exams only shifted down by 1% and our WTL activities were given 5% of the total grade. While there were not large shifts in the large scale grading scheme, the instructional team felt that 5% was significant enough to motivate students to participate in the assignments.

1.3 My Role

I joined the Physics 140 research group in the Spring 2018 term. I had previously worked as a Learning Assistant for the course since Fall 2016. As a part of joining the research team, I was tasked with coordinating the implementation of the WTL activities into the Spring term course. Along with my research position, I worked as a Learning Assistant during class time and was hired as a Writing Fellow to grade and give feedback on our WTL activities.

Going into the Spring term, no concrete prompts or rubrics for the activities had been created. I took on the creation of these prompts and rubrics with help and feedback from various members of our research group and the Foundational Course Initiative team. The prompts for our activities were created as the Spring term was in session. The structure of these prompts was based on examples from other course that had previously used M-Write, such as Chemistry and Economics courses. The topics and learning goals were set by the instructional team. I brainstormed scenarios based around these topics and would present them to the team. Once a scenario was agreed upon, I drafted up prompts based on the scenarios and these received edits from the team. The final drafts of each prompt were completed the week that they were sent out. As with many other aspects of the Spring course, it served as the testing space for our prompts.

As I was the sole grader during the Spring term, I was able to design a grading rubric that worked well for me and the students. However, this grading rubric did not go into enough detail to be used in a course of over 600 students and multiple graders. Therefore, the time in between the Spring term and Fall term courses was time spent on revising both the prompts and rubrics for the WTL activities. Based on our experience with WTL in Spring, each prompt was either completely rewritten with a different scenario, or edited to increase the engagement that students had with the prompt. Once again, I led these changes and received input from the instructional team on spots that needed to be clarified for students. To generate a rubric that can be used by many different graders constantly is a challenge. I did not have prior experience in this sort of work. To create the rubric that we used in the Fall term, I worked with a member of the Foundational Course Initiative. They were able to provide advice and guide me in the best practices of a writing assignment rubric. Our final Fall rubrics were modeled off of the rubrics used in the Economics course that used M-Write and is a part of the Foundational Course Initiative.

Normally, in the courses that have participated in M-Write, it is the instructor that is responsible for the generation of prompts, creation of rubrics, and coordination of the team of Writing Fellows. As instructors of our Physics 140 course are often busy with the many other aspects of the course, I took on these responsibilities. My work during the Spring term prepared me well for the first two tasks, however, the coordination of a team of 14 Writing Fellows was one that was new to me. Writing Fellows are hired by the M-Write system and are trained in a 1 credit seminar course that is ran during the semester. In our case, many of the Writing Fellows came from the Spring Term course. This was helpful as they were already familiar with WTL activities in Physics 140. The Fellows who were seeing WTL activities for the first time benefited from the others experience.

During the Fall and Winter terms, I worked closely with both the instructional team and Writing Fellows in my role as the Lead Writing Fellow. I worked with the instructional team to create the assignments in Canvas and set up the Peer Review tool on our course site. We planned the schedule of when the activities would be conducted in the course timeline and worked around other aspects of the course to do so. I worked with Writing Fellows in preparing them for each prompt by discussing the prompt with them as a group. I also created breakdowns of each prompt to highlight the learning goals and how we expected students to show their progress in these goals. This was to help the Writing Fellows in their grading and feedback to students.

Another aspect of my role as Lead Writing Fellow was to schedule office hours and grading sessions during the Fall and Winter semesters. I coordinated the scheduling of Office Hours with the Physics Student Services Center so that we could host our WTL specific office hours in the Physics Help Room. I worked with the Writing Fellows to fill these times. After each First Draft was due, I also scheduled grading sessions in which all the Writing Fellows would come together to start their grading. These were typically a few hours in which I would give an example of how to grade the specific prompt, followed by the Fellows beginning their grading. This gave them the opportunity to communicate with each other and myself about the grading and ask any questions they may have.

Over this past year, I have been deeply involved in the implementation of our WTL activities. I was there at their conception and have brought them to their current form. Thus, the work needed to research these activities and evaluate their impact on student learning naturally fell to me. In this thesis, I will present how effective our activities were through the lens of the specific learning goal, "Understanding the importance of defining an energy system and how to choose the entities to include in it".

1.4 Literature Review

In this Literature Review, I present the current body of work surrounding both writing-to-learn and students understanding of Energy. These two areas of research are imperative for this thesis as I am directly investigating how well a WTL activity can support student learning around a specific learning goal related to Energy. In each section, I present an overview of the current findings and explore how different studies compare to each other and the conclusions we can take from them. I end with a description of how the work in this thesis fits into the current literature and the motivations of performing such a study.

1.4.1 Writing-to-Learn

Writing-to-learn (WTL) activities are based in cognitive theory. Janet Emig theorized that writing uniquely serves learning because it naturally follows the structure of recollection, synthesization, and revision. She believed that through this model, students were able to come to a more nuanced understanding of the material (*Emig* (1977)). When Janet Emig wrote "Writing as a Mode of Learning" in 1977, WTL theory had not been widely empirically explored. Since then, WTL has been implemented in a wide variety of contexts and has been analyzed in a number of large scale meta summaries (*Ackerman* (1993), *Rivard* (1994), *Bangert-Drowns et al.* (2004), *Prain and Hand* (2016), *Gere et al.* (2019)). Through the years, the way that WTL activities have been implemented, their effectiveness and the critiques of them have changed.

Most recently, in "Writing and Conceptual Learning in Science: An Analysis of Assignments" (*Gere et al.* (2019)), the current landscape of WTL in the sciences is evaluated. In their review of WTL activities, the authors evaluated each activity by identifying four major components and examining how much these impacted students' learning. These four components were *meaning-making writing tasks* (including analysis, evaluation, argument, and explanation), *interactive writing processes* (including discussion with a peer or instructor, receiving feedback from an instructor or peer), *clear writing expectations* (explanation of what students are to learn, genre, rhetorical context, and criteria to be used in grading) and *metacognition* (planning, monitoring, evaluating, and adapting cognitive strategies).

One of the important conclusions that the authors reach is that the way the components of assignments are implemented has a significant effect on the quality of learning. Simply the presence of these components does not directly lead to meaningful student learning. The authors found that the activities were most effective when students are exposed to multiple opportunities to learn what is expected and engage them with multiple forms of each of the components. A conclusion that can be drawn from this is that it is ineffective to only implement a single WTL activity within a course and expect to see large learning gains. Students must become familiar with the practice and have opportunities of reflection.

An important piece of the WTL activity that has been identified by Gere et al. is the Peer Review aspect of activities. There is evidence to support that student understanding improves more when Peer Reviews are given as a form of feedback, rather than just the instructor alone (*Patchan et al.* (2009), *Cho and MacArthur* (2010)). Specific to a physics context, it has been shown that students are able to evaluate their peers' work to the same degree as experts, (*Price et al.* (2013)). In a study of student misconceptions in introductory biology, Halim et al. showed that the prevalent mode of revision arose through the directed peer-review comments. However, for just as many misconceptions were remedied, new ones would be introduced into students work (*Halim et al.* (2018)). While Peer Review has been established as an important part in satisfying the *interactive writing process* component, mixed results have been seen in its implementation.

The review by Gere et al. encompasses a wide range of WTL activities that were conducted in a variety of contexts spanning multiple scientific disciplines and level of study. However, out of the 46 total studies that were synthesized in the report, only 9 were identified with physics as the primary discipline with only 6 studies conducted in a university level environment. It is important to recognize the limited scope of each individual study as WTL can be better suited for certain content areas than others. In the body of Physics Education Research (PER) literature, there is support for WTL activities, but no clear consensus of what topics WTL activities can best service. While it is clear that the type of WTL activity and how it is implemented has an impact on student learning, it has been shown that certain topics see larger learning gains than others while using the same format and conducted in the same environment (*Cummings and Murphy* (2007)).

In the study by Cummings and Murphy, performance on a subset of questions from the Force and Motion Conceptual Evaluation and the Energy and Momentum Conceptual Survey was used post-instruction to measure the conceptual understanding of both Newton's 3rd Law and Impulse and Momentum. The researchers showed that there was no difference in performance on Newton's 3rd Law questions between students who participated in WTL activities with those who did not. However, the students who had experience with the WTL activities did show a higher performance on Impulse and Momentum questions than the control group of students.

It has also been shown that WTL activities can be useful in students learning of electrostatics. In a study conducted in an electrostatics unit, the WTL activities that were given to a treatment group of students in resulted in higher learning gains on the Electrostatics Conceptual Test when compared to students in the control group (*Atasoy* (2013)). These WTL activities took place in class with peer discussion prior to writing and instructor feedback provided post writing.

While WTL theory has been around for many years, and the evaluation of it's successes has been reflected on throughout this time (*Rivard* (1994)), the body of literature on WTL in physics is still thin. The logistical implementation of the activities is important, however, there does seem to be a correlation with the specific topic being addressed by the activity (*Cummings and Murphy* (2007), *Atasoy* (2013)). It is within this space that this thesis fits. Thus, the goal of this thesis is to not only evaluate the WTL process through the development of a prompt and the components of the activity, but also to address how well the practice of WTL can influence student's progression towards a specific learning goal.

1.4.2 PER: Energy Systems

The learning goal that we are investigating with our WTL activity is associated with the concept of energy. Energy can be a difficult concept to explain, and often instructors fall into using metaphors and words that could potentially cause confusion for students (*McIldowie* (1995)). A recent study found that novices explain energy in a wide range of ways, including "the ability to do work" as well as treating energy as a substance (*French et al.* (2015)). These are not necessarily incorrect ways of interpreting energy and are often used in intro level textbooks to explain different phenomena (*Young et al.* (1996), *Knight* (2017), *Chabay and Sherwood* (2015)), but they have the potential to cause confusion in certain situations.

The ways of talking about and categorizing energy, ontologies, have been studied previously. A study found three predominate ways students discussed the concept of energy; energy as a substance, stimulus to action, and as a vertical location (*Scherr et al.* (2011)). The researchers found that students who thought of energy as a stimulus that has an effect on objects often conflated energy and forces. They found that treating energy as a substance best supported the idea of energy conservation as students used this ontology to track energy during an interaction. However, they noted that using a substance metaphor could seed the idea that Potential Energy is held in objects if used too broadly. This fear was shown to be true where researchers have seen an increase in referring to Potential Energy as being located inside an object (*Goodhew and Robertson* (2014), *Lindsey et al.* (2012)).

Goodhew and Robertson identified six advantages of the substance metaphor; energy is conserved, energy transfers among objects, energy is localized, energy can be located in objects, energy can change form, energy can accumulate in objects. As well as four limitations of the metaphor; energy does not share all qualities of substances, potential energy is not located in a single object, energy is frame-dependent, energy can be negative. The researchers found through their study that when the substance metaphor was explicitly included in the activity, both the advantages and limitations increased within students. The limitation that increased the most was students who stated that potential energy was held within one object. However, Goodhew and Robertson noted that a greater number of instances occurred in which the advantages increased and thus concluded that the advantages out weigh the limitations.

The use of the substance metaphor is accepted as a useful tool in teaching students energy. It is the most heavily used tool in energy tracking diagrams (*Gray and Scherr* (2016), *Scherr et al.* (2012)) in which students create graphical and pictorial representations of energy during interactions. Using the substance metaphor to create these diagrams is positively correlated with successful energy tracking and conservation. However, identifying the system in which the energy is being tracked is often lacking from the representations.

Without relating to students the importance of identifying a system, they may struggle with certain scenarios, including those involving Gravitational Potential Energy. Students have been found to struggle with understanding the implications of a chosen system (*Sabo et al.* (2016)), especially as it has to do with GPE (*Lindsey et al.* (2012)). Lindsey found students tend to associate potential energy with a single object, consistent with the other studies. They also found that students assumed that energy of any system was constant and that they failed to recognize that any group of objects could be considered for a system.

The current literature surrounding Energy strongly favors using the substance metaphor when teaching students how Energy is transferred during interactions. However, as is pointed out in many of the studies, this idea often leads to students believing that Potential Energy is held within objects. As well as students being unable to show understanding of the importance of selecting an energy system and the implications it has on using the Conservation of Energy. The work presented in this thesis falls into this gap in the current body of literature nicely. Methods for helping students understand how to define energy systems and why it is important are needed. This work does not seek to replace the current support for using the substance metaphor, but instead used to compliment the current teaching methods that are used. WTL activities offer a possible method to help students learn about energy systems.

CHAPTER II

Methods for Collecting and Analyzing Student Understanding

In this chapter, I introduce the various methods of implementation and analysis that were used in the process of this thesis. First, I explain how the M-Write process works. I cover all three components of the activities, including their various motivations, how they are designed to help students learn, and the grading rubric used for each. I then present the specific prompt that was used for our second WTL activity, which is the focus of this thesis. Next, I discuss the Natural Language processing method that was used to analyze the Spring term data set. I present the tools and statistical tests used.

I then move to introducing the techniques that were used to investigate students' progression towards our learning goal in the Fall term. I explain how we collected and analyzed our data through quantitative and qualitative methods. Finally, I conclude with discussing the post-exam surveys that were used to measure students' attitudes towards the activities.

2.1 M-Write

M-Write is a socio-technical program created by a team in the Sweetland Writing Center at the University of Michigan that is designed to bring writing-to-learn assignments into STEM courses. The program is funded by U-M through a five-year, $1.8 million grant. Various courses across the university have implemented the program with great success, including Chemistry, Materials Science, and Statistics. The goal of M-Write assignments is for students to learn more as they think deeply about the topics they write about.

M-Write Fellows are a part of this program. They are undergraduate students who have done well in the course previously. Fellows offer technical and conceptual support for students through office hours and email. Each fellow reads and grades the submissions of 42 students along with providing written feedback of both the first and revised draft. Fellows also contribute to research projects conducted in our course.

In the Fall 2018 term, the 650+ students completed 3 different WTL activities. Each activity consists of three parts; the First Draft, the Peer Reviews and the Revised Draft. The goal of this process is for students to learn and reflect on their knowledge.

The students had approximately 2 weeks for the entirety of each activity. They would complete the First Draft during the first week and the Peer Reviews followed by the Revised Draft during the second week. Submissions were required to be within 300 to 500 words and Peer Reviews were expected to be thoughtful and substantive.

Each of the 3 different activities had to do with a topic covered in Physics 140 and the instructions for each activity were given through prompts. Prompt 1 asked students to explain a fleas flight trajectory using Kinematics. Prompt 2 guided students to apply their knowledge of Energy and Systems to a real-world energy storage facility. Finally, in Prompt 3, students were asked to use their knowledge of Forces and Conservation Laws to explain an elliptical orbit.

2.1.1 First Draft

The first step in the M-Write process is the First Draft. In P140, we saw this as an opportunity for students to attempt to answer the prompt to the best of their abilities. The philosophy of WTL is that students learn through the whole process, and so we wanted students to be able to complete the First Draft without worrying about whether or not they were completely correct. In this First Draft, students were expected to address a list of "be sure to's..." that are related to specific learning goals. The students were not expected to answer each one with complete accuracy.

During both the Fall 2018 and Winter 2019 terms, students were given one full week to complete their First Draft. The prompt would be released on a Thursday and the following Thursday the First Draft would be due on Canvas by 11:59 pm. During the week leading up to the due date, Office Hours were held by Writing Fellows in the Physics Help Room, 1416 Randall Lab.

Students were given access to the rubric that they would be graded on. These can be viewed in figures 2.1 and 2.2.

FIRST DRAFT RUBRIC

		Strong	Sufficient	Weak	Insufficient	Points awarded
Physics Concepts		*7 points* **Attempts** to address **all** parts of the prompt using relevant physics concepts	*5 points* **Attempts** to address **most** parts of the prompt using relevant physics concepts	*3 points* **Attempts** to address **few** parts of the prompt using relevant physics concepts	*0 points* **Attempts** to address **no** parts of the prompt using relevant physics concepts	
		Strong	Weak	Insufficient		
Style of Writing	Perspective	*3 points* **Appropriate** voice for the role of the writer and target audeince	*2 points* Voice is **inconsistent** with the role of the writer or target audience	*0 points* Voice is **inappropriate** for the role of the writer and target audience		
	Clarity	*3 points* Writing is **easy to follow** for a non-specialist and is cohesive and connected	*2 points* Writing is **inconsistant**, possible to be followed by someone trained in the field but not a non-speicalist	*0 points* Writing is **confusing** and difficult to understand		
	Structure	*2 points* 300 - 500 words and follows the correct format	*1 points* 300 - 500 words but does not follow the correct format	*0 points* Outside of word limits		
					Total	

Figure 2.1: The rubric for the First Draft for the Fall 2018 term. This rubric was used both by students to keep track of their own writing and for Writing Fellows to grade student submissions. The total amount of points awarded in the First Draft was 15 points. Out of the total 100 for an entire M-Write activity, this is worth 15%.

We learned through grading during the Fall 2018 term, that we had too high of a resolution in our grading scale. There was not a clear difference between taking a single point off or taking two points off. Therefore, in the Winter 2019 term, we moved to a grading scale that had less resolution, but each point carried more weight. Along with changing the resolution, we revised the descriptions of each grading category to further help both students and Writing Fellows understand the rubric.

First Draft Rubric (6 points → 15%)

Physics Concepts (max. 4 pts)	4 Points	3 Points	2 Points
	Attempts to address **all** of the "Be sure to's…"	*Attempts* to address **most** of the "Be sure to's…"	*Attempts* to address **few** of the "Be sure to's…"
Audience Consideration (max. 1 pt)	1 Point	0 Points	
	• Writing is **clear**, so someone who is not an expert **would understand** the physics. • Response **takes into account** audience's physics background.	• Writing is **confusing,** so someone who is not an expert would have **difficulty understanding** the physics. • Response **does not take into account** audience's physics background.	
Conciseness (max. 1 pt)	1 Point	0 Points	
	Response is **within** 300 to 500 words while **staying relative** to the prompt.	Response is **outside of** 300 to 500 words and **strays away** from the prompt.	

