Writing-to-Learn in introductory Materials Science and Engineering

L. Marks[1,2], H. Lu[1,2], T. Chambers[1], S. Finkenstaedt-Quinn[3], and R.S. Goldman[1,4]

[1]Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, U.S.A.

[2]Sweetland Center for Writing, University of Michigan, Ann Arbor, MI 48109, U.S.A.

[3]Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, U.S.A.

[4]Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A.

Sunday, May 2, 2021

Abstract

We present a study on the impact of Writing-to-Learn (WTL) assignments on student learning in introductory materials science/engineering. WTL promotes deeper thinking by asking students to address "real-world" situations via writing. The inclusion of peer review and revision processes in the WTL assignments allows students to give and receive feedback and critically assess their work. Through analysis of writing products and using pre/post assessments, we examine student gains in conceptual understanding and critical reasoning. Gain distributions across topics suggest that highly-effective assignments require students to map between qualitative and quantitative representations of phenomena and to connect their microscopic and macroscopic understandings.

Keywords: Education, crystallographic structure, corrosion, phase transformation, polymers

Introduction

Writing has long been used to support learning across a range of contexts and disciplines.[1,2,3] One such writing-based instructional practice, Writing-to-Learn (WTL), has been incorporated into classrooms in a number of forms—spanning reflective writing to long, scaffolded writing assignments—and used to support a number of instructional goals—from developing disciplinary thinking to conceptual learning.[4,5] Within science education, WTL assignments have been used to support students' development of scientific argumentation, metacognition, and conceptual understanding.[5,6,7] These goals are also represented in the WTL assignments described in the engineering education literature.[8,9,10,11] However, only a few studies of WTL in materials science have been reported to date.[10, 16] In one case, the effect of shorter in-class writing assignments on student learning within an introductory materials science course was explored.[10] In another case, we examined student responses to a context-based WTL assignment that consisted of a draft, peer review, and revision cycle, emphasizing its usage and efficacy of supporting conceptual learning of polymer properties within an introductory materials science course.[16] Here, we expand upon our prior work by considering student responses across a comprehensive set of WTL assignments spanning materials classes and functionalities to gain insight into student gains across a semester and inform future use of WTL in introductory materials science and engineering.

For these studies, we utilize a WTL process in which students apply content knowledge to "real world" situations by writing a response to an authentic scenario, performing content-focused peer review, and finally revising their initial response.[12] This WTL process incorporates the key elements for effective WTL assignments identified by meta-analyses of the WTL literature, namely clearly-defined and interactive writing expectations that incorporate meaning-making tasks and support metacognition.[6,] **Error! Bookmark not defined.**[,] **Error! Bookmark not defined.** This WTL process also aligns with cognitive theories of learning such as social constructivism,[1,6,10, 13 , 14] which posits that students learn within their individual social environments by restructuring existing knowledge to incorporate new knowledge.[13,14] Indeed, research has shown that this WTL assignment process has enabled students to constructively engage with the peer review and revision processes, thereby supporting them in learning challenging content in a wide range of introductory STEM courses, including biology, chemistry, and statistics.[15,16, 17,18,19,20,21]

The core objective of the introductory materials science and engineering course is to introduce the principles of engineering materials, with an emphasis on fundamental relationships between internal structure, properties, processing, and performance of materials that are essential for understanding the role of materials in the design of engineering systems. A second objective is to introduce materials classes (metals, ceramics, polymers, semiconductors, and composites) and their distinctive chemistry and internal structure. A third objective is to illustrate the role of thermodynamics (via phase diagrams) and kinetics (via diffusion) to the design of materials. The final objective is to introduce materials functionality, with an emphasis on understanding connections between internal structure (microstructure and defects) and macroscopic properties/performance.

Here, we examine the influence of WTL assignments that incorporate an authentic context and social elements, in the form of peer review and revision, on student understanding of key concepts in introductory materials science. This study is guided by the following research questions:

1. Do students' descriptions of the WTL-assignment-targeted content improve between their drafts and revisions?
2. Do students develop more robust understandings of the WTL-assignment-targeted content?
3. Which learning goals are best supported by the WTL assignment design?

Methods

Implementation

Data in this study came from two main sources: students' written responses to the WTL assignments, which were quantified using a rubric generated by the research team, and students' responses to concept-inventory-style assessment external to required coursework. Quantitative analysis of student writing was used to examine whether peer review and revision contributed to improvements in students' conceptual descriptions. Quantitative analysis of students' external assessment question responses were used to compare the learning gains of students under two conditions: students who participated in WTL and students who participated in a guided group discussion.

This work was performed at a midwestern university in a lower-level Materials Science Engineering (MSE) course during four separate semesters. The course consisted of lecture and recitation sections