

Investigating the mechanisms of peer review

Robert Dalka and Timothy McKay

Department of Physics, University of Michigan, 450 Church St, Ann Arbor, MI 48109

As a part of a larger study of Writing-to-learn (WTL) in introductory physics, we present the effects of the peer review process on student understanding of energy systems. This study looks at student writing from two consecutive semesters, Fall 2018 with 670 students, and Winter 2019 with 603 students. In one of our three WTL activities, students were asked to list the entities that would be included in the energy system of a pumped water storage facility. In our analysis, we looked at both the peer feedback that a student received and the peer work which they read. From the Fall 2018 results, we have found that the students revise their energy systems more often after receiving a peer review comment advising a change to the energy system. We have further found that the written feedback students receive from members of the instructional team directly correlates with *how* students revise their energy system. However, the analysis of our Winter 2019 data does not support these same conclusions. Our results present the complex, but significant, role that a peer review process plays in building students' understanding of energy systems.

I. INTRODUCTION

Writing-to-learn (WTL) is based in cognitive theory [1]. The practice of it is believed to uniquely serve learning as writing naturally follows the structure of recollection, synthesis, and revision [2]. Through the years, the way that WTL activities have been implemented, their effectiveness and the critiques of them have changed.

Gere et al. recently evaluated the current landscape of WTL activities in the sciences [3]. 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 are meaning-making writing tasks, interactive writing processes, clear writing expectations and metacognition.

In this study, our focus is placed on the process of peer review, or the *interactive writing processes* component identified by Gere et al. 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 [4], [5]. Specific to a physics context, it has been shown that students are able to evaluate their peers' work to the same degree as experts [6]. 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 were introduced into students' work [7]. This study focused deeply on identifying specific misconceptions, and how those misconceptions were affected through the feedback written. Yet, other inputs a student may receive during peer review were not taken into account, such as first drafts that students read.

The literature surrounding WTL in physics is scarce. Out of the 46 total studies that were synthesized in the report by Gere et al., only nine were identified with physics as the primary discipline and only six of those nine were conducted in a university-level environment. Along with studying what happened in our course, the purpose of this paper is, in part, to bolster the literature surrounding peer review in WTL activities, especially in physics.

A. Our Course Context

Over two recent academic terms, Fall 2018 and Winter 2019, the instructional team for the calculus based Introduction to Mechanics course (Physics 140) at the University of Michigan, has implemented WTL activities. 670 students were enrolled in the Fall 2018 term and 603 students were enrolled in the Winter 2019 term. Three sections were offered each semester, two taught by a lecturer and one by a tenured faculty member. Both instructors work collaboratively as part of a larger instructional team, using the same evaluative tools and coordinating lectures.

There were two motivating factors for using WTL activities in our course. First was to engage students in a mode

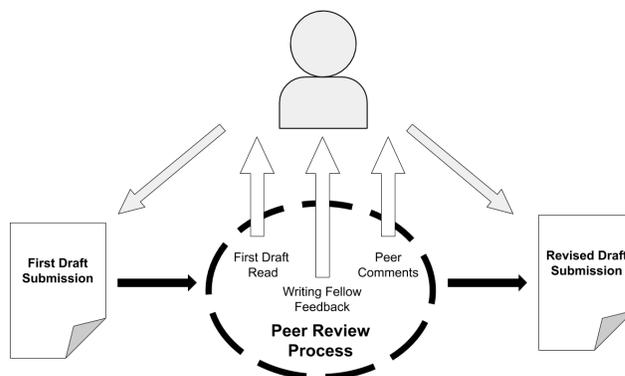


FIG. 1. Figure above shows the process used in our WTL activities. A student is shown writing the First and Revised Drafts, along with receiving all the inputs given during the Peer Review process.

of learning they may not experience otherwise in the course. Our second motive was to give the instructional team a different perspective of student knowledge, which they could use to pinpoint the learning goals students were struggling to reach.

The implementation of WTL activities into our large, foundational course was facilitated by a socio-technical system called M-Write. The M-Write project works to bring WTL pedagogy into many STEM courses through working with faculty to design and implement discipline and topic specific activities [8]. The WTL pedagogy that M-Write promotes includes the four components identified by Gere et al. First, students receive a prompt that places them in a tangible scenario in which they must write about a specific topic to an identified audience. In our course, students had a week to write a 300 - 500 word first draft. Next, each student is responsible for reading three anonymous first drafts and responding to specific questions that guide their written feedback. The students have four days to complete the peer review. Finally, the students must use the feedback they received and submit a revised version of their response for final grading. Students are given three days to make these revisions. These three steps will be referred to as the *first draft*, *peer review*, and *revised draft*.