Figure 2.2: The rubric for the First Draft for the Winter 2019 term. This rubric was used both by students to keep track of their own writing and for Writing Fellows to grade student submissions. The total amount of points awarded in the First Draft was 6 points. Out of the total 40 for an entire M-Write activity, this is worth 15%.

2.1.2 Peer Reviews

The next step in the M-Write process is the Peer Reviews. The process of conducting Peer Reviews has been shown to be heavily impactful on student learning (*Halim et al.* (2018), *Moon et al.* (2018)).

The M-Write team has developed a Canvas plug-in that automates the distribution of student work and collects the feedback that students give each other. A typical student would be assigned three other First Drafts to read and review. Within the Peer Review tool, specific questions were asked that related to the learning goals of the prompt. Guided by these questions, students were able to provide more relevant and helpful feedback. Students were usually given the weekend after the First Draft was due to complete their Peer Reviews. They were typically due by 9:00 am on the Monday following the First Draft due date.

The Peer Reviews were also graded. The grading of the Peer Reviews was to motivate students to provide thoughtful feedback that actively helped their peers.

PEER REVIEW RUBRIC

	Strong	Sufficient	Weak	Insufficient	Points awarded
Relevance	*7 points* Answers each rubric question with relative and helpful comments for **all** responses	*6 points* Answers each rubric question with relative and helpful comments for **most** responses	*5 points* Answers each rubric question with relative and helpful comments for **about half of** responses	*0 points* Answers each rubric question with relative and helpful comments for **few to none** responses	
Specificity	*7 points* Answers each rubric question with specific examples from the writing for **all** responses	*6 points* Answers each rubric question with specific examples from the writing for **most** responses	*5 points* Answers each rubric question with specific examples from the writing for **about half of** responses	*0 points* Answers each rubric question with specific examples from the writing for **few to none** responses	
Structure	*6 points* Meets word count for each rubric question in **all** responses	*5 points* Meets word count for each rubric question in **most** responses	*4 points* Meets word count for each rubric question in **about half of** responses	*0 points* Meets word count for each rubric question in **few to none** responses	
				Total	

Figure 2.3: The rubric for the Peer Review for the Fall 2018 term. This rubric was used both by students to keep track of their own writing and for Writing Fellows to grade student submissions. The total amount of points awarded in the Peer Review was 20 points. Out of the total 100 for an entire M-Write activity, this is worth 20%.

As with the First Draft rubric, this rubric was also changed for the Winter 2019 term. Again, this was done to give more importance to each point and further clarify expectations for students. The categories for the Peer Review rubric were changed more dramatically than for the other two components of the activities. The "Relevance" and "Specificity" categories were combined into one category, while the "Structure" category became the "Thoroughness" category.

Peer Review Rubric (8 Points → 20%)

Content of Review (max. 5 pts)	5 Points	4 Points	3 Points
	All responses are... ● relevant to the question. ● helpful and respectful. ● specific and use examples from text.	**Most** responses are... ● relevant to the question. ● helpful and respectful. ● specific and use examples from text.	**Few** responses are... ● relevant to the question. ● helpful and respectful. ● specific and use examples from text.
Thoroughness (max. 3 pts)	3 Point	2 Points	1 Points
	All responses to each question are longer than 30 words.	**Most** responses to each question are longer than 30 words.	**Few** responses to each question are longer than 30 words.

Figure 2.4: The rubric for the Peer Review for the Winter 2019 term. This rubric was used both by students to keep track of their own writing and for Writing Fellows to grade student submissions. The total amount of points awarded in the Peer Review was 8 points. Out of the total 40 for an entire M-Write activity, this is worth 20%.

2.1.3 Revised Draft

The final component of the M-Write process is the Revised Draft. This is where students are able to take the feedback from both their peers and from their Writing Fellow and revise their original submission. The goal is for students to change what they originally wrote and better address the learning goals that we set. Through this revision, they are also revising their knowledge to better align with the correct physics concepts. Even if a student had correct explanations in their First Draft, they were encouraged to find ways to refine their explanations to bolster their understanding of the physics.

Students were typically given about three days to make these revisions after receiving their feedback on Monday morning. The Revised Draft was due on Thursday night at 11:59 pm. Again, students were given a rubric that they could use to guide their revisions and make sure they were meeting the grading expectations.

REVISED DRAFT RUBRIC

		Strong	Sufficient	Weak	Insufficient	Points awarded
Physics Concepts		*40 points* **Accurately** addresses **all** parts of the prompt using relevant physics concepts	*32 points* **Accurately** addresses **most** parts of the prompt using relevant physics concepts	*24 points* **Accurately** addresses **few** parts of the prompt using relevant physics concepts	*0 points* **Accurately** addresses **no** parts of the prompt using relevant physics concepts	
		Strong	Weak	Insufficient		
Style of Writing	Perspective	*7 points* **Appropriate** voice for the role of the writer and target audeince	*5 points* Voice is **inconsistent** with the role of the writer or target audience	*0 points* Voice is **inappropriate** for the role of the writer and target audience		
	Clarity	*7 points* Writing is **easy to follow** for a non-specialist and is cohesive and connected	*5 points* Writing is **inconsistant,** possible to be followed by someone trained in the feild but not a non-speicalist	*0 points* Writing is **confusing** and difficult to understand		
	Structure	*6 points* 300 - 500 words and follows the correct format	*4 points* 300 - 500 words but does not follow the correct format	*0 points* Outside of word limits		
Revision		*5 points* **Meaningful** revision is made in the physics or writing that shows thought and reflection	*3 points* Revisions made are **unthoughtful** and do not boost the quality of the response	*0 points* **Insufficient** revisions made for a final draft		
					Total	

Figure 2.5: The rubric for the Revised Draft for the Fall 2018 term. This rubric was used both by students to keep track of their own writing and for Writing Fellows to grade student submissions. The total amount of points awarded in the Revised Draft was 65 points. Out of the total 100 for an entire M-Write activity, this is worth 65%.

Similarly to the First Draft and Peer Review rubrics, the Revised Draft rubric was changed to decrease the sheer number of points while retaining the same relative weight. The Revised Draft rubric in the Fall term had too fine of a resolution and was the main reason for transforming all of the rubrics. The rubric used in the Winter term can be seen in Figure 2.6. This rubric is more clear in its expectations.

As an additional source of guidance, the rubric was broken down into a grading guide. The grading guide for the Fall 2018 term can be found in Appendix A. The guide was used to give more specific instructions on how to award points to students. Overall, the rubrics used in our WTL activities were designed to help both students during their writing and Writing Fellows during their grading.

Revised Draft Rubric (26 Points → 65%)

Physics Concepts (max. 15 pts)	15 Points	12 Points	9 Points
	Accurately addresses **all** of the "Be sure to's…"	*Accurately* addresses **most** of the "Be sure to's…"	*Accurately* addresses **few** of the "Be sure to's…"
Audience Consideration (max. 5 pts)	5 Points	3 Points	0 Points
	• Writing is **clear,** so someone who is not an expert **would understand** the physics. • Response **takes into account** audience's physics background. • **Minimal** equations are used to explain physics concepts.	• Writing is **sometimes confusing,** so someone who is not an expert would have **difficulty understanding** the physics. • Response **attempts to take into account** audience's physics background. • Equations are **sometimes** used to explain physics concepts.	• Writing is **confusing,** so someone who is not an expert **would not understand** the physics. • Response **does not take into account** audience's physics background. • Equations are used to explain physics concepts a **majority** of the time.
Conciseness (max. 1 pt)	1 Point	0 Points	
	Response is within 300 to 500 words while **staying relative** to the prompt.	Response is outside of 300 to 500 words and **strays away** from the prompt.	
Revision (max. 5 pts)	5 Points	3 Points	0 Points
	• **Clear** revisions have been made that **actively improve** on the First Draft. • Corrections are made to physics concepts (if applicable). • Correct explanations **are improved upon** to be more clear and precise.	• **Some** revisions have been made that **do not always improve** on the First Draft. • Corrections are made to physics concepts (if applicable). • The correct explanations **are not improved** and do not become more clear or precise.	• **No** revisions have been made.

Figure 2.6: The rubric for the Revised Draft for the Winter 2019 term. This rubric was used both by students to keep track of their own writing and for Writing Fellows to grade student submissions. The total amount of points awarded in the Revised Draft was 26 points. Out of the total 40 for an entire M-Write activity, this is worth 65%.

As we created the WTL activities used in our course, it was important to keep the four major components identified by Gere in mind. These included meaning-making writing tasks, interactive writing processes, clear writing expectations, and metacognition (*Gere et al.* (2019)). We sought to meet each one of these criteria in our activities. Our use of real world examples to connect to physics concepts and placing the student in a specific role writing to an audience aimed to create meaning-making writing tasks. The inclusion of Peer Review and Writing Fellow feedback created an interactive writing process. Our rubrics were created to make expectations clear for students. Finally, the requirement of a Revised Draft with thoughtful revision was included in our activities to have students engage in metacognition.

2.2 The Prompt

In this section, I outline how the prompt used in the second WTL activity came to its current form. I discuss the original conception of the prompt and its evolution along the learning goals. The title of our second prompt is “A Watershed Moment in Energy Storage”. The focus of this prompt is for students to think about defining energy systems and tracking the energy that flows into and out of the system that they define. This prompt was implemented before the second exam in our course time line and is released immediately following the lectures on Energy.

In this second prompt, the student is placed in the role of a consultant working at a renewable energy firm. In the prompt, they receive a frantic phone call from their boss. Their boss is interested in the Ludington Pumped Storage Plant, which uses pumped water that is held above a bluff on the coast of Lake Michigan to store energy. Their boss tasks the student with writing a memo describing the physics of the storage facility.

The students must define the system that the energy is flowing into and out of and describe the energy transfers that occur while the water is moving through the facility. They also were asked to explain how energy can be lost while being consistent with the Conservation of Energy.

Figure 2.7: Aerial image of the Ludington Pumped Storage Facility. The light blue "lake-type" body of water is the stored water that has been pumped up from the large lake below (on the far right of the image).

2.2.1 Spring Term Context

This prompt was created during the Spring 2018 term of Physics 140. The Spring 2018 term was a time of change. Not only was it the first time that M-Write would be used in a physics course, but it was also the pilot version of a studio-style version of Physics 140. The focus of our studio-style course was mastery focused learning and the majority of time in class was dedicated towards students working through problems.

A class day began with a short lecture, typically around 15 minutes before students began their group work. The group work consisted of four, or five, higher level physics problems that had been broken up into guided steps. After the group work for the day, students then worked on their individual work that was made up of conceptual questions, a mastery question related to the group work questions, and a challenge problem. Since the majority of the time spent in class was centered around the group work, we were able to gain a considerable amount of buy-in from the students as they felt a greater connection to the class and each other. It was under these condition that the prompts were first developed.

2.2.2 Spring Term Prompt and Reflection

This particular prompt, “A Watershed Moment in Energy Storage”, was created with the intention of featuring a local example of Gravitational Potential Energy (GPE) storage. Upon a lengthy brainstorming session, the example of the Ludington Pumped Storage Facility was settled on.

The next step was to clearly state the learning goals that we wanted students to work towards through the M-Write process. We decided on the following two learning goals:

1. How Gravitational Potential Energy related to the separation distance.

2. How to recognize energy transfers and the ability to track the energy.

We then needed to set the students in a realistic scenario where they would need to explain the conceptual physics of the Pumped Storage Facility. This is when we decided to place the students in the role of someone working at a energy consulting firm and that they would be writing to their boss about the facility. We felt that this position would be a realistic position for most of the students in our course. The the full prompt used in the Spring 2018 term can be seen in Appendix B.

The deployment of this activity went smoothly. Students’ First Drafts where mostly reasonable responses with small suggestions being made by both myself (the M-Write Fellow), and their peers during Peer Reviews. A common comment made was along the lines of “loss of energy”. The following are feedback comments that were commonly made in response to students’ writing about loss of energy.

> **Comment 1:** "In your response, think about where energy might be converted to forms that are not easily recovered. What would these forms be and when would the energy be converted to these forms?"
>
> **Comment 2:** "Be careful about how you talk about the transformation of energy into forms that are not easily recovered. Using the word "lost" can be misleading as energy can never be destroyed. Where are they being "lost" from?"

In their Revised Drafts, many students adjusted the way they were talking about the "loss" of energy and were more tactful in their treatment of losing energy. However, this lead us to think about adjusting our learning goals of the prompt. We noticed that students were having trouble with discussing energy loss and what it meant for there to be an energy system. Therefore, we structured our new learning goals around this issue.

2.2.3 Fall Term and Beyond

We knew that we had to change our learning goals to better match what students were struggling with. If our learning goals were set along the concepts that students needed the most help with, our activities would be more helpful for students and their learning. The following became our three new learning goals:

1. Understanding the importance of defining an energy system and how to choose the entities to include in it.

2. How to recognize energy transfers and the ability to track the energy.

3. How the Conservation of Energy can be consistent with a "loss" of energy.

Learning goal 1 was used to replace the first learning goal from Spring. To talk about GPE, you need the two entities in your system that are being separated, such as the Earth and the Water. Otherwise, you must instead discuss the Work done by the Gravitational Force, which students rarely did. This new learning goal also functioned as a base to build learning goals 2 and 3 upon. Learning goal 2 was a hold over from Spring as we felt that it was still an important part of the prompt. Learning goal 3 was introduced to directly address the issues we were seeing students having during the Spring term.

Along with our re-defined learning goals, we expanded on the prompt to give the prompt a more realistic feeling. We gave the student's boss a personality by writing out a phone call in which the boss would lay out their questions and help contextualize the situation. Without the group work in class creating buy-in for students during the Fall and Winter terms we were worried we may lose some of the investment from students. We felt that the more detailed prompt would help create some of the buy-in we would be missing out on due to the lecture-style class. The prompt that was used in Fall and Winter can be seen in Appendix 2.

The development of this WTL activity and the prompt that is used for it was not just a single day of brainstorming and writing. It took testing our original ideas in the Spring term and refining them for the Fall and Winter semesters. The work is never done either. There will always be something about the activities that can be improved. Work should be continuously put into improving all assignments and tools we provide to students.

2.3 Linguistic Inquiry and Word Count (LIWC)

The Linguistic Inquiry and Word Count program, from now on referred to as LIWC, is an algorithm that turns any piece of writing into a set of variables that aim to describe the text that it is given. It does this through a "bag of words" approach, in which it takes each word and matches it to an internal dictionary of categorized words. From this it builds the variables that describe text. The internal LIWC dictionary has a total of 6,400 words, word stems, and a select number of emoticons (*Pennebaker et al.* (2015)). For each text file, the algorithm creates about 90 variables that describe the text. This includes 4 language variables that describe the type of language used (analytical thinking, lout, authenticity, and emotional tone), 3 general descriptor categories (words per sentence, percent of words matched, and percent of words with six or more letters), 21 grammar dimensions, 41 categories about psychological constructs (cognition, drives, etc.), 6 personal concern categories (work, home, etc.), 5 informal language markers, and 12 punctuation categories.

LIWC also has the ability to use a user created dictionary in place of its' own to generate variables describing the text. In an exploratory effort, a "Physics" dictionary was created. This dictionary was used along with the LIWC dictionary to investigate the Spring term work. This dictionary included categories such as Energy, Kinematics, Motion, Interactions, ect.

Once variables had been created for each students' First and Revised drafts, a Multivariate Analysis of Variance (MANOVA) test was ran. The goal of this analysis is to find what variables set two, or more, groups of submissions apart. This can be done multiple times with multiple splits among the submissions. After choosing the type of groups to organize the submissions into, the MANOVA test is used to find the largest variance between the groups. The test takes into account multiple continuous dependent variables, and bundles them together into a weighted linear combination or composite variable.

After the largest contributing composite variables are found, a Primary Component Analysis (PCA) test was ran. The PCA plots all of the submissions on a multi-dimensional space where the dimensions are the variables that were determined to be significant by the MANOVA test. Then, the variables that contribute most to the separation of groups are identified as the primary components. You can then take 2-dimensional slices and plot all individual submissions along the axes of primary components. If there are differences between the groups of submissions that were identified, these differences would be seen in the groupings of submissions along the plot.

Using this process, one is able to identify what characteristics of the writing are most different between groups of submissions. The hope of this type of analysis is to gain insight into what might separate students' understanding of different topics and what types of groups could be identified.

2.4 Qualitative analysis of writing supported by automated methods

The majority of the analysis done in this thesis focuses on how students made progress towards a specific learning goal. To study this, all of the students submissions needed to be categorized into groups depending on their specific responses that related to the learning goal in question. To efficiently categorize the submissions, a Python script was created to pull the significant passages from the submissions. This cut down the amount of reading needed significantly and made the task of reading every students response a manageable task.

For the learning goal that was the focus of this thesis, the significant passages where those that had to do with energy systems. Codes were created as a quick overview of the passages was taken. These codes were used to categorize students' definitions of the system for both their First and Revised Drafts. The distribution of these categories was used to understand the range of student understanding in our course. The distributions for the First and Revised drafts were combined into heatmaps that show the movement of students from one system definition to another. Different definition migrations were investigated for different student populations.

To compliment the data from the First and Revised Drafts, the reasons for student's revisions were also explored. This focused on the Peer Review process and the feedback that students both gave and received. This analysis adds more context to the revision process and how students progress towards the particular learning goal that was under investigation. For this analysis, a similar technique was employed to pull the Peer Reviews that gave certain common feedback notes. The percentage of students who gave certain types of feedback was addressed and this was related to how students revised their energy system definition.