This process is overseen by a group of near-peer mentors, called *writing fellows*. The writing fellows are responsible for grading all three steps, holding office hours, and providing their own feedback during the peer review stage. They are undergraduates who have previously taken the course and are trained by WTL pedagogy experts who are a part of the M-Write team. These writing fellows were essential to the process we had set up for these activities. Along with their responsibilities with the students, our writing fellows provided feedback on prompts to the instructional team before they went out to students. They also helped identify patterns in student responses that informed investigations such as the one presented in this paper.

B. The Writing Prompt and Learning Goal

Each WTL activity begins with students receiving a prompt. In the specific activity discussed in this paper, the prompt places the student in the role of a consultant working at a renewable energy firm. The prompt describes a frantic phone call from their boss, who discusses a pumped water energy storage facility in Ludington, MI. The boss tasks the student with writing a memo explaining the physics of how the energy is stored. The learning goal for this prompt is “Understanding the importance of defining an energy system and how to choose the entities to include in it.” For the specific scenario described in the prompt, we expected the students to reach the explanation that in order to discuss Potential Energy, one needed to include both the water being pumped up and the Earth in their system as the Potential Energy is stored within the interaction.

We received a wide variety of students’ descriptions of the energy system and the entity they included. These were first noticed by the writing fellows in the Fall 2018 semester while reading students’ first draft submissions. This prompt showed us that confusion existed for students concerning energy systems and the entities to talk about in their system. This is why we decided to focus on this prompt in our study. We have looked at data from both Fall 2018 and the following semester, Winter 2019, to check if the results of the activity are consistent.

II. METHODS OF ANALYSIS

To understand how the peer review process influences students’ revision of their described energy system, we first defined what measurable inputs a student might receive during the peer review period. These inputs are first drafts that students were assigned to read, peer review comments they receive written by different students, and feedback written by the writing fellows, who are a part of the instructional team. To study how these inputs to students correlate to the way they define their energy system, we classified how they wrote about the system in their first and revised drafts.

All first and revised drafts from both semesters were deductively categorized into groups based upon the system that was described by the student. These groups emerged as follows; Earth and Water, Earth and Plant and Water, Plant and Water, Water, Plant, Plant and Water and Electrical Grid, No System, and Not Submitted. Once these groups were created, we saw a clear split between drafts in which students had described a system where Potential Energy was held within the interaction of entities and systems where Potential Energy was held within the entities themselves. Therefore, the categories of *Earth* and *No Earth* were the dominant way we grouped student responses.

Next, we gathered and read all of the peer reviews. Feedback from students was guided by specific questions, so we isolated the student responses to the question that directed

them to evaluate the energy system. Each peer review comment was labeled as either in support of the described system, or not. If the comment was not in support, it was also labeled with the alternative system presented. In a similar process, the writing fellow comments were examined as students would receive this feedback at the same time as their peer reviews. These comments were categorized as feedback that either directed students to include the Earth in their system, directed students to generally rethink their system, or did not touch on the energy system at all. To make the naming of inputs names, we have a shorthand for our terms. Peer reviews suggesting revision with alternative system (PR-S), the reading of a peer’s first draft system that is different than the student’s own system (Read-S), and a writing fellow comment that suggests that the student should revise their energy system (WF-S).

From here on, only students who had completed both a first and revised draft were analyzed. Anyone who did not submit both of these parts was dropped from the pool. This left 620 individual students from the Fall 2018 term and 514 students from the Winter 2019 term.

We identified two questions to direct our investigation of this complex process. Both questions concerned the students who did not include the Earth in their first draft energy system. The second concerned those students who changed their energy system in the revised draft and specifically focuses on investigating whether or not students included the Earth because it changes the way Potential Energy can be talked about in a system. The questions are,

- What inputs significantly correlate to students revising their energy system?
- What inputs significantly correlate to students revising their energy system to include the Earth?

To address the first question, we looked at each student and the measurable inputs they received from the peer review process. Then found whether they fell into the *revised system* or *remained system* groups. From there, we used a two-sided t-test for the null hypothesis that the two groups of *revised system* and *remained system* had consistent distributions of students who had received the same inputs. The significant inputs were determined by the resulting p-value using a three tier α of 0.05, 0.01, and 0.001. This process was then repeated for the second question with groups of *revised Earth system* and *revised not Earth system*.

III. RESULTS

In the Fall 2018 term, 79 students included the Earth in their first draft system. The 541 students who did not include the Earth varied in their revised draft. 186 of these students did not revise at all, 203 of these students revised their system, but did not include the Earth, and only 152 students revised to include the Earth in their system. In the Winter 2019 term, more students included the Earth in their first draft system with 114 students doing so. Of the 400 students who did not

include the Earth in their first draft, 99 of them did not revise their system, 86 of these students did revise their system, but did not include the Earth. However, in stark contrast to the previous semester, 215 of these students did revise to include the Earth in their system.

In the Fall 2018 semester, students who did revise, received an average of 1.7 peer reviews advising them to revise. This is significantly different (p-value < 0.001) when compared to students who did not revise, who received an average of 0.93 peer reviews advising them to revise. Students who revised also read an average of 2.4 first drafts with systems different than their own, which is significantly different (p-value < 0.001) compared to the students who did not revise. Those students read an average of 2.1 first drafts with different systems. The students who revised received more peer reviews telling them to change their system and read more first drafts that explained different systems than their own.

In the Winter 2019 semester, the students who did revise received an average of 1.6 peer reviews telling them to change. This was not significantly different than students who did not revise, who received an average of 1.4 peer reviews that advised them to revise. Students who revised also read an average of 2.3 first drafts with systems different than their own, which was significantly different from those students who did not revise (p-value = 0.039). For the students who remained with their original system, the average number of first drafts they read with different systems was 2.1. These are similar results to the Fall 2018 term, however, the differences between those that revised and those that did not are not as significant.

Many inputs significantly correlate to whether or not a student will revise their system for the Fall 2018 term, as can be seen in Table I. The most significant correlations appear to be how many peer review comments a student receives advising them to revise. If the student had received no peer reviews telling them to rethink their system, most of these students remained with their system. However, if a student received more peer reviews advising them to change, the student was more likely to revise their system. A similar pattern emerges for the students who read a system different than their own, however, this is not as correlated to whether or not they revised as the peer review comments were. In the Winter 2019 term, none of the inputs measured had a significant correlation with the distribution of students who revised or remained.

Although the Winter 2019 data returns a null result for the correlation of inputs to if students revised, the Fall 2018 term paints a complex picture. One in which the reading of different systems might also interact with receiving peer review comments. We took a closer look at these inputs and analyzed students grouped by the cross section of reviews received and first drafts read (Table II). If a student did not receive a peer review advising them to revise, they are more likely to remain with their system no matter how many times they read a first draft system different than their own. The input categories that correspond significantly to more likelihood of revision are when a student receives both two or three peer reviews advising revision and reads two or three first drafts with sys-

Input Type	Revised	Remained	p-value
Fall 2018			
(N = 355) (N = 186)			
PR-S = 0	52	74	<0.001***
PR-S = 1	95	68	0.018*
PR-S = 2	115	27	<0.001***
PR-S = 3	93	17	<0.001***
Read-S = 0	5	8	0.037*
Read-S = 1	41	32	0.068
Read-S = 2	121	74	0.19
Read-S = 3	188	72	0.0016**
WF-S = 1	319	156	0.043*
Winter 2019			
(N = 301) (N = 99)			
PR-S = 0	50	20	0.42
PR-S = 1	93	36	0.31
PR-S = 2	98	30	0.68
PR-S = 3	60	13	0.13
Read-S = 0	8	6	0.11
Read-S = 1	39	18	0.19
Read-S = 2	96	31	0.91
Read-S = 3	158	44	0.17
WF-S = 1	286	90	0.14

TABLE I. This table presents the number of students who experienced different inputs during the peer review process. These students included in this table did not have the Earth in their first draft system.

tems different than their own.

Following through to our second question of what inputs correlate to a student including the Earth into their system, the results are a little less complicated. In both semesters, the overwhelming factor was whether or not the writing fellow specifically mentioned the Earth in their feedback to the students. In Fall 2018, there were 152 students who revised to add the Earth and 203 students who revised without the Earth. A total of 200 of the students who revised received a writing fellow comment mentioning the Earth, with 109 of these students revising to add the Earth to their system. The distribution of students receiving a writing fellow comment mentioning the Earth is significantly correlated, with a p-value less than 0.001, to students including the Earth in their revised system. In Winter 2019, there were 215 students who revised to add the Earth and 86 students who revised without the Earth. A total of 200 of these students received a writing fellow comment mentioning the Earth, with 157 of these students revising to add the Earth to their system. This distribution is also highly significant, with a p-value less than 0.001. The input that students receive from other students may be enough to encourage them to revise their system, but

	PR-S = 0			PR-S = 1			PR-S = 2			PR-S = 3		
	Revised	Remained	p-value	Revised	Remained	p-value	Revised	Remained	p-value	Revised	Remained	p-value
Read-S = 0	1	4	0.031*	2	4	0.094	1	0	0.47	1	0	0.47
Read-S = 1	11	18	0.0012**	15	10	0.55	11	4	0.52	4	0	0.15
Read-S = 2	21	31	<0.001***	32	28	0.034*	36	8	0.018*	32	7	0.025*
Read-S = 3	19	21	0.012*	46	26	0.74	67	15	<0.001***	56	10	<0.001***

TABLE II. The population of students in this table are those who did not have the Earth in their first draft system in Fall 2018. Moving horizontally represents the number of peer reviews that a student received advising them to revise their system. Moving down the table represents the number of different systems that students read in the first drafts. Each cell holds the number of students who Revised or Remained and the p-value when comparing the two groups.

the highest correlated input to whether or not they include the Earth in writing fellow feedback.