2.5 Qualitative Student Examples

The quantitative data gives an overall picture of what happened during these activities. From the large statistics with our high number of participants, we can notice large scale trends and the patterns we do notice have more weight behind them. However, without actually looking at what student wrote individually, the collective statistics may not be enough to truly understand what happened in these activities. Therefore, a qualitative investigation of individual students has been added to this thesis to give the additional context needed.

For these case studies, the full progression of single students is covered. For each case study, we looked at how the student addressed the specific learning goal we are measuring. The reason for the categorization of their definition is explained. Any additional information about the student's First Draft is also explained.

Then, we begin looking at the Peer Reviews that they received as well as the other First Drafts that they were assigned to read during the Peer Reviews. The extent to which the feedback that the student received helped them in their Revised Draft is assessed. The same is done for the First Drafts that the student read. In this section, the feedback from the student's Writing Fellow is included and the possible impact of this feedback is assessed as well.

Finally, the relevant passage in the student's Revised Draft is analyzed and commented on. Through this process, we are able to contextualize the overall trends by highlighting specific examples. Through this type of analysis, we are able to understand what the larger trends are actually telling us about students' progression towards our learning goal.

2.6 Post-exam Surveys

In the Fall 2018 term of Physics 140, extra time was given on exams. The exams that were created used an estimated time of examination of 90 minutes. In Fall 2018, the time given was 1.5 times longer, resulting in exams that had time restraints of 135 minutes. This extra time gave the instructional team an opportunity to collect answers to survey questions by adding them as questions to the end of the exam as questions 21 - 25.

In the post-exam surveys for Exam 2 and Exam 3, a question about the M-Write process was included. The question in the Exam 2 survey asked students about their opinion of the M-Write process as a whole and whether or not they felt it had an impact on their learning. While answers to this type of question may be skewed one way or another, the opinion of students greatly impacts the buy in from them and the effectiveness of the WTL process.

In the Exam 3 survey, a question was included asking students about their participation in the M-Write Office Hours. In this questions, we wanted to evaluate the support we were giving to students for these new type of assignments.

The distribution of answers to each question was looked at and was further broken down into groups of students. Students were then grouped by which response they gave to the specific survey question, and their grade on the corresponding exam was investigated. This was done to check if the different opinions that students had about M-Write or the ways they interacted with office hours had any correlation to their exam performance.

CHAPTER III

Results of Student Learning Progression

In this chapter, I present the results of our various investigations into our WTL activity. I begin with the Natural Language Processing analysis that was performed using the LIWC program and various tests of variance. I show how the analysis of the Spring Term data set came back with null results, so we abandoned this mode of analysis for the Fall 2018 data set. Next, I present the grading results of our second WTL activity. I discuss how each Writing Fellow's grading compared to the average grading, thus providing evidence for the validity of our process as an assessment.

Then, I move into the bulk of our results in looking directly at how students progress towards the learning goal. I look at this progress through the First Draft, Peer Reviews, and the Revised Drafts. Student trends are identified and results are presented on the final distribution of how students defined their system and reasons for doing so. As an extension of this quantitative analysis, I present three examples of students moving through the WTL activity as a whole, tracking their learning progression.

I report on the survey results from post-exam surveys conducted during the Fall 2018 term. These provide context for what students thought about the activities and how they engaged in M-Write specific office hours. I then conclude our results with an analysis of the data that we collected during the Winter 2019 term. This includes grade data along with a learning goal analysis.

3.1 Spring Term LIWC

In the Spring Term, we used three total WTL activities. While only the second activity's scenario remained in the Fall and Winter terms, the concepts covered by the first and third were similar. The student submissions that were collected during the Spring term served as our initial data set. To analyze this data, we tested out a Natural Language Processing technique that was described earlier in the Methods chapter. For this analysis, the program LIWC was used along with MANOVA and PCA tests. The results from this preliminary study were not promising, and therefore this technique was not used to investigate the WTL activities during the Fall term. However, it is important to include these results in this thesis to highlight our reasons for not continuing the technique.

As described in the Methods chapter, once we had collected all of the student submissions, the files were processed by the LIWC software and given variable measurements based on the LIWC internal software. We then decided to take a look at different ways to divide the submissions into groups to find possible differences between them. The groupings that were investigated were First and Revised Drafts, division by gender, and division by performance split at the mean grade. In our analysis, we were not able to find any statistically significant variance between the overall First and Revised Drafts nor the divisions by performance and gender. We did not have a large enough data set from the Spring term to look into other possible groupings.

Each time that new a new grouping method was investigated, the MANOVA and PCA tests were ran to check for the differences between these groups. This means that different principle components were found for each grouping method.

Our first attempt at using this method was to find differences in the First and

Revised Drafts. In this early attempt, all of the submissions were used to see if there was any overall trend in how our students revised. The first principle component identified was characterized by the usage of words having to do with describing time and motion. This means that when students used a lot of words describing time or motion, their submission would be plotted high along this dimension. The second principle component that was found to separate the groups was when the submission contained a large number of words associated with cognitive processes. This includes words such as "cause", "would", and "should". The first two principle components were found to be responsible for 34.2% of the explained variance between the First and Revised Drafts. Even though the first two principle components are the strongest in terms of explaining the differences between the two groups, they do not overwhelm the remaining principle components.

One would typically expect to see distinct groups when plotting the individual submissions along the first two principle components. However, due to the fact that the two principle components did not heavily outweigh the others, we end up without any statistical difference between the mean of the two different groups. Using LIWC, we were unable to detect any large scale differences between the First and Revised Drafts.

Figure 3.1: Plot of each submission for all First and Revised Drafts along first two dimensions. The blue circles represent the submissions that are First Drafts. The yellow triangles represent the submissions that are the Revised Drafts. The ellipses around each mean point represent the error on the mean.

We figured that it was probably a little ambitious to expect that there would be wide-scale differences in First and Revised Drafts when comparing across all three activities. Therefore, we took a look at the variance between the First and Revised Draft for each individual activity. Similarly to the initial investigation, the first WTL activity showed no difference between the two drafts when plotting along it's first two principle components. When looking at only the second WTL activity, the only significant variable between the First and Revised Draft groups was the pure word count. When we investigated the third WTL activity, a significant difference in the mean position of First and Revised Drafts was found when plotted along the first two principle components. A simple explanation for why this difference showed up in this activity is that most students did not understand the initial set up of the prompt given to them. When they revised, some started from scratch and most had to make a large amount of revision.

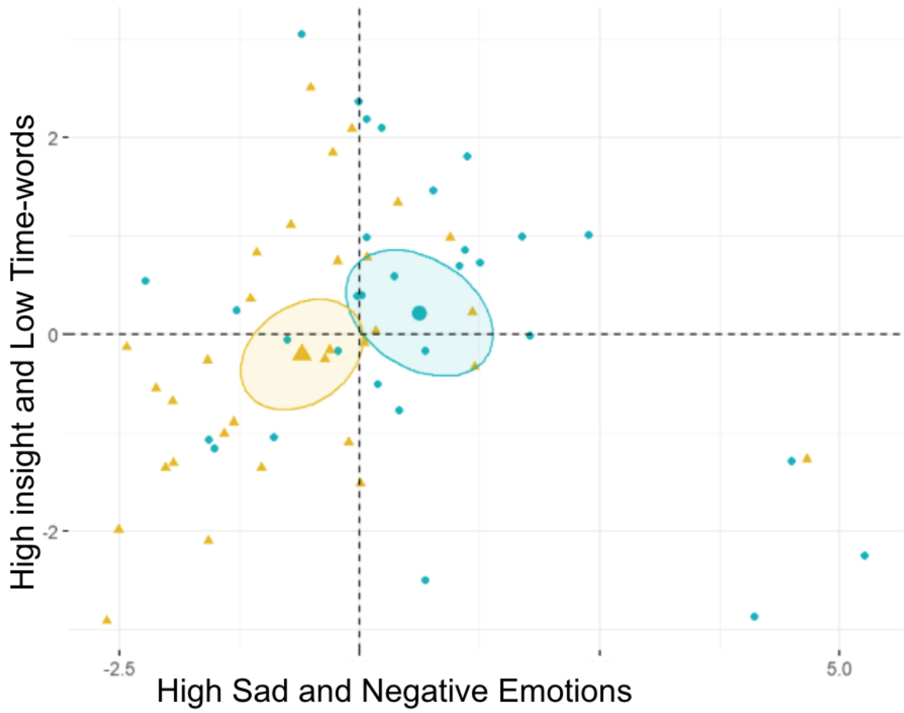

Figure 3.2: Plot of each submission for third prompt First and Revised Drafts along first two dimensions. The blue circles represent the submissions that are First Drafts. The yellow triangles represent the submissions that are the Revised Drafts. The ellipses around each mean point represent the error on the mean.

Even though we were able to see this difference, the two greatest principle components started to give us some doubt as to whether or not the internal LIWC dictionary would be able to be used for our purposes. This first principle component in this analysis was characterized by a high number of words associated with negative emotions. Words such as "loss" are included in this pile of words. When talking about physics, the word "loss" may be used in different contexts than in every day life. The negative emotion categorization would not necessarily be appropriate for a physics context.

A useful aspect of LIWC is the ability to create and upload new dictionaries. For Physics, we currently have a first draft of a dictionary that we applied to the Spring data. This dictionary was built through looking at both words that students used in their writing and words the University Physics book (*Young et al.* (1996)) included in the definitions section at the end of chapters. When using the "Physics" Dictionary, a strong difference between each of the prompts can be seen.

Figure 3.3: Plot of each Revised Draft compared between WTL activities using the created Physics dictionary. The blue circles represent the first WTL activity (Kinematics). The yellow triangles represent the second WTL activity (Energy). The red squares represent the third WTL activity (Orbits).

This is outcome is not surprising for two reasons. First, each of the activities concerns a different physics concept. Second, we created the dictionary using the writing, so these results simply confirm that we correctly categorized the data set. Although it was gratifying to finally spot a true difference, the technique was not carried on to the Fall data as it did not offer any specific insight to students' writing on the whole.

3.2 Grading the Activities

It is important to assess how any activity or tool is used in a course, especially when it is being implemented for the first time. A part of this assessment is investigating how grades are given to students and if the process is fair to all and that the grading is not wildly skewed in any one way. That is why in this section, I present the grading data of our second WTL activity. I specifically show our investigation into how grades compare across different Writing Fellows and how well this matches up to the average scores. This analysis is done for all three components of the activity.

3.2.1 First Draft

A total of 621 students turned in their submission on time for the First Draft. The average grade of these students was 97.1%. The high average score is expected as the First Draft grading rubric only takes into account if they attempted to address everything asked of them, not if they correctly addressed everything. The distribution of these scores can bee seen in Figure 3.4.

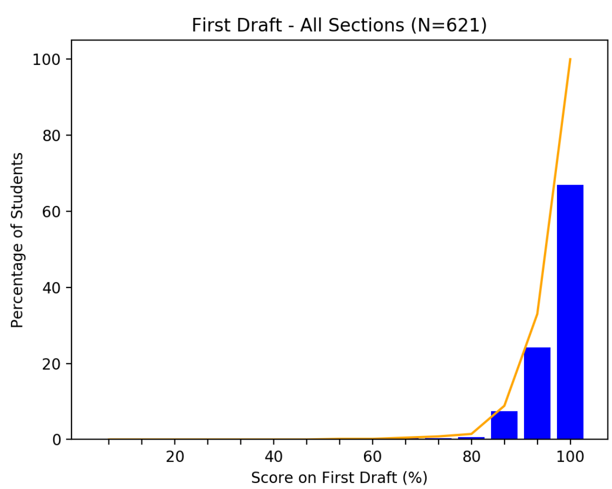

Figure 3.4: Distribution of scores on the First Draft for all students. The blue bars represent the percentage of students with that score. The orange line shows the integrated number of students along the distribution.

These scores were given by the 14 Writing Fellows. One possible challenge that we faced was to ensure that all Writing Fellows graded consistently. To better understand how the Writing Fellows were grading, individual grading distribution was also found for each Writing Fellow. The distributions of the Writing Fellows grades resembled the overall distribution. While there were some outliers in which 100% scores were much more common than others, the overall average scores were not worrying. All average Writing Fellows scores were within one standard deviation of the overall average score. These average scores are compared to the overall average in Figure 3.5.

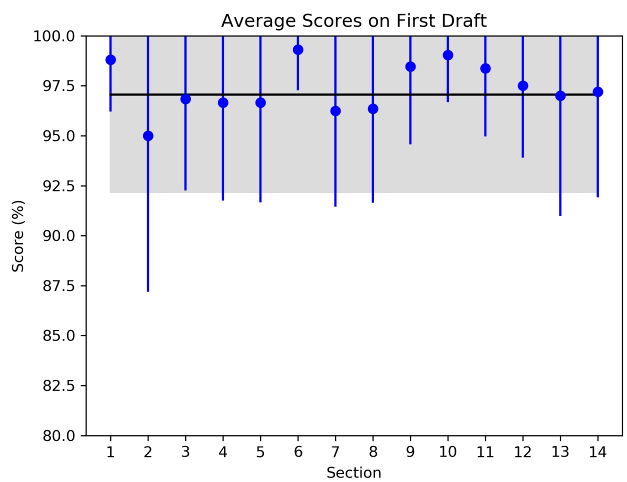

Figure 3.5: Comparison of Writing Fellow's First Draft scores to the overall average score with the error bars being standard deviations from the average. The blue dots show each Writing Fellow's average. The black line denotes the overall average score with the gray region showing the standard deviation of the overall average.

3.2.2 Peer Review

Not every student who completed a First Draft completed the Peer Reviews that were assigned to them. In fact, only 573 students completed the Peer Reviews. The average score of students who turned in the Peer Reviews was 96.5%. The distribution curve for the Peer Reviews had a longer tail than that of the First Drafts. This is due to some students not completing all of their reviews and is responsible for the lower average score. The distribution of these scores can be seen in Figure 3.6.

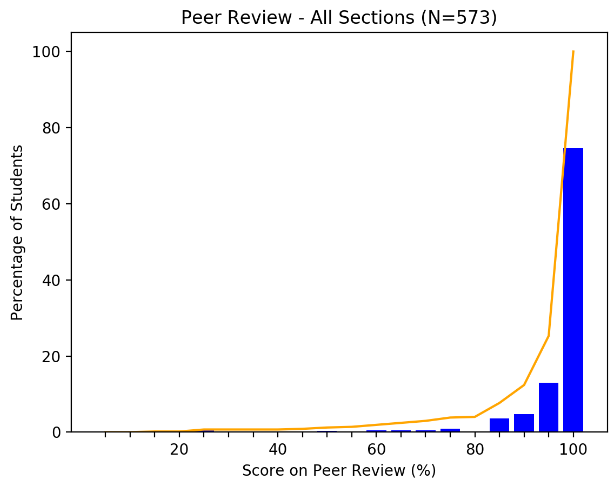

Figure 3.6: Distribution of scores on the Peer Review for all students. The blue bars represent the percentage of students with that score. The orange line shows the integrated number of students along the distribution.

As was with the First Draft, all of the Writing Fellows gave average scores that were within one standard deviation of the overall average.

3.2.3 Revised Draft

At 627, the number of students who submitted their Revised Draft was higher than both the number of students who completed their Peer Reviews and those who submitted a First Draft. The average overall score for the students who submitted the Revised Draft was 94.0%. This average score is lower than the First Draft, however, this is not necessarily unexpected as the requirements of the Revised Draft are more strict. The distribution of scores also turned out to be slightly larger than previous distributions, with less than 40% of students earning full points. The complete distribution of the Revised draft scores can be seen in Figure 3.7.

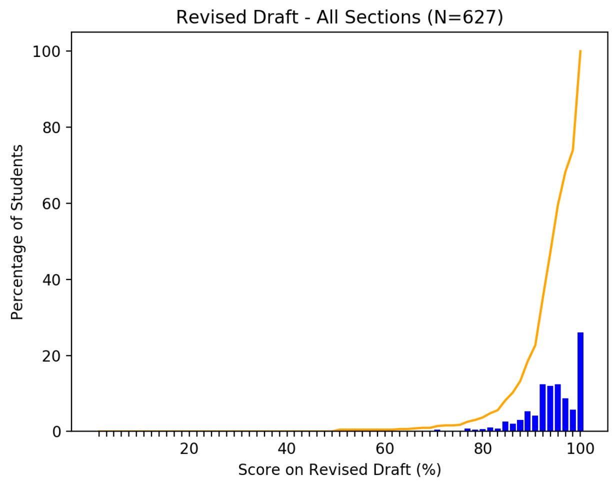

Figure 3.7: Distribution of scores on the Revised Draft for all students. The blue bars represent the percentage of students with that score. The orange line shows the integrated number of students along the distribution.

Once again, all Writing Fellows gave average scores that were within one standard deviation of the overall average score. However, the way in which Writing Fellows scored varied slightly from Fellow to Fellow. Some Writing Fellows, such as those for Sections 2 and 11, had a very spread out score distribution. However, other Writing Fellows, such as those for Sections 6, 12 and 13, feature very large percentages of students earning the same score. This pattern could have been due to the ultra-high resolution of our grading rubric and gave us another reason to create a less fine grain scale in the Winter semester.

The largest differences in the grading occured during the Revised Draft. Overall, grading went smoothly with Writing Fellows staying consistent in their average scores.

3.3 Learning Goal Analysis

In this section, I discuss how well students met our learning goal of "Understanding the importance of defining an energy system and how to choose the entities to include in it". I look at how students defined their system in their First Draft, along with how students gave their written feedback during the Peer Review section. I then present how students defined their energy system in their Revised Draft and how it is different than the First Draft. Finally, I present the overall trends found in students' progression to our learning goal. Reasons for the trends observed are investigated and discussed.

3.3.1 First Draft

In the First Draft, we found that students widely varied when it came time to decide what entities to include in their system. In order to quantitatively analyze the students definitions of systems, categories needed to be made to code student's writing. These categories were created by reading student's submissions.

> **Earth and Water:** "The system we are working with is in reference to the mass of the water and the Earth as a whole unit." – Student 227, First Draft

> **Earth, Water and Plant:** "We can define the system that the energy is flowing into and out of as the Earth, the water, the reservoir, and the turbines." – Student 375, First Draft

> **Water:** "The system that the energy flows in and out of is the lake water itself." – Student 155, Revised Draft

> **Plant:** "When they referred to an energy system, they were talking about how their system converts energy to help their needs, in this case electricity." – Student 391, First Draft

> **Water and Plant:** "Let me start with defining the system that your example refers to. The system relates the bodies water and the turbines." – Student 27, Revised Draft

> **Water, Plant, and Electrical Grid:** "The Plant system is comprised of the power grid, turbines, water, and reservoir." – Student 539, Revised Draft

The students who identified the system as the "Water, Plant, and Electrical Grid" made up less than 5% of the total number of students. The number of students who did not identify a system and the number of students who did not turn in a First Draft were both greater than those who identified the “Water, Plant, and Electrical Grid” system. Therefore, these students will be grouped into one “Other” category for the remainder of the analysis.

In the First Draft, the highest percentage of students defined the system to include the water and the plant itself and fell into the “Water-Plant” category. Less than 15% of students included the Earth in their energy system. The full distribution of students can be seen in Figure 3.8.

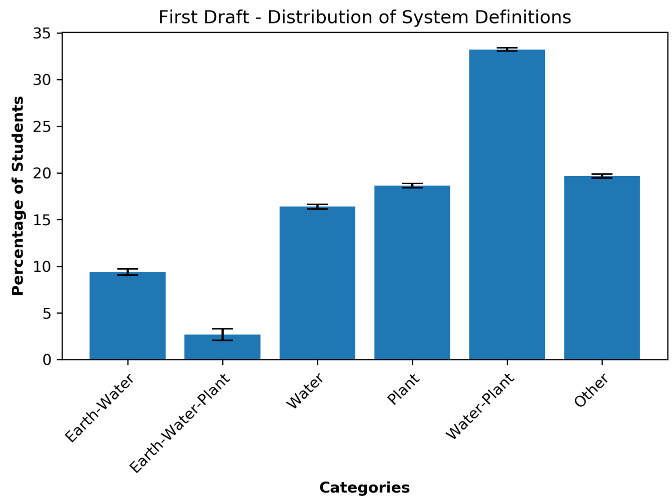

Figure 3.8: Distribution of students’ First Draft energy system definitions. Each bar represents the percentage of total students who defined their energy system in the corresponding category.

In the distribution shown, the second greatest category is the “Other” category. As mentioned earlier, only 5% of students used the “Water, Plant, and Electrical Grid” definition for their system. Most of this “Other” category is made up of students who either did not define their system in the First Draft, thus losing those points, or did not submit their First Draft at all. Tracking how these students learn is difficult as they did not originally define a system. The only progression that would be visible to us would be from nothing to whatever they wrote in their Revised Draft.

The next two most popular categories were the “Water” and “Plant” categories.

The "Water" category is not an unexpected answer from students. The way in which Gravitational Potential Energy is usually introduced is in the form "mgh". This equation focuses the attention on whatever the "m" is, in this case water. The "Plant" category is a little more surprising. This category is populated mostly by students who thought that the "system" referred to the facility itself. The thinking being that the Ludington Pumped Storage Facility was generating and storing the energy, so that is the energy system. This confusion happens due to the flexibility of the word "system" and is an issue that often arises in physics.

In this First Draft, we saw many different responses from students about what the system was. Some were expected, while others made us think about how we were phrasing the question. However, most students did not correctly identify the system in their First Draft.

3.3.2 Peer Review

In the First Draft submissions, only 81 total students had included the Earth in their defined system. This accounts for both students who used the "Earth-Water" and "Earth-Water-Plant" systems. Each of these students were assigned three other student's submissions to give feedback on during the Peer Review portion. In the Peer Reviews, 65.4% of students who had the Earth in their First Draft system did not advise any of their assigned peers to include the Earth in their system. Only 6.2% of these students gave this feedback to all three of their assigned peers, and only 11.1% of the students told two of their assigned peers to include Earth in their system. Another 17.3% of these students advised only one of their peers they were assigned to incorporate the Earth into their system. In total, 34.6% of students who included Earth in their First Draft system advised at least one other student to do the same.

The hope was that if a student had understood how to define a system, that they would help others do the same through their written feedback. However, most did not do this, almost two-thirds. This shows that even if a student could correctly understand how to define their system, it did not mean that they were confident enough to advise someone else to do the same. There is a value in having students confront these difficult concepts in a social setting. These are complex ideas and students may not always be confident in their ideas alone.

Of the remaining students who did not have the Earth in their First Draft system, 98.3% of them did not tell their peers to include Earth in the system. This statistic is more in line with what was expected. The students who did not correctly identify the system did not tell others to include the Earth in theirs.

At a first look, it seems that the Peer Reviews may not have made as big of an impact as we had hoped. With only a small population of students who correctly identified the system in the First Draft, it did not leave many in the position to help their peers correctly identify the system. Despite this, we saw a large rise in the number of students who included the Earth in their Revised Draft system.

3.3.3 Revised Draft

Although the system of the "Water-Plant" was still the dominate definition of the system among students in the Revised Draft, the number of students who included the Earth rose. The "Earth-Water-Plant" and "Earth-Water" systems were the second and third most common categories. In context this is a large improvement. Beforehand, these categories ranked at the bottom, but in the Revised Draft we saw a surge of students including the Earth in their definition.

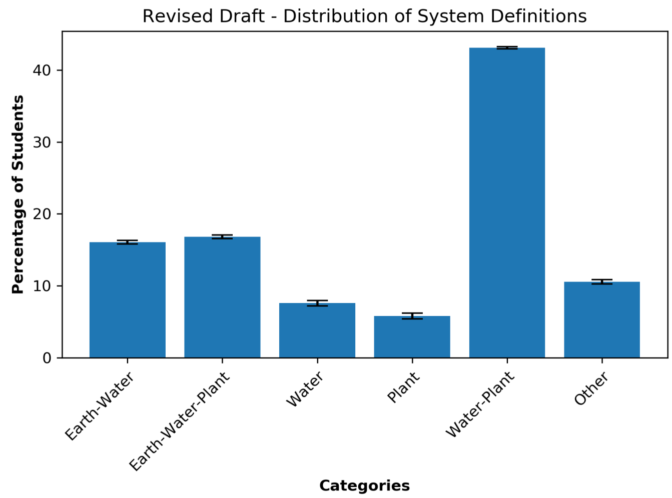

Figure 3.9: Distribution of students' Revised energy system definitions. Each bar represents the percentage of total students who defined their energy system in the corresponding category.

While we did see the categories that include the Earth rise, the "Water-Plant" system category also rose. This is most likely due to the large number of students giving Peer Reviews who had used the "Water-Plant" system in their First Draft. If a student received multiple Peer Reviews from students who advised a "Water-Plant" system, they would be more likely to change to that system. Even though the categories of "Water", "Plant", and "Other" decreased in the number of students, the number of students who included the Earth in their system was still in the minority. Less than 40% of student had the Earth in their system in the Revised Draft. This highlights the difficulty that students have in defining a system, and thus the importance of this activity.

3.3.4 Overall Progression

Looking at each individual component of the WTL activity can give some insight into what happened during it. However, to get a full picture, it is best to look at the overall progression of students and their writing. After all, the goal of the activity is for students to learn and progress towards a learning goal, so looking for their progress and the possible reasons of why they made that progress is essential.

As a reminder, the learning goal that we are investigating is "The importance of defining an energy system and how to choose the entities to include in it". In our WTL activity, the most simplistic system to choose while still being able to discuss Potential Energy was the "Earth-Water" system, however we still accepted the "Earth-Water-Plant" definition as a valid alternative.

When comparing the distribution of system definitions in the First and Revised Drafts, students' progression became clearer. From Figure 3.10, one can see that both the definitions which include Earth increased in popularity in the Revised Draft. Both the definition of just the "Water" or just the "Plant" decreased from First to Revised Drafts. While these changes seem like students were moving towards the right path, there is still the sizeable increase in students who define the system as the "Water-Plant".

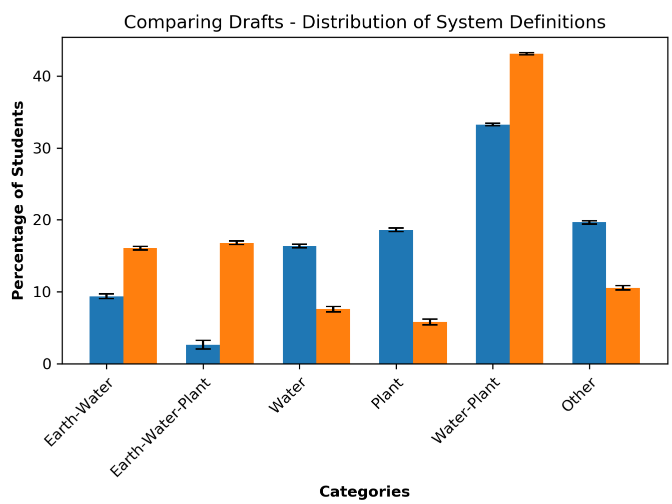

Figure 3.10: Comparison of the distribution of students' First Draft and Revised Draft energy system definitions. The First Draft definition is in blue, while the Revised Draft definition is in orange.

Through the direct comparison of First and Revised drafts we were able to identify some trends in students' behavior, however the distribution comparison alone leaves us with more questions. The heatmap that can be seen in Figure 3.11 adds a greater context to what we are seeing in the distribution comparison.

In this heatmap, the darker shades represent a higher percentage of students who fall into that specific path. The system definition that a student began with in their First Draft is plotted on the vertical axis. On the horizontal axis, the system definition present in a student's Revised Draft is plotted. This results in a traceable heatmap of students' progression during the prompt.

The path that most students followed was where the system was defined as the "Water-Plant" in their First Draft and stuck with this definition in their Revised Draft. The "Water-Plant" definition also gained students who originally defined the system as either just "Water" or just "Plant". This trend explains the increase that we see in the distribution comparison.

Although the signal is weaker than the two paths described above, the majority of students who are adding the Earth into their system are coming from those who originally indicated the system as "Water-Plant" in their First Draft. This was the largest population of students in the First Draft, so it makes sense that these students would account for the largest shifts in the Revised Draft.

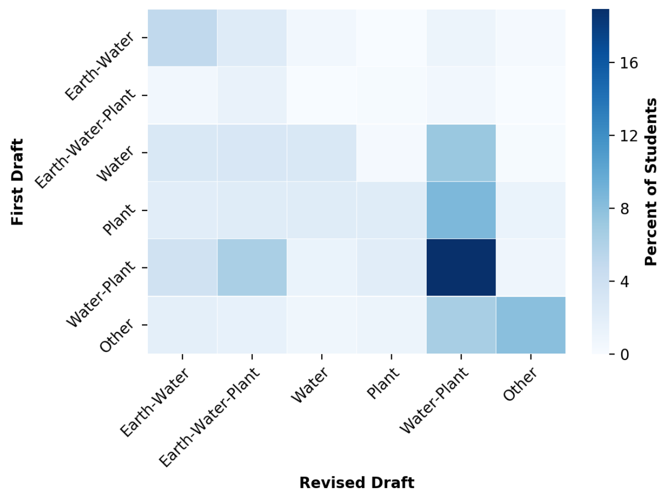

Figure 3.11: Heatmap of student revision moving from the First Draft to the Revised Draft. Students original definition of the system is along the y-axis, labeled by First Draft, and the students revised definition is along the x-axis, labeled by Revised Draft. The darker blue represents a larger percentage of students who fall into that category.

Through the heatmap, we are able to see how students revised their knowledge and we can get a sense of their progression towards the learning goal. However, we are not able to see *why* students are making these changes. We hypothesized that students would be adding Earth into their system due to one of several possible mechanisms; receiving a Peer Review which advised them to include the Earth in their system, reading a peer's First Draft in which Earth was a part of the defined system, receiving feedback from their Writing Fellow to include the Earth in their system, and coming in for Office Hours to discuss the system. We do not have data for the last two possible reasons, however, we can investigate first two.

The total number of student who included the Earth in their Revised Draft system was 221. Of these 221 students, 28.5% of them had already had the Earth in their First Draft system. 4.98% of the 221 students that included Earth in the Revised Draft only received a Peer Review that told them to include the Earth in their system. However, a sizable 22.6% of these same students read a peer's First Draft that had Earth in the system. 5.43% of the 221 students that had the Earth in their Revised

Draft both received a Peer Review that told them to do so and read another student's First Draft that included the Earth. Due to our inability to investigate Writing Fellow comments or Office Hour attendance quantitatively, 38.5% of the 221 students who included the Earth in their Revised Draft are unaccounted for and included it in their Revised Draft for unknown reasons.

Generating heatmaps in relation to the specific reasons for change adds additional confirmation to these results. The total number of students who received at least one Peer Review that advised them to include Earth in their system was 57. We separated students into two groups, one which received these specific Peer Reviews and one that did not. In doing this, we can see subtle differences in the patterns of the heatmaps. For students who received this specific feedback, a larger percentage of people who originally began with a "Water-Plant" definition added Earth to their system than those who did not receive the feedback. The students who received the specific feedback also did not shift to "Water-Plant" as frequently if they did not have it in their First Draft.

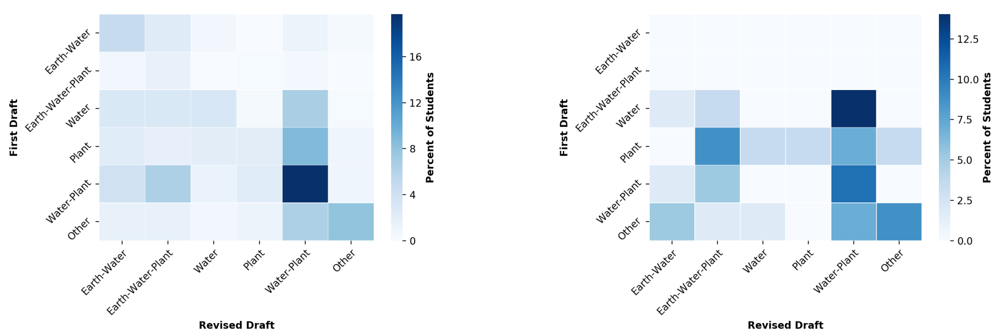

Figure 3.12: Heatmaps comparing student revision for students who did and did not receive a Peer Review telling them to include Earth. Students who did not receive a Peer Review advising to include Earth in their system is on the left (N = 614). Students who did receive such a review are on the right (N = 57).

The total number of students who received Peer Reviews advising them to add the Earth into their system is relatively small when compared to the total number of students. Earlier we showed that only a small number of students who had the Earth in their Revised Draft system had received these reviews anyway. The more interesting variable to look at is those students who read another student's First Draft that included the Earth in their First Draft system. This is a much higher number of students (N = 210) and as shown previously, had a larger impact on the inclusion of the Earth into Revised Draft systems.

So now, by separating the students into two groups, one in which students read at least one peer's First Draft with the Earth included in the system and another group where students did not read this, we can once again compare the heatmaps of two groups. In the group where students read a peer's First Draft with the Earth in the system, a wider range of students are editing their drafts to include the Earth in their system. We can see students shifting form all original definitions, including "Water", "Plant", and "Water-Plant".

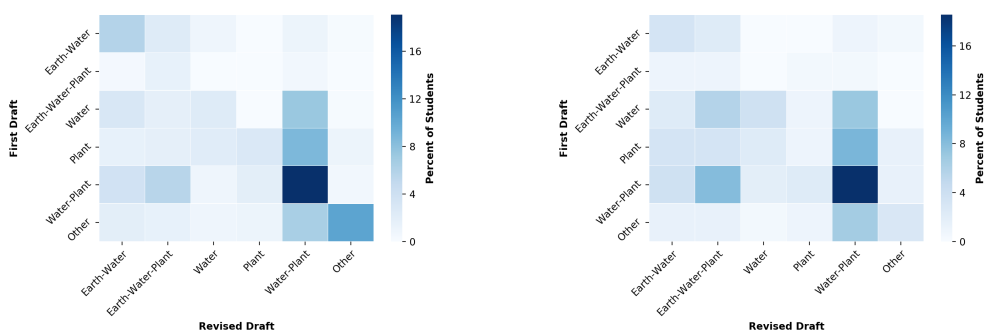

Figure 3.13: Heatmaps comparing student revision for students who did and did not read a First Draft with the Earth included in the system. Students who did not read a First Draft that included the Earth in the system is on the left (N = 461). Students who did read such a First Draft are on the right (N = 210).

We have shown how students shift their definitions of a system on a large scale. We found that while not a majority, a significant number of students were able make positive progression in our learning goal. The students who added the Earth to their system in the Revised Draft had been significantly influenced by reading other students' work. The trends identified are visible in the heatmaps shown. The trends offer insight into how students refine their understanding about systems, specifically those that involve the Earth.

3.4 Student Examples

The quantitative data gives a good picture of how students progressed towards the learning goal overall. We can notice trends in the data that show promising results from the WTL activities. However, to better understand how the mechanisms of learning really work in these activities, it is useful to look directly at the progression of students in a qualitative way. The selected examples serve that purpose as they show how different students may progress in the learning process.

3.4.1 Example Study 1: A student who added the Earth due to the Peer Review feedback

Student 70 originally defined their system as the "Water-Plant" system, but changed their system in the Revised Draft and defined it as a "Earth-Water-Plant" system. They begin their First Draft by defining what the energy system is.

> "The Ludington Pumped Storage Plant is a hydroelectric plant that harnesses the energy from water and turns it into electricity. The energy system that the team keeps mentioning consists of the water, the six turbines that the plant has, and the generators to which the turbines are connected."

The entities of the turbines and generators are considered all part of the Plant when the categorization is done. There is no mention of the Earth in this definition, however, when Student 70 discusses the energy transfers that occur within the system, they make explicit use of Gravitational Potential Energy.

> "Initially, the water has gravitational potential energy when it is sitting in the reservoir high above the shore. Then, when the water flows downward, the potential energy is converted to kinetic energy."

This interpretation of Potential Energy is consistent with what has been reported by the PER literature. Student 70 states that "the water *has* gravitational potential energy" (emphasis mine). This ontology places the Potential Energy within the object, and not within the interaction.

During the Peer Review process, Student 70 read three other student's First Drafts. Perhaps surprisingly, all three drafts that the student read described the system as just the "Plant" system. Student 70, did not receive any beneficial examples through reading other students' First Drafts.

However, Student 70 did receive two Peer Reviews that gave feedback advising the student to add the Earth into their system. The Peer Review that did not give this advice was confirmatory of their original system definition.

> **Student 173:** "The system is well defined. The reference to both the turbines and the generators is essential for understanding how the plant works, so it is great that you mentioned them in your paper. This was well done."
>
> **Student 292:** "The author defines the what he/she thinks the system is clearly. However, the book talks about systems of energy occurring between the object that gravity is working on and the Earth. So in this case, I believe the big picture system is between water and the Earth, as the water is what PE is actually stored in. The turbines, generators, and reservoir are part of the smaller picture system in which the water with energy flows through"
>
> **Student 492:** "The system that the energy flows in and out of is defined explicitly. However, an area of improvement that could be made is to include the Earth in the system because otherwise there is no potential energy."

Student 70 was a part of Section 2, and thus received feedback on their First Draft from the Writing Fellow responsible for Section 2. In this feedback, the Writing Fellow did not explicitly direct the student to include the Earth in their system, but rather told the student to think about their system in broader terms.

> "-Elaborate on the energy transfers when the water flows from the lake to the reservoir.
>
> -Think of the system in broader terms."

This indirect prompt from the Writing Fellow, paired with the direct feedback advising the inclusion of the Earth in the system from two of the Peer Reviews, Student

70 was able to revise their definition and make progression towards the learning goal. In their Revised Draft, they defined the system as the "Earth-Water-Plant" system.

> The Ludington Pumped Storage Plant is a hydroelectric plant that harnesses the energy from water and turns it into electricity. The energy system that the team keeps mentioning consists of the water, the six turbines that the plant has, the generators to which the turbines are connected, and the Earth whose gravity causes water to flow.

The revision that the student did make in this Revised Draft is not necessarily significant in word count, but represents a shift in understanding how Potential Energy relates to the system that is chosen. However, the student did not revise the words they used when they tracked how energy was transferred through the system. The student still claims that "the water *has* gravitational potential energy energy' (emphasis mine). They did not shift where the Gravitational Potential Energy is accounted for and still tracked it as being held *within* the water. Through this careful reading, we can see that the student made progression towards the learning goal, but failed to reach the point where they understand the implications of choosing their system.

3.4.2 Example Study 2: A student who added the Earth due to the First Drafts they read

Student 346 originally defined their system in a way that became categorized as simply the "Plant" definition. The student claimed that the system was the turbine generator and that this system moved the water around as a way of moving energy.

> "The system that the energy is flowing into and out of is the turbine generator. The turbine generator converts the energy that flows into it from the lake into the electric energy that flows out of it. The turbine also pumps water from the lake into a reservoir. The reservoir is able to store a large amount of the lakes water."

When Student 346 discussed the transfer of energy within the system however, they referred to the water as a sort of fuel, or physical manifestation of the energy that was in the “Plant” system. The way that this student discussed the transfer of energy shows that they did not understand how their defined system would affect the way they should discuss energy within it.

> “When the water is sitting in the reservoir, it is stored as gravitational potential energy. When the water is moving through the system it is transferred to kinetic energy. It is also further transferred into the electrical energy that the process produces.”

The Peer Reviews that this student received are vague and could have possibly lead to more confusion on the part of Student 346. Each of the three Peer Reviews are complimentary on how they defined their system, but offer soft suggestions of ways to improve their system definition.

> **Student 39:** “The author does a good job clearly stating in their first sentence what the system is. The author could improve by incorporating the reservoir and the lake more into the explanation.”
>
> **Student 76:** “You described how the system flows in and out. You showed how the system is from the turbine generator, you could improve on the specifics of what is the energy that flows in and out.”
>
> **Student 393:** “Well done explaining both aspects of the system and how they move through the turbine itself. However you might want to rephrase when you said the turbine itself is what pumps the water into the reservoir.”

The feedback from Student 39 asks the student to include the reservoir and lake into the explanation. It is unclear if Student 39 is referring to the water itself when they mention the reservoir and lake, or if they mean the containers that hold the water. They also state that Student 346 should incorporate this into their explanation, leaving it ambiguous if Student 346 should add these entities to their actual system definition. Both Students 76 and 393 ask Student 346 to either give more specifics or rephrase certain aspects of how the energy is moving in the system, but did not advise Student 346 to do anything about the system definition itself.

Student 346 was given three other students’ First Drafts to read during the Peer Review process. Unfortunately, one of these First Drafts failed to define a system in the First Draft. Student 346 recognized this failure and gave feedback that advised the student to include this portion of the assignment in their Revised Draft. The second First Draft that Student 346 reviewed, a draft written by Student 498, had no clear definition of the system either. Student 346 asked in their feedback for this student to clarify what Student 498 meant as they did not understand what Student 498 was attempting to define as the system. The First Draft that Student 346 read was written by Student 294, and this draft included the Earth in their system. The defined system fell into the "Earth-Water" category of systems. An excerpt from Student 294's First Draft is included below.

> "The system that you are speaking of is an Earth-water system. In order to use the law of conservation of energy, the Earth-water system is closed, which means that the Earth and the water are isolated from other factors of the environment."

In response to this First Draft, Student 346 gave the following Peer Review;

> "The author did a good job explaining and defining the system that energy is flowing into and out of. They stated what the system is and in what conditions the system is present in. Maybe elaborate more on what a closed system is and how that applies to this project."

Student 346 recognized the clear definition of the "Earth-Water" system and did not have an issue with Student 294 defining the system this way. However, Student 346 did find the claim that the system was closed a bit confusing and asked Student 294 to clarify what they meant by this. In Student 346's Revised Draft, they defined the system as "Earth-Water". It is plausible from that Student 346 created their Revised Draft system definition based on what they read from Student 294.

Before we can make this claim, we must take a look at the Writing Fellow feedback that Student 346 received. Student 346 was a part of Section 9 and thus received feedback from the Writing Fellow responsible for this section's grading. The feedback from this Writing Fellow can be seen below.

> "Good first draft! Please be sure to meet the minimum word requirement before submitting your final draft! In addition, you may want to reconsider the system

> that the energy is flowing in and out of. We are thinking broader terms here! If you were to throw a ball in the air, the system that creates the KE and GPE would be the ball and the Earth! This is similar to the situation we are looking at. Also, it may help to explain the energy transfers in a little bit more depth!"

In this feedback, the Writing Fellow does not explicitly tell Student 346 that the system must include the Earth. However, the Writing Fellow does give the student an analogous situation of the ball and the Earth system. The analogy given in the Writing Fellow feedback could have coupled with what Student 346 read from Student 294's First Draft to result in Student 346's Revised Draft system definition of "Earth-Water". Below is Student 346's Revised Draft system definition.

> "The system that the energy is flowing into and out of is the water and the Earth. The turbine generator converts the energy that flows into it from the lake into the electric energy that flows out of it. The system also pumps water from the lake into a reservoir. The reservoir is able to store a large amount of the lakes water."

In this excerpt, the only two changes that were made was a swap of "turbine generator" to "water and Earth" and a deletion of "turbine" in the third sentence. These changes may reflect a progression in the learning goal of the how to define a system of energy, yet the limited amount of revision does not make an overall convincing argument. Furthermore, the student did not revise how the energy moves in the system. This points to the possibility that while Student 346 may have a better grasp on how to define systems, they do not quite understand the importance or implications of choosing a specific system.

3.4.3 Example Study 3: A student who added the Earth due to Writing Fellow feedback

Student 405 had originally defined the energy system as a "Plant-Water" system. Through the course of the WTL activity, the student ended with an "Earth-Water" definition in their Revised Draft. The way in which Student 405 originally defined their system is included in the excerpt below.

> "The energy is stored in the water and the system created by the six turbines. The six reversible turbines that were discussed in the article allow for the water to be transported up to the reservoir to be stored as the article described by pumping the water upwards to the reservoir and then releasing it later to produce electricity."

This student defined the Potential Energy held in the system in relation to the height of the water in the system. The higher the water was, the higher the Potential Energy became. This is consistent with the "vertical location" ontology identified in the PER literature.

Student 405 only received two out of the three Peer Reviews that they should have. This is a result of one student not completing their Peer Reviews. Both Peer Reviews that Student 405 received criticized the way that they defined their system, but did not offer concrete ways to improve the definition.

> **Student 434** "The author mentioned the systems, but did not define the system very well. It is not made clear what exactly makes out the systems."
>
> **Student 585** "The author does not directly identify what the system is. However, they do explain how the system works with the water flow and how the energy is changed. One improvement would be to define what a system is so that the reader has a clearer understanding when reading the rest of the memo."

Neither of the two Peer Reviews offered any suggestions that would have directly led Student 405 to include the Earth in their defined system. Now we looked to the First Drafts that Student 405 read during the Peer Review process. Two of these First Drafts, written by Students 56 and 118, shared the same system definition as Student 405, the "Plant-Water" definition. Student 405 gave affirmatory feedback to both Students 56 and 118. The third First Draft that Student 405 read had no system defined. Student 405 did write feedback that advised this student to include a definition in their Revised Draft, but did not write what this definition should exactly be.

Student 405 is in Section 10 and thus received feedback on their First Draft from the Writing Fellow responsible for this section. The feedback from this Writing Fellow is given below.

> “-Explain the system explicitly. When a ball is thrown (or moved) up, the energy system is described as Ball-Earth system. What is the system for the water?
>
> -The energy conversions need to be described as to what happens after the energy is converted to its potential form.”

Similar to Writing Fellow 9, Writing Fellow 10 gives the ball-Earth system analogy in their feedback. It seems that this is enough to encourage Student 405 to change their system definition in their Revised Draft to a “Earth-Water” definition. Even though Student 405 received two Peer Reviews that did not offer this specific advice, the importance placed by the student on the Writing Fellow feedback was enough to motivate their revision. Below is Student 405’s Revised Draft system definition.

> “In this case, the system can be defined in multiple ways. The simplest of such is defining the system as the water and the Earth. In this Water-Earth system, the water is moved up and down using the turbines of the plant. ”

In their Revised Draft, Student 405 acknowledged that the system could be defined multiple ways, but pointed out that they would define it in the simplest way. This shows that the student recognized the implications of selecting an energy system. Student 405 reached the full learning goal in that they understood the importance of defining an energy system and how to do so correctly.

3.5 Post-exam survey results

For the final part of the analysis of our WTL activity we examined student responses to post-exam surveys that were conducted during the Fall 2018 term. Our first question was asked on the post-Exam 2 survey. The goal of this question was to ask students how they felt about M-Write and if they believed that M-Write helped them learn. A total of 424 students answered this survey question, about two-thirds of the total number of students.

23. What is your opinion about M-Write activities thus far?

a. I **like** M-Write, and the activities are **useful** for my understanding of physics
b. I **like** M-Write, but the activities are **not useful** for my understanding of physics
c. I **dislike** M-Write, but the activities are **useful** for my understanding of physics
d. I **dislike** M-Write, and the activities are **not useful** for my understanding of physics
e. I am neutral towards the M-Write activities

Figure 3.14: Survey question on student opinion of M-Write. This was included on the post-exam survey for Exam 2. N = 424.

Perhaps to slight dismay, the response that had the highest number of students was selection D, that they disliked the activities and did not see them as useful for their learning. Responses A, C, and E gathered about 20% of students each. Response B had just over 10% of students. A single student selected B,D, and E. This could be the result of confusion as other survey questions instructed the students to select all that apply.

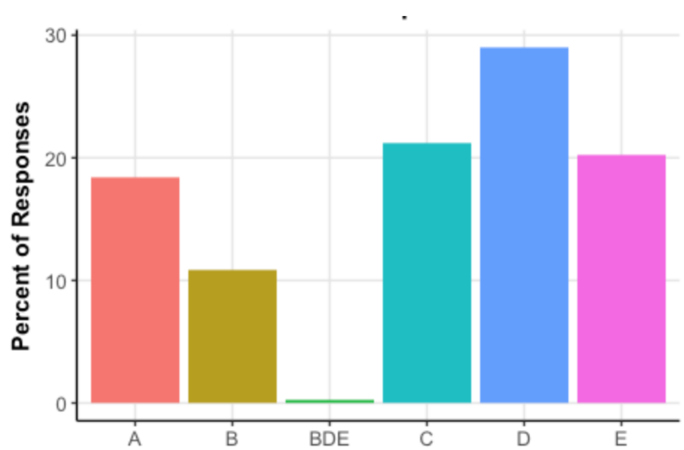

Figure 3.15: Survey responses of student opinion of M-Write. *Figure provided by Nita Kedharnath.*

These results show that the majority of students did not like engaging in the WTL activities and felt they were not learning through them. These results could suggest that simply adding the activities on top of their other work while not devoting a large portion of their grade to them is detrimental to their effectiveness. It could cause students to be both unhappy with the workload and do not see them as important as other aspects of the course.

One might expect that students who feel differently about the WTL activities may perform differently on the exam. Therefore, the Exam 2 scores for each group of respondents was then plotted using a box plot. Each of the groups' average Exam 2 scores ended up being within one quartile of all other groups. This shows that how students thought about the M-Write activities did not significantly effect their performance on Exam 2.

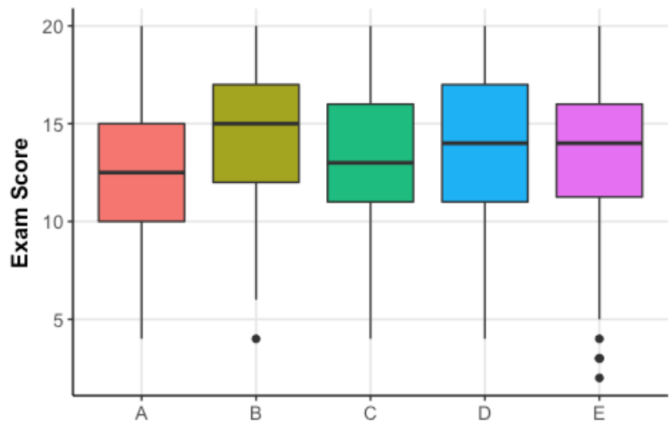

Figure 3.16: Survey responses of student opinion of M-Write related to grade data. The one student who selected BDE was taken out of this data set. Each is a boxplot showing the scoring quartiles. The exam was out of a total 20 points. *Figure provided by Nita Kedharnath.*

Those students who selected response B, that they liked M-Write but did not find them useful, had a slightly higher exam score than the other groups. However, this is not in any sense significant and musing on reasons for this would not yield any important findings.

The second question that we asked about M-Write was included on the post-Exam 3 survey. In this question, we wanted to gauge how students were using the M-Write specific office hours. A total of 370 students answered this survey question, a decrease of 54 from our previous M-Write survey question.

24. For the M-Write assignments, M-Write Fellows hold office hours. Have you used these office hours?

a. Yes, I have used the M-Write Office Hours and find them helpful
b. Yes, I have used the M-Write Office Hours, but they are not helpful
c. No, I have not used the M-Write Office Hours because I didn't have questions on the assignments
d. No, I have not used the M-Write Office Hours because I can't make it at the times they are offered
e. No, I have not used the M-Write Office Hours because I was not aware of them

Figure 3.17: Survey question on student usage of M-Write office hours. This was included on the post-exam survey for Exam 3. N = 370.

An overwhelming number of students responded that they had never used the M-Write office hours because they did not have questions during the M-Write activities. While the office hours had a consistent flow of students during the semester, this large number of students responding they had never came was not unexpected. It seemed to be a select number of students who came for office hours. The next highest number of students responded that they had come to office hours and they found them helpful. Very few people had come to office hours and found them unhelpful.

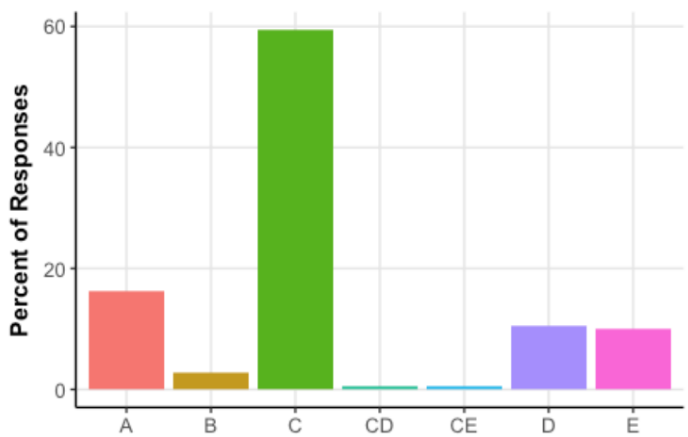

Figure 3.18: Survey responses of student usage of M-Write office hours. *Figure provided by Nita Kedharnath.*

These results show that the office hours that were being held were helpful for those who came. Selections D and E had a number of students responding for each. These responses caused some worry as it means that the students who responded that they were too busy for our office hours could have benefited from them if they were offered at available times for them. A similar dilemma is present for students who said they had never heard there were M-Write specific office hours.

We again wanted to check if there was any correlation between student responses and the exam grades of students. The average Exam 3 score for each group of respondents was calculated along with the standard deviation of this average. The only significant difference in average sore was the average score of the students who responded that they were not able to attend office hours because they could not make the times that the office hours were held. These students scored significantly worse on the exam compared to the other group of respondents.

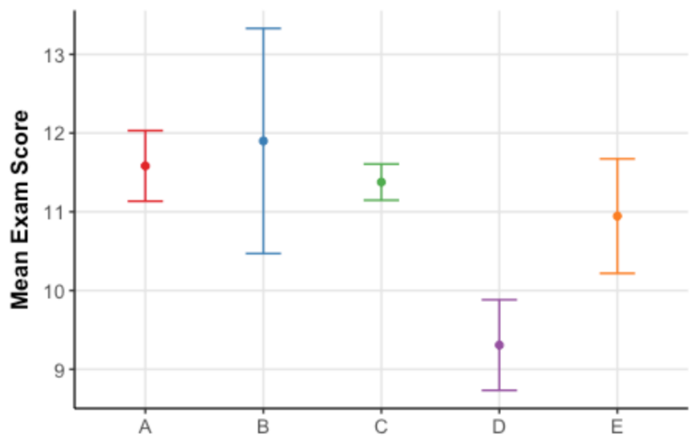

Figure 3.19: Survey responses of student usage of M-Write office hours related to grade data. Students who selected more than one survey response are not included in this data set. The average exam score for each respondent group is represented by the dot with the standard deviations on each average represented by the error bars. *Figure provided by Nita Kedharnath.*

The statistically lower average exam score for the students who could not make the M-Write office hours due to the times they were offered is worrisome and could be explained due to several reasons. It could be either that these students do not make enough time in their schedule for the course, and thus struggle in the course in general. However, it could also be that these students are truly too busy and do not have the time to commit to coming to office hours and studying in general. Either way, the solution to help these students is outside the scope of the WTL activities, but is something that should be addressed elsewhere.

Overall, the results from the post-exam surveys give us a look into what students were thinking about M-Write and how they interact, or don't interact, with in-person portion of the activities. From these survey results, we were able to adjust for future semesters to ensure the student experience will continue to improve.

3.6 Winter 2019 Preliminary Analysis

In this section, I discuss how students addressed our learning goal in the Winter 2019 term. I first give a brief overview of the grading in this term, along with average scores and distributions of scores given. Then, I examine how students progressed towards our learning goal during this term and compare the Winter 2019 results to the Fall 2019 results.

In the Winter term, we had the Writing Fellows themselves categorize the student submissions for the learning goal analysis. While this proved an effective way of generating our data, not all Writing Fellows completed their categorization of both the First and Revised Drafts. Therefore, only 345 submissions of the possible 603 student submissions are able to be used in this analysis. Another limitation of this Winter 2019 analysis is that we do not have full access to the Peer Review depository as that only becomes available after the semester. Therefore, we are able to track how students progressed towards the learning goal, but we are not able to investigate why this happened.

3.6.1 Grade Distributions

In Winter 2019, we used a rubric that kept the percentages that were given to each component the same, but decreased the point values in each. For example, the Revised Draft was now out of 26 points, but it still accounted for 65% of the total points of any on WTL activity. This was done to address the problems associated with having a grading scale that had too fine of a grain. In this term, we had 603 total students enrolled. This is less than Fall 2018, which had a total of 671 students.

On the First Draft, 572 students turned in their submission on time. The average score for students who submitted on time was 97.5%. This is comparable to the average percentage for students in the Fall term, which was 97.1%. So even though we changed to point values from 15 points to 6 points, the average grades on the First Draft stayed constant. The full First Draft grading distribution for the Winter term can be seen in Figure 3.20.

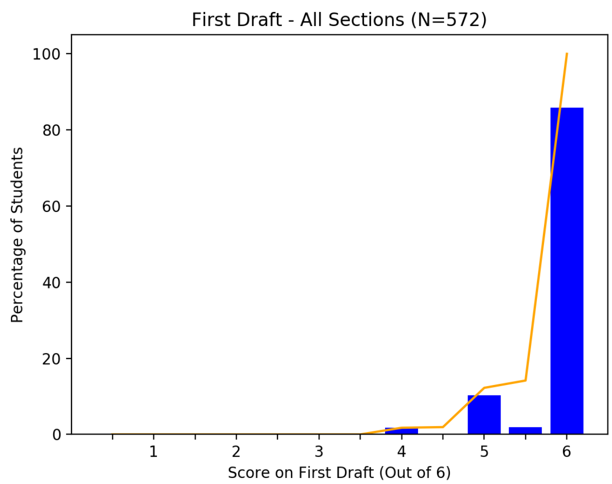

Figure 3.20: Distribution of scores on the First Draft for all students in Winter 2019 term. The blue bars represent the percentage of students with that score. The orange line shows the integrated number of students along the distribution.

Most of the students earned the full 6/6 on the First Draft, with about 10% receiving the 5/6 score. A few scores of 4/6 were given as well as 5.5/6. A higher percentage of students earned 100% scores in this First Draft than in the Fall semester. However, the loss of a single point carried more weight than it did in the Fall semester and this resulted in the overall average grade being comparable.

In the Peer Reviews for the Winter 2019 term, we had 477 students complete their feedback to their fellow students. This means that only 83.4% of students who completed a First Draft also completed their Peer Reviews. In Fall term, this percentage was 92.3%. It is surprising that the completion percentage dropped and shows that more time needs to be spent to motivate students to participate in all portions of the activity. Yet the average score of those who completed their Peer Reviews stayed about the same. In the Winter 2019 term, the average grade on the Peer Reviews was 97.2%, compared to Fall 2018's average grade of 96.5%. The distribution curve of the grades is also similar and can be seen in Figure 3.21.

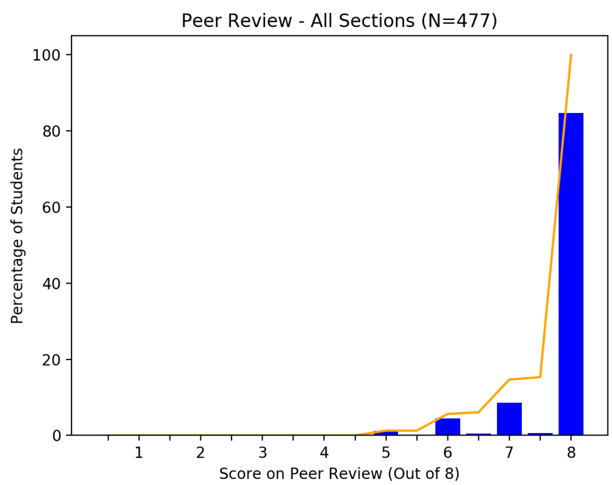

Figure 3.21: Distribution of scores on the Peer Reviews for all students in Winter 2019 term. The blue bars represent the percentage of students with that score. The orange line shows the integrated number of students along the distribution.

A total of 551 students submitted their Revised Draft on time for the Winter 2019 term, earning an average grade of 92.8%. This is lower than the Fall 2018 average score of 94.0%, but not by a large margin. Each are lower than the First Draft score in their respective term. A reason for the slightly lower average could be that Writing Fellows felt more comfortable deducting points when they had a clearer rubric to use. The distribution curves of these grades are slightly different. While they both have the 100% score as the most popular grade, the distribution curve in Fall 2018 had a feature that resembled a bell curve with a peak in the low-mid 90%. Whereas the distribution curve for the Winter 2019 term has grades that continue to increase in popularity as you move up the grading scale.

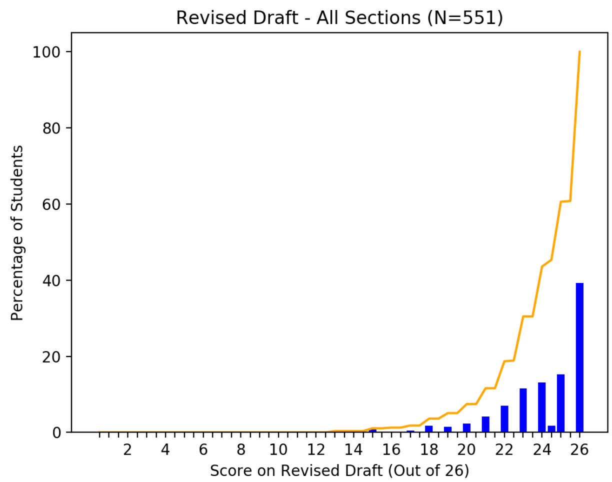

Figure 3.22: Distribution of scores on the Revised Draft for all students in Winter 2019 term. The blue bars represent the percentage of students with that score. The orange line shows the integrated number of students along the distribution.

Although certain features of the grading distributions may be different between the Fall and Winter terms, the average scores remained constant. This is most likely a result of changing the grading rubrics. It should not be a cause for concern. Through our Winter term, the grading of our WTL activities has been shown to be consistent and fair to all students.

3.6.2 Learning Goal Analysis

In Winter 2019, the same categories for the system were used as in Fall 2019. Therefore, direct comparisons between semester can be made as well as tracking student learning during the Winter semester. In the Winter semester, the largest category of students defined their system as the "Water-Plant" system. The percentage of students in the Winter semester who did this is slightly larger than that of the Fall term, but not by a significantly large margin. The distribution, compared with the Fall term, can be seen in Figure 3.23.

Although the "Water-Plant" system category is similar, there are striking differences in many of the categories. The "Other" category in the Winter term is much larger than that of the Fall. This is most likely due to the fact that the categorization was done by Writing Fellows. They were instructed that if they did not know how to categorize a certain student's response to place it in the "Other" category.

Both the "Water" and "Plant" categories account for a much smaller percentage of students than in the Fall term, while both the categories in which the Earth is a part of the system have more. It is difficult to point to a reason for this difference. A possible factor that contributed to this feature is instructors providing more introduction to the activities in class, or Writing Fellows being more aware of how to instruct students during office hours. However, no direct correlation can be found, and these are just possible contributing factors.

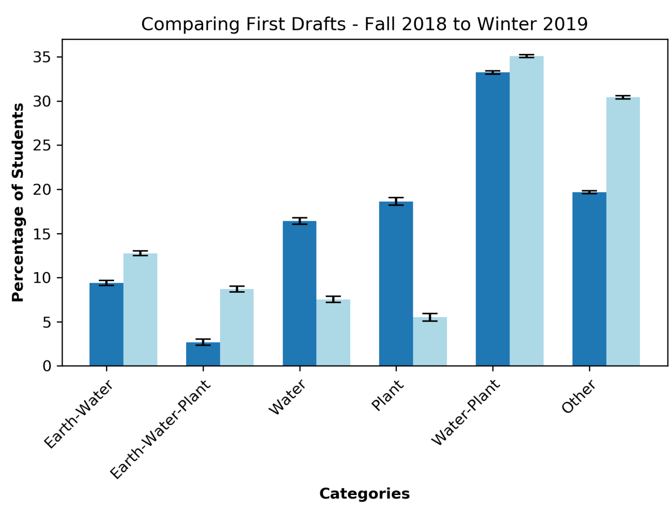

Figure 3.23: Comparison of the distribution of students' First Draft energy system definitions from Fall 2018 to Winter 2019. The Fall 2018 definition is in dark blue, while the Winter 2019 definition is in light blue.

In the Revised Draft of the Winter term, we saw more than half of the students include the Earth in their system. When comparing the category distribution to the

Fall term, it is clear that there are more students in the "Earth-Water" and "Earth-Water-Plant" system categories. While there were more to begin with in the First Draft as well, the margin in the percentages increased in the Revised Draft. Similarly to the First Draft, both the "Water" and "Plant" system groups still have less students in the Winter term than the Fall term. Unlike in the First Draft comparison, the "Water-Plant" category in the Revised Drafts is much less in Winter than in Fall. This is a significant difference and is one of the more defining features of the Winter distribution.

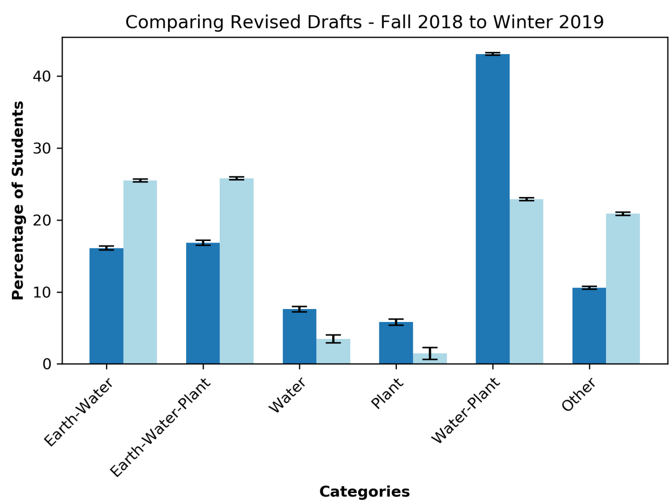

Figure 3.24: Comparison of the distribution of students' Revised Draft energy system definitions from Fall 2018 to Winter 2019. The Fall 2018 definition is in dark blue, while the Winter 2019 definition is in light blue.

As was true in the Fall term, it is necessary to look at how students progressed from the First to Revised Draft in the Winter term. Unfortunately, we cannot use any Peer Review data to gain further insight into reasons for students change, we are still able to look at how students changed. When comparing the system definition distribution between the First and Revised Drafts for the Winter term, we see a much clearer picture than we did in the Fall. The "Earth-Water" and "Earth-Water-Plant" system categories both increased dramatically, while all other categories decreased. This can be seen in Figure 3.25.

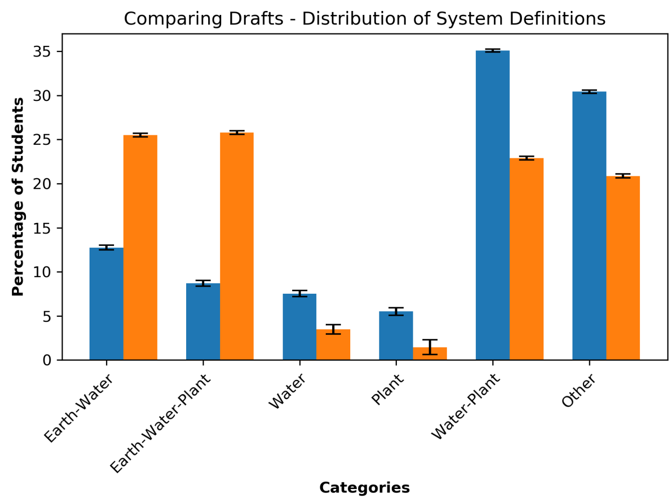

Figure 3.25: Comparison of the distribution of students' First Draft and Revised Draft energy system definitions in Winter 2019. The First Draft definition is in blue, while the Revised Draft definition is in orange.

To better understand how our students progressed from the First Draft to the Revised Draft, we again looked at a heatmap comparing the two. The progression of students in the Winter semester was much different than that in Fall. The largest number of students started in the "Other" category and stayed in the "Other" category. Again, this points to the fact that Writing Fellows were not always confident in their ability to categorize a student and is not necessarily a reflection of students learning progression.

The next largest amount of students actually moved from the "Water-Plant" system definition to the "Earth-Water-Plant" system definition. Unlike in the Fall 2018 term, in the Winter term, we see students moving away from this system definition adding the Earth into their system. A significant number of students did stick with the "Water-Plant" definition, but it is not as dominating of a trend as in Fall.

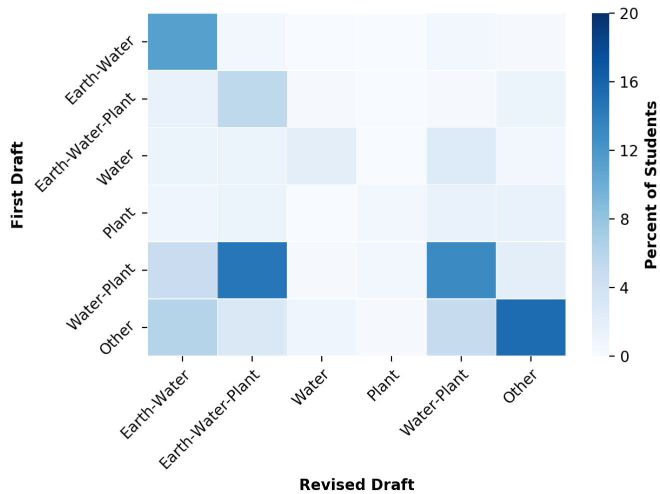

Figure 3.26: Heatmap of student revision moving from the First Draft to the Revised Draft for Winter 2019.

These early Winter 2019 results are promising. However, without the Peer Reviews data and no real significant way of understanding why students made the progression they did, it is tough to draw any concrete conclusions. It does seem that in the Winter 2019 term, students were able to better progress towards our learning goal and that the repeated implementation has improved our use of the WTL activity. These results offer a promising future for the WTL activities, specifically those surrounding our learning goal of “Understanding the importance of defining an energy system and how to choose the entities to include in it”.

CHAPTER IV

Discussion

In this final chapter, I discuss the conclusions that can be drawn from what we found in our Fall 2018 analysis. I evaluate how well our current WTL activity helped students progress towards our learning goal of "Understanding he importance of defining an energy system and how to choose the entities to include in it." I draw from both our quantitative and qualitative results to make these conclusions. I then show how our early analysis of the preliminary Winter 2019 data supports our findings which shows larger learning gains over the Fall 2018 semester. After which, I provide possible paths for future work. Including how we could improve the activities themselves to better lead students in their progression towards the learning goals, as well as other types of measurements that could be made using different methods in future work.

4.1 Conclusions

When this project initially began, the idea of using prominent Natural Language Processing techniques seemed very promising. However, it is clear that we were not able to use the methods that come straight out of the box with the LIWC software. The LIWC dictionary did not result in any meaningful difference between groups of submissions from the Spring term. It did not matter how these groups were created, whether it was by draft version, student demographics or performance. In the end, this method of analysis was abandoned and no clear conclusions could be drawn from the data. If these techniques are to be used successfully, a robust "Physics dictionary" would need to be created. We were able to separate out our three activities using the rudimentary dictionary that was created in this study. However, if one hoped to use this method to spot potential differences in student performance or other dimensions, the dictionary would need to be expanded and trained on a large data set.

In this thesis, it was important to include the grading methods and data for the WTL activities. This was done in order to give a full picture of what happened and to support the validity of our claims. The grades given in both the Fall 2018 and Winter 2019 terms resemble a homework-like distribution. This is not surprising as the First Draft and Peer Review were graded based on criteria that supported completion of the activity more than correctness. The Revised Draft has the lowest overall average score due to the shifted focus towards the correctness of the students response.

One may worry about how the grading compares between Writing Fellows. We found that the average scores given by our Writing Fellows tended to be consistent to the overall average. While some of the score distributions of certain Writing Fellows fell out of the norm, their average score stayed consistent. This did not create a large worry on the part of the instructional team.

An issue that did prove to be significant was students attitudes towards the WTL activities during the Fall 2018 term. As shown by the post-exam surveys, a large number of students did not enjoy the activities or feel that they were effective for their learning. The majority of students also did not attend the office hours to discuss the activities in person. While we still were able to measure students learning and the activities were shown to be effective to a degree, this lack of buy in from students could have caused the activities to not be as effective as they could have been. We had hoped that students would enjoy these activities more than other work in the course and thus put in more effort to them, however, this was not the case.

The main goal of this project was to investigate if WTL could service as a learning and evaluation tool for the specific learning goal of "Understanding the importance of defining an energy system and how to choose the entities to include in it." Part of the second WTL activity was structured around this learning goal and we looked specifically at how students addressed these parts. Due to WTL activities creating a large amount of student writing, we were able to clearly see how well students had addressed the learning goal and their progression through the activities. Evaluating this learning goal through WTL activities is possible and can be successful.

However, from our Fall 2018 results, it seems the success of WTL as an learning method for our specific learning goal is not quite as strong as its success as an evaluation tool. A very small percentage of students included Earth in their initial definition of the system. This is okay, as we hoped to see an improvement when the students wrote their Revised Drafts. We did see this improvement, but not as large of one as we would have hoped. In the end, less than 40% of students included Earth in their defined system in the Fall semester. There was a shift in students definitions of the system that show progress towards the learning goal, but it was not a majority of students. This percentage increased in the Winter semester. Our preliminary data from this semester shows that a majority of students did end the WTL activity with the Earth in their defined system and that the shifts in students understanding were clearly measurable. According to our early data, a majority of students in the Winter 2019 term were able to make progress towards our learning goal.

A considerable amount of importance is placed on the Peer Reviews in our WTL activities. They are 20% of each overall activity grade. This is under the guidance of the M-Write program and previous results of Peer Reviews impact on student learning (*Halim et al.* (2018), *Moon et al.* (2018)). However, we have shown that the students who made progress towards the learning goal in the Fall term often did not make this progress through direct feedback from their peers. The larger impact came from students reading other students work that they were assigned to review. This shows that the Peer Review process is valuable for student learning, but receiving direct comments from peers may not be as valuable as previously thought.

This result may be due to the small percentage of students who actually gave feedback that advised their fellow students to include Earth in their system. In the Peer Reviews from the Fall semester, people who included Earth in their system seem reluctant to tell others to do the same. This could show that although students are able to correctly identify a system themselves, they may not understand the importance of doing so. It could also be that these students are not confident in their own definition of the system and do not feel comfortable advising students to revise what they wrote.

Upon our qualitative investigation done through example students from the Fall 2018 term, the progression towards the learning goal shown through the quantitative analysis may not be as meaningful of a progression as originally thought. As seen through the student examples, the revisions made, while they correctly re-define the system, are sometimes superficial. Some students will only change the few words that move them into the correct defined system, but they do not continue this revision work throughout their submission. This shows that although the student may now understand how to correctly identify the system in this situation, they do not necessarily understand the importance or implications of defining the system in that way. While they have made some progress, they may not have reached the full learning goal. However, this same analysis is not yet able to be done for the Winter 2019 term, in which we saw more significant quantitative improvement in students definitions of the system.

Overall, we have found that WTL activities are a good fit in terms of an evaluation tool for the learning goal. The specific implementation used in the Fall 2018 term did see limited successes in terms of a learning tool for our specific learning goal. In the Fall term, students did make progress towards our learning goal, but it is unclear how much progress was truly made by each individual student. In contrast, a majority of students were able to correctly identify the system in the Winter term. The learning progression in this term was clear from students shifting their system definitions to include the Earth. The results of this thesis show a promising start for WTL activities in Physics 140. Specifically concerning the learning goal of "Understanding he importance of defining an energy system and how to choose the entities to include in it". We were able to clearly see which students struggled initially as well as those who were able to make progress in it. Future work should be done to continue development and research surrounding WTL in Physics 140 to improve how we support students and their learning.

4.2 Future Work

For the immediate future, work should be devoted to the development of the next generation of WTL activities for the Physics 140 course. We have shown that WTL can be effective for evaluating and learning physics, but more work needs to be done to refine our activities. A possible path for this work could be to make the WTL activities more precise and aim to tackle a single learning goal in each activity. This would shorten each activity, giving the instructional team the ability to conduct more WTL activities in the semester that cover a wider range of topics. Shortening and creating more precise activities may also increase student attitudes towards the activities. This would lead to more buy in from the students and possibly make them more effective.

The next generation of activities could also be muli-step. What we saw in the Fall and Winter terms was a reluctance from students to revise their original drafts. When they did revise, most revisions were surface level. Using an approach where you ask students a new question in the Revised Draft would serve to create more revision work and perhaps be a better guide towards a learning goal. This could take the form of asking students to add on to what the First Draft asked of them or to specifically clarify what they said in a different way. This type of structure could vary depending on which learning goal was being addressed.

From the Fall 2018 analysis, we found that reading other students work was a significant factor in how students revised their own work. Using the model where everyone is assigned three other peers at random results in some students never reading a First Draft that uses the correct physics. A possible improvement one could make to foster more revision that advances every students understanding is to plant an instructor created, or selected, First Draft into every students assigned Peer Reviews. This would ensure that everyone is able to read a submission that uses correct physics. The instructor planted draft could be set so that it is indistinguishable from the other drafts. The students would still have to compare the drafts they read and decide what is the correct physics for themselves, but with the added benefit of receiving at least one correct version. Using this model, it would also be possible to plant two draft, one with correct physics and another with incorrect physics reasoning. A possible study could be investigating how students digest the different interpretations and tracking in which ways they progress towards the learning goal.

Along with the continual development of the prompts used in the WTL activities, the creation of pre/ post evaluations would be beneficial in studying the WTL activities effects on student learning. A good model for this can be found in the study by Moon et al., in which the pre/ post assessment consisted of multiple choice questions paired with a confidence rating and space for a written explanation for each answer choice (*Moon et al.* (2018)). Questions could be created by researchers for each learning goal associated with the created prompts, or be adapted from standard assessments, such as the Force Concept Inventory.

WTL activities could be implemented all across the landscape of physics courses. The large introductory course setting gives a researcher a large set of data points for measuring the learning impact, as well as providing a wide range of possible interpretations of physical problems. However, the activities would still be useful in both upper level courses as well as courses with smaller class sizes. In a smaller course, Peer Review does not necessarily need to be coordinated by an online distribution system. The smaller scale could benefit from in class discussions surrounding particular prompts or even Peer Reviews written in class and discussed in person. In this smaller class, the instructor could also provide feedback for students. This would have both benefits and costs. While the feedback could be more precise if given by an instructor, the students may feel more scrutinized if they knew the instructor was reading their First Drafts. With a near peer mentor providing feedback on the First Draft, students may feel more comfortable and not afraid to make mistakes. In any case, WTL activities have the ability to see success in both courses of small and large class sizes.

There is a bright future for WTL in physics. In introductory courses, it is not often that students are asked to specifically explain fundamental concepts in words. Through these activities, we have the opportunity to do just that. We can see where students struggle and will be able to continue to improve the support for helping students reach our learning goals.

APPENDICES

APPENDIX A

Fall 2018 Grading Guide

M-Write Grading Guide - Physics 140 Fall 2018

First Draft

In grading the First Drafts, there are four main things you should be looking for as an M-Write Fellow. These four questions will be used to assign points in each of the four categories in the rubric, which has been included below.

1. Did they attempt to address all of the physics questions (did they include all of the “be sure to’s”)?
2. Did they write in a way that seems appropriate for the situation they are in?
3. Can you understand and follow the physics that they describe in their response?
4. Did they meet the word requirements and write in an acceptable format?

FIRST DRAFT RUBRIC

		Strong	Sufficient	Weak	Insufficient	Points awarded
Physics Concepts		*7 points* **Attempts** to address **all** parts of the prompt using relevant physics concepts	*5 points* **Attempts** to address **most** parts of the prompt using relevant physics concepts	*3 points* **Attempts** to address **few** parts of the prompt using relevant physics concepts	*0 points* **Attempts** to address **no** parts of the prompt using relevant physics concepts	
		Strong	Weak	Insufficient		
Style of Writing	Perspective	*3 points* **Appropriate** voice for the role of the writer and target audeince	*2 points* Voice is **inconsistent** with the role of the writer or target audience	*0 points* Voice is **inappropriate** for the role of the writer and target audience		
	Clarity	*3 points* Writing is **easy to follow** for a non-specialist and is cohesive and connected	*2 points* Writing is **inconsistant**, possible to be followed by someone trained in the field but not a non-speicalist	*0 points* Writing is **confusing** and difficult to understand		
	Structure	*2 points* 300 - 500 words and follows the correct format	*1 points* 300 - 500 words but does not follow the correct format	*0 points* Outside of word limits		
					Total	

Did they attempt to address all of the physics questions (did they include all of the “be sure to’s”)?
This is the one you want to focus your most attention to, as it has to do specifically with the physics. Use this question for yourself when you are assigning points for the “Physical Concepts” category. This category is out of 7 total points. A few examples scenarios are laid out below.

A student addresses...

- Everything in the “be sure to’s” —— 7 points
- Almost everything, but forgets a part of one of the “be sure to’s” —— 6 points
- All but one of the “be sure to’s” —— 5 points
- About half of the “be sure to’s” —— 4 points
- Less than half of the “be sure to’s” —— 3 points

These guidelines will be helpful, and their will always be edge cases. If you are not sure, trust yourself and/or reach out to your Returning M-Write Fellow for any help. Remember, we are **not** grading what they get “right” or “wrong” for the First Draft. However, keep track of things they need to fix so you can include these in your feedback comments!

Did they write in a way that seems appropriate for the situation they are in?
This question will help you assign points for the “Perspective” category of the rubric. This is out of 3 points. We are including this in the grading so that students keep in mind who they are writing to and what role they are in. This is to help make the assignment more “real” in a sense and to have students think carefully about how they are writing their response.

Points will be deducted may be deducted in two ways. One way is if a student **only** uses equations to describe the physics and does not have any written description to go along with it. The second is if a student uses language that is clearly inappropriate for the interaction. The inappropriate language should not be a large factor for most students, but take points off accordingly and consult with your Returning Writing Fellow if need be. Below are some example scenarios.

A student uses...

- Only an equation (without written description with it) to explain a concept once —— 2
- Only an equation (without written description with it) to explain two or more concepts —— 1

Can you understand and follow the physics that they describe in their response?
Thinking about this as you read student’s submissions will help you assign points for the category “Clarity”. This is out of 3 points. This is not about how they write stylistically necessarily, but more so if the physics that they describe can be understood. This could either be a concept that a student explained in a confusing manner, or perhaps the logic and order of ideas seems disjointed. Here are a few example scenarios that will help guide your assignment of points.

- You easily understand whether or not a student is “correct”/“incorrect” —— 3 points
- It takes you awhile to understand what a student is trying to say —— 2 points
- You cannot understand what the student is attempting to explain —— 0 points

These scenarios apply to the response as a whole. If there is just one spot where you are confused, you do not need to deduct points, just mention this in your feedback. However, if there are 2 or more times where you find yourself in the second or third scenario, you should deduct points accordingly.

Did they meet the word requirements and write in an acceptable format?
This question will help you give points for the “Structure” category. This is out of 2 points. As for the word count part, you do not need to be using a word counter on every single response you grade. However, if a student’s submission seems a lot shorter or a lot longer than the others you have read, put the text in a word counter of your choice and check the actual number. If a student is a few words above (525 for example) you do not need to take points off, but make a comment in your feedback. Use the 300 word count as a strict limit, where if you find a student has written less than this, points should be deducted.

The correct “format” is a pretty loose description. Check if they have a “hello” and a “goodbye”, but only take points off if the format is especially strange. If you run across one where you are not sure if the format is wrong, contact your Returning Writing Fellow for a final decision.

Peer Reviews

In grading the Peer Reviews, there are three categories that you must consider. The following three questions will help you assign points to each of the categories.

1. Did the Peer Reviewer answer each question with feedback that applies to that question?
2. Did the Peer Reviewer offer specific suggestions that relate to the submission they reviewed?
3. Did the Peer Reviewer write more than 30 words for each response?

PEER REVIEW RUBRIC

	Strong	Sufficient	Weak	Insufficient	Points awarded
Relevance	*7 points* Answers each rubric question with relative and helpful comments for **all** responses	*6 points* Answers each rubric question with relative and helpful comments for **most** responses	*5 points* Answers each rubric question with relative and helpful comments for **about half of** responses	*0 points* Answers each rubric question with relative and helpful comments for **few to none** responses	
Specificity	*7 points* Answers each rubric question with specific examples from the writing for **all** responses	*6 points* Answers each rubric question with specific examples from the writing for **most** responses	*5 points* Answers each rubric question with specific examples from the writing for **about half of** responses	*0 points* Answers each rubric question with specific examples from the writing for **few to none** responses	
Structure	*6 points* Meets word count for each rubric question in **all** responses	*5 points* Meets word count for each rubric question in **most** responses	*4 points* Meets word count for each rubric question in **about half of** responses	*0 points* Meets word count for each rubric question in **few to none** responses	
				Total	

Did the Peer Reviewer answer each question with feedback that applies to that question?
If a student has responded to each of the Peer Review questions genuinely answering each with comments that apply directly to those questions, then they should receive full points for this. If you notice a pattern of a student not doing this, you may follow the rubric above to deduct points.

The first four/three questions will be concerning the four/three “be sure to’s…”, while the last question has to do with improvements overall that do not have to deal with the “be sure to’s…”

Did the Peer Reviewer offer specific suggestions that relate to the submission they reviewed?
This question is to check if the Peer Reviewer is giving suggestions that would be useful for the student they are reviewing. These could use examples from the student’s draft, or mention specific corrections that they should make.

Take points off if you notice a pattern of unhelpful feedback that is not specific to the student. This indicates that the Peer Reviewer did not read the student’s response closely, if at all.

Did the Peer Reviewer write more than 30 words for each response?
This is not that hard to check either. 30 words is really only about 2 substantive sentences. You do not need to use a word counter for each response, just if it seems like a student really did not write much. This is similar to our other policies surrounding word count. For reference the following feedback is exactly 30 words…

“The author made it clear in which direction the acceleration was, except the sentence about free fall is not imperative to the understanding of the email and can be omitted.”

Late or Missing
As it has been laid out here, grading Peer Review is pretty straight forward. The tricky part is if a student turned in Peer Reviews late or only turned in a portion of their Peer Reviews. For everything M-Write we will be using a policy of late means missing and only making exceptions for extreme circumstances.

If you see that the student you are grading turned in some of their Peer Reviews late or not at all, grade the ones that were turned in on time by the normal scale. Once you have the grade, take a fraction based on how many Peer Reviews were completed on time. If 2/3 Peer Reviews were completed, multiply that fraction by the score you totaled earlier. The process is similar for if they only completed 1/3 Peer Reviews on time.

Revised Draft

While grading the Revised Drafts, there are a few things you should think about. First off, we are now checking for correct and accurate physics. You will be taking points off if they get things wrong. Also, when giving your feedback, you can explicitly tell them what the correct answer is. The following set of questions will help you use the rubric to give scores. These are similar to the First Draft, but are useful so they are repeated here.

1. Did the student correctly address all of the physics questions (did they use correct physics when answering all of the "be sure to's")?
2. Did they write in a way that seems appropriate for the situation they are in?
3. Can you understand and follow the physics that they describe in their response?
4. Did they meet the word requirements and write in an acceptable format?
5. Did the student make appropriate revisions?