IV. DISCUSSION

Although there were two instructors involved in this study, we treated each semester as one group of students. We see this as valid because all students were evaluated across the sections through the same methods, each writing fellow was responsible for students from all sections, and the work for these activities was completed outside of lecture. A more significant limitation on our results is that, on occasion, a student did not receive all three peer reviews. As well as when students are given incorrect feedback that leads them down a wrong path in their revision. In a peer review process, this is expected, and is why we ensured that all students received feedback from a writing fellow.

In both the Fall 2018 and Winter 2019 terms, a majority of students revised their energy system. In this investigation, we have identified three measurable inputs that students experience during the peer review process. These being, comments from peers, readings of other students' first drafts, and feedback written by writing fellows. Without direct knowledge of what the students thought processes were, we cannot know completely what influenced them. We instead have presented correlated factors that lead to common outcomes.

In the Fall 2018 term, a student receiving no peer comments telling them to revise was strongly correlated to a student remaining with their system. This is true no matter how many different systems a student read. So while students who revised read a first draft that had a different system more often than received a peer review comment telling them to change, the peer review comments are more strongly correlated to a student revising their energy system. The strongest combination of inputs that was correlated to a student revising their system was when a student would receive both two or more peer reviews advising them to revise and that they read two or more first drafts with systems different than their own.

The Winter 2019 results do not offer these same conclusions. In this term, no input factor was significantly correlated to if a student revised their system. Similar trends to

the Fall 2018 term appear, however, these trends do not represent any significant or dominating pattern. This null result shows how important outside, currently unmeasured, inputs may be. Future studies could be done to perhaps understand what outside influence to revision do exist. In future studies, a reflection assignment or student interviews about their thought process and reasons for revision would be helpful in determining currently unmeasured factors in revision.

Our investigation into the second question yielded more direct results. When looking at those students who did revise their energy system, we see that what the writing fellows suggest in their feedback correlates to whether or not a student includes the Earth in their revised draft system. This correlation was significant in both Fall 2018 and Winter 2019 terms, providing evidence that feedback from writing fellows may have a strong influence on how students revise. This result makes sense when placing it in the context of our situation. Since there was such a wide variety of systems, it is likely that the peers a student receives input from will have a different system than them. These inputs, at least in the Fall 2018 term, are correlated with revision of the system. However, since not many students included the Earth in their initial system, a particular student would not be exposed to that system from their peers. Students may look towards the writing fellow as the deciding factor for *how* they should revise, if they have already decided to revise. The writing fellows have thus proven to be essential to how we have conducted our WTL activities. If one wanted to implement these WTL activities without the writing fellow input, steps would need to be taken to better leverage the other types of input students receive to make up for the writing fellow influence.

ACKNOWLEDGMENTS

Thank you to the Physics 140 instructional team for the support needed to implement these activities. Thank you to the M-Write team for their help in developing the prompts and providing writing fellow training. Finally, thank you to our writing fellows who have made the process of conducting these activities possible.

-
- [1] L.S. Vygotsky, Thought and language, in *Annals of Dyslexia* **14**, 1 (1964).
- [2] J. Emig, Writing as a mode of learning, in *College composition and communication* **28**, 2 (1977).
- [3] A.R. Gere, N. Limlamai, E. Wilson, and S. MacDougall, K. Saylor, and R. Pugh, Writing and Conceptual Learning in Science: An Analysis of Assignments, in *Written Communication* **36**, 1 (2019).
- [4] M.M. Patchan, D. Charney, and C.D. Schunn, A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor, in *Journal of Writing Research* **1**, 2 (2009).
- [5] K. Cho and C. MacArthur, Student revision with peer and expert reviewing, in *Learning and instruction* **20**, 4 (2010).
- [6] E. Price, F. Goldberg, S. Patterson, and P. Heft, Supporting scientific writing and evaluation in a conceptual physics course with calibrated peer review, in *AIP Conference Proceedings* **1513** 1 (2013).
- [7] A.S. Halim, S.A. Finkenstaedt-Quinn, L.J. Olsen, A.R. Gere, and G.V. Shultz, Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review, in *CBE-Life Sciences Education* **17** 2 (2018).
- [8] <https://lsa.umich.edu/sweetland/m-write.html>. Retrieved 6/12/2019.