REVISED DRAFT RUBRIC

		Strong	Sufficient	Weak	Insufficient	Points awarded
Physics Concepts		*40 points* **Accurately** addresses **all** parts of the prompt using relevant physics concepts	*32 points* **Accurately** addresses **most** parts of the prompt using relevant physics concepts	*24 points* **Accurately** addresses **few** parts of the prompt using relevant physics concepts	*0 points* **Accurately** addresses **no** parts of the prompt using relevant physics concepts	
		Strong	Weak	Insufficient		
Style of Writing	Perspective	*7 points* **Appropriate** voice for the role of the writer and target audeince	*5 points* Voice is **inconsistent** with the role of the writer or target audience	*0 points* Voice is **inappropriate** for the role of the writer and target audience		
	Clarity	*7 points* Writing is **easy to follow** for a non-specialist and is cohesive and connected	*5 points* Writing is **inconsistant**, possible to be followed by someone trained in the feild but not a non-speicalist	*0 points* Writing is **confusing** and difficult to understand		
	Structure	*6 points* 300 - 500 words and follows the correct format	*4 points* 300 - 500 words but does not follow the correct format	*0 points* Outside of word limits		
Revision		*5 points* **Meaningful** revision is made in the physics or writing that shows thought and reflection	*3 points* Revisions made are **unthoughtful** and do not boost the quality of the response	*0 points* **Insufficient** revisions made for a final draft		
					Total	

Did the student correctly address all of the physics questions (did they use correct physics when answering all of the "be sure to's")?

This is the one you want to focus your most attention to, as it has to do specifically with the physics. Use this question for yourself when you are assigning points for the "Physical Concepts" category. This category is out of 40 total points. A few examples scenarios are laid out below.

A student uses correct physics to accurately addresses...

- All of the "be sure to's" —— 40 points
- Almost everything, except part of one of the "be sure to's" —— 33-39 points
- All, but one of the "be sure to's" is missing or incorrect —— 30-32 points
- Only about half of the "be sure to's", with the others missing or incorrect —— 24-30 points
- Less than half of the "be sure to's", with most of them missing or incorrect —— 23 or less points

These point divisions are pretty fluid, so make your best judgement when deciding what to give for a score. Be confident, but reach out for any help.

Did they write in a way that seems appropriate for the situation they are in?
This question will help you assign points for the "Perspective" category of the rubric. This is out of 7 points. We are including this in the grading so that students keep in mind who they are writing to and what role they are in. This is to help make the assignment more "real" in a sense and to have students think carefully about how they are writing their response.

Points will be deducted may be deducted in two ways. One way is if a student **only** uses equations to describe the physics and does not have any written description to go along with it. The second is if a student uses language that is clearly inappropriate for the interaction. The inappropriate language should not be a large factor for most students, but take points off accordingly and consult with your Returning Writing Fellow if need be. Below are some example scenarios.

A student uses only an equation to (without written description with it)...
- To explain a concept once —— 6-7 points
- To explain two or more concepts —— 3-5 points

Can you understand and follow the physics that they describe in their response?
Thinking about this as you read student's submissions will help you assign points for the category "Clarity". This is out of 7 points. This is not about how they write stylistically necessarily, but more so if the physics that they describe can be understood. This could either be a concept that a student explained in a confusing manner, or perhaps the logic and order of ideas seems disjointed. Here are a few example scenarios that will help guide your assignment of points.

- You easily understand whether or not a student is "correct"/"incorrect" —— 7 points
- It takes you awhile to understand what a student is trying to say —— 4-6 points
- You cannot understand what the student is attempting to explain —— 0-3 points

Did they meet the word requirements and write in an acceptable format?
This category will be scored the same as it was in the first draft, only now it is out of 6 points. Reach out if there are any concerns about the structure.

Did the student make appropriate revisions?
This can be tough to judge, but there are tips and tools that you may use to help you assign points for the Revision section (total 5 points). Check to see what comments you originally gave on their draft, this will give you a hint of things that you should check that they changed and fixed. It will be helpful to use the "compare documents" functionality provided by Microsoft Word. This will highlight differences between documents for you. If you want to use this tool, but are having trouble with it, contact your Returning Writing Fellow with questions. "Meaningful revision" really just means did the student make changes that truly impact part of their draft. This can usually be categorized by a 3 or more sentences being different. However, if a student replaced a lot of confusing explanations and made a more clear, concise explanation using only one or two sentences, this is also meaningful revision.

Late Submissions
Accept all late submissions you grade for 50% the grade that you would have given otherwise.

APPENDIX B

Spring Term - Prompt

Article to read:
http://www.mlive.com/business/west-michigan/index.ssf/2011/02/ludington_pumped_storage_plant.html

A Watershed Moment in Energy Storage

You are a consultant for a renewable energy firm here in Ann Arbor. Your firm works with companies such as DTE and Consumers Energy to help create more renewable energy projects around Michigan.

Prompt:
Your boss comes across the following MLive article on the Ludington Pumped (Energy) Storage, and thinks it is a brilliant idea. However, your boss is skeptical that the fundamental physics hold up, and they want you to walk them through all of the different energy transformations that occur in this process (i.e. where the energy initially comes from, how it is stored, and how it is transformed back into energy that can be used by the consumers). Your boss also wonders if they'd be able to recover all of the energy they started with, and asks you to conclude with your personal judgement - is this a sensible way for Michigan utility companies to store energy?

In your report be sure to:

- Use the concept of gravitational potential energy to explain how elevating water would store energy.
- Describe the processes of energy transformation involved in raising the water up.
- Describe the processes of energy transformation involved in recovering the stored energy.
- Close your report with a short description of your personal judgement. Is this a sensible way for Michigan utility companies to store energy?

Rubric:

- Did they successfully describe gravitational potential energy and how how elevating water uphill would store energy?
- Did they include an accurate description of the energy transformation involved in raising the water up?
- Did they include an accurate description of the energy transformations involved in recovering the energy stored?
- Did they address whether all the energy could be recovered?
- Did they include a judgement about the sensibility of the energy storage method?

Peer review guidelines:

- Print and read over your peer's essay to quickly get an overview of the piece.
- Read the essay more slowly with the rubric in mind.
- Highlight pieces of texts that let you directly address the rubric prompts in your online responses.
- In your online responses, focus on larger issues (higher order concerns) of content and argument rather than grammar and spelling.
- Be very specific in your responses, referring to your peer's actual language, mentioning terms and concepts that are either present or missing, and following the directions in the rubric.
- Use respectful language whether you are suggesting improvements to or praising you peer.

Revision Prompt:

- Re-read the prompt and rubric to think about what an effective response should include.
- Read feedbacks you received from your peers.
- Keep these in mind, re-read your draft and highlight places where you can improve
- Revise and submit your response

APPENDIX C

Prompt

MWrite Prompt 2: *A Watershed Moment in Energy Storage*

You are a consultant working at a renewable energy consulting firm here in Ann Arbor. Your firm works with companies such as DTE and Consumers Energy to help create more renewable and sustainable energy projects around Michigan.

Your boss, Mr. Green, is a marketing wiz, but is not always up to speed on the technical side of the projects. He is on a visit of the Ludington Pumped Storage Plant. He sends you an email saying that it is going well and shares with you an MLive article from 2011, discussing the 6-year upgrade project that increased the efficiency of the plant (see attached pages if link does not work). Mr. Green is pretty sure that your firm would be able to oversee and advise on the construction of a project that would be very similar to the Ludington Pumped Storage Plant, but he is unsure on how the system completely works.

Later that day, you get a call from Mr. Green.

> *"Hey, this place is really cool, we should definitely send you out here sometime. Anyway, the reason I called is because while the team here has been great at going over the logistics and, like, the financial aspects of this place, they haven't really gotten all that technical with me, maybe they don't think I'll understand?*
>
> *Who knows, but I was wondering if you could have a memo, or something, memo would probably be good. If you could have a memo ready for me when I get back tomorrow. Could ya go over some of the physics behind this sort of facility? I really want to know where this energy is coming from, I mean, I know it's off the grid, but thats electricity. How is that supposed to lift the water in the first place, and where does it go after? Is this the 'moving between systems' thing you were talking about the other day?*
>
> *They keep telling me that the energy is stored in this water and the energy system, but what does that even mean? What is the system they are talking about? Then they said the energy could be uh, I think they said 'lost'? How does this happen, does it like, leak out of the water? I've never seen energy leaking out of water. And lost? Where does it go? Last week you were going on and on about how energy couldn't be created or destroyed, but they are telling me they lost this energy and its now gone? Were you serious about that energy conservation stuff, what's going on here?*
>
> *I do really like this place though, it's really nice. Do you think that this is something that is uh, that is sensible for state of Michigan to store energy elsewhere in the state? Where do you think these kind of facilities would work the best?*
>
> *Hope that's not too much. I'd really appreciate it if you could do this for me."*

You reply that you will be able to draft a memo that answers all of his questions about the process while explaining the physics behind it. Mr. Green thanks you and hangs up the phone.

In your report be sure to…

- Define the system that the energy is flowing into and out of
- Describe the energy transfers that occur while the water is moving in the system
- Address Mr. Green's concerns about the 'loss' of energy, explain how this is consistent with the Conservation of Energy
- Close your report with your judgement of whether this type of project is physically and ecologically sound and where these projects would work best.

Keep in mind…

- Make sure to convey your physics understanding clearly and precisely
- Do not include your name (to preserve anonymity for peer reviews)
- Stay within 300-500 words
- Write your response to resemble the body of a memo
- Use an appropriate tone of voice for your role and audience
- Outside sources are not needed, but if used, please cite them in a citation style of your choice

Ludington Pumped Storage Plant to receive $800 million upgrade over six years[1]

Updated February 7, 2011 at 6:07 PM; Posted February 7, 2011 at 3:32 PM

By Dave Alexander
LUDINGTON — Consumers Energy and Detroit Edison will invest $800 million over six years to upgrade and lengthen the life of the Ludington Pumped Storage Facility.

The Ludington Pumped Storage Plant has six huge turbine generators that will be replaced in the project announced Monday by Consumers Energy and Detroit Edison.
The innovative electric "battery" built in 1973 on a bluff overlooking Lake Michigan south of Ludington is a key component for the public utilities as they move into alternative sources of energy, such as wind farms.

Replacement of the plant's six massive turbines will increase the plant capacity by 16 percent — from the current 1,872 megawatts to 2,172 megawatts, after the replacements are installed by 2019. The construction project is estimated to create 100 jobs during the six years, utility officials said in a public announcement at the Ludington-Scottville Chamber of Commerce offices Monday afternoon.

"The upgrades will improve its efficiency, increase its role in support of clean-energy sources for Michigan, create jobs and ensure that the plant will continue to contribute to the economy of the Ludington area and Michigan for many more decades," Detroit Edison President Steve Kurmas said in a prepared statement.

Detroit Edison owns 49 percent of the plant and the Jackson-based Consumers Energy — the public electric utility for most of western Michigan — owns the other 51 percent. The Pumped Storage Plant was built to provide a low-cost, reliable source of electricity for Michigan customers during high electric demand times.

"The Ludington Pumped Storage Plant has proven its value over several decades of service, providing millions of Michigan electric customers with outstanding performance and dependable reliability," Consumers President and CEO John Russell said in a prepared statement. "This major investment is a sign of our commitment to Michigan's economic development and points to the state's future growth."

When combined with the planned $232 million, 56-turbine Lake Winds Energy Park near the Pumped Storage Plant, Consumers Energy would be investing more than $1 billion in Mason County. The Lake Winds Energy Park still needs a critical Mason County zoning approval as strong opposition to the wind farm has formed in recent weeks.

The facility includes a 842-acre reservoir sitting atop a bluff overlooking Lake Michigan.

The Ludington Pumped Storage Plant is a 1,000-acre site four miles south of the city of Ludington. The facility includes a 842-acre reservoir perched atop the bluff that is able to hold 27 billion gallons of Lake Michigan water.

When electric demand is low and the electric rates are cheaper, such as during the overnight hours, lake water is pumped 372 feet up to the reservoir. When electric demand is high and rates increase during the day, the water is released back down to Lake Michigan to produce electricity that has been “stored” in the reservoir like a giant battery.

When the plant opened, it was named one of the Top 10 civil engineering projects of the 20th Century in Michigan by the state section of the American Society of Civil Engineers.

The Ludington Pumped Storage Plant is able to produce enough electrical power to supply a community of 1.4 million people, Consumers Energy officials say. The upgrade will increase the number of people served by the plant to 1.65 million, officials said.

The Ludington plant plays an increasingly important role as a storage facility for renewable energy produced during off-peak periods, thereby making renewable energy more affordable and reliable, Consumers officials said.

Not only has the utility looked at Mason County for a land-based wind farm, but Scandia Offshore Wind has explored a huge Lake Michigan wind farm for the waters off Mason and Oceana counties due to the proximity of the Pumped Storage Plant. Both county boards of commissioners have rejected the Scandia plans.

In addition to the electricity, the plant generates $10 million a year in property taxes for area governments and schools.

The Ludington maintenance and upgrade project is scheduled to begin in 2013. It will need 100 workers each year in the construction trades, including electricians, welders, crane operators, pipefitters, millwrights and carpenters, company officials said.

The plant upgrade includes installation of six new turbines from the Toshiba International Corp. of Japan, a global leader in electrical generation technologies. Toshiba won the turbine contract through a competitive bid, Consumers officials said.

Consumers Energy and Detroit Edison are both publicly traded companies and regulated utilities in the state of Michigan. Consumers Energy is the principal subsidiary of CMS Energy, providing power to 1.8 million customers in 68 Lower Michigan counties, while Detroit Edison is the subsidiary of DTE Energy, providing electricity to 2.1 million customers in Southeast Michigan.

[1]Alexander, D. (2011, February 07). Ludington Pumped Storage Plant to receive $800 million upgrade over six years. Retrieved from https://www.mlive.com/business/west-michigan/index.ssf/2011/02/ludington_pumped_storage_plant.html

BIBLIOGRAPHY

BIBLIOGRAPHY

Ackerman, J. M. (1993), The promise of writing to learn, *Written communication*, *10*(3), 334–370.

Atasoy, Ş. (2013), Effect of writing-to-learn strategy on undergraduates conceptual understanding of electrostatics, *The Asia-Pacific Education Researcher*, *22*(4), 593–602.

Bangert-Drowns, R. L., M. M. Hurley, and B. Wilkinson (2004), The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis, *Review of educational research*, *74*(1), 29–58.

Chabay, R. W., and B. A. Sherwood (2015), *Matter and interactions*, John Wiley & Sons.

Cho, K., and C. MacArthur (2010), Student revision with peer and expert reviewing, *Learning and instruction*, *20*(4), 328–338.

Cummings, K., and M. Murphy (2007), The effectiveness of incorporating conceptual writing assignments into physics instruction, in *AIP Conference Proceedings*, vol. 883, pp. 61–64, AIP.

Emig, J. (1977), Writing as a mode of learning, *College composition and communication*, *28*(2), 122–128.

French, T., A. Sanchez, L. M. Brousil, and E. Balison (2015), Compartmentalization of energy concepts–definitions, ontologies, and word associations, PER Conference, pp. 115–118, College Park, MD.

Gere, A. R., N. Limlamai, E. Wilson, K. MacDougall Saylor, and R. Pugh (2019), Writing and conceptual learning in science: An analysis of assignments, *Written Communication*, *36*(1), 99–135.

Goodhew, L. M., and A. Robertson (2014), Investigating the proposed affordances and limitations of the substance metaphor for energy, PER Conference, pp. 103–106, Minneapolis, MN.

Gray, K., and R. Scherr (2016), Drawing energy: Evidence of next generation science standards for energy in diagrams, PER Conference, pp. 128–131, Sacramento, CA.

Halim, A. S., S. A. Finkenstaedt-Quinn, L. J. Olsen, A. R. Gere, and G. V. Shultz (2018), Identifying and remediating student misconceptions in introductory biology via writing-to-learn assignments and peer review, *CBELife Sciences Education*, *17*(2), ar28.

Knight, R. D. (2017), *Physics for scientists and engineers*, Pearson Higher Ed.

Lindsey, B. A., P. R. L. Heron, and P. S. Shaffer (2012), Student understanding of energy: Difficulties related to systems, *American Journal of Physics*, *80*(2), 154–163, doi:10.1119/1.3660661.

McIldowie, E. (1995), Energy transfer-where did we go wrong?, *Physics Education*, *30*(4), 228–230, doi:10.1088/0031-9120/30/4/006.

Moon, A., E. Zotos, S. Finkenstaedt-Quinn, A. R. Gere, and G. Shultz (2018), Investigation of the role of writing-to-learn in promoting student understanding of light–matter interactions, *Chemistry Education Research and Practice*, *19*(3), 807–818.

Patchan, M. M., D. Charney, and C. D. Schunn (2009), A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor, *Journal of Writing Research*, *1*(2), 124–152.

Pennebaker, J. W., R. Boyd, K. Jordan, and K. Blackburn (2015), The development and psychometric properties of liwc2015, *Austin, TX: University of Texas at Austin*.

Prain, V., and B. Hand (2016), Coming to know more through and from writing, *Educational Researcher*, *45*(7), 430–434.

Price, E., F. Goldberg, S. Patterson, and P. Heft (2013), Supporting scientific writing and evaluation in a conceptual physics course with calibrated peer review, *AIP Conference Proceedings*, *1513*(1), 318–321, doi:10.1063/1.4789716.

Rivard, L. O. P. (1994), A review of writing to learn in science: Implications for practice and research, *Journal of Research in Science Teaching*, *31*(9), 969–983.

Sabo, H. C., L. M. Goodhew, and A. D. Robertson (2016), University student conceptual resources for understanding energy, *Phys. Rev. Phys. Educ. Res.*, *12*, 010,126, doi:10.1103/PhysRevPhysEducRes.12.010126.

Scherr, R., H. Close, and S. B. McKagan (2011), Intuitive ontologies for energy in physics, pp. 343–346, Omaha, Nebraska.

Scherr, R. E., H. G. Close, E. W. Close, and S. Vokos (2012), Representing energy. ii. energy tracking representations, *Phys. Rev. ST Phys. Educ. Res.*, *8*, 020,115, doi:10.1103/PhysRevSTPER.8.020115.

Young, H. D., R. A. Freedman, T. Sandin, and A. L. Ford (1996), *University physics*, vol. 9, Addison-Wesley Reading, MA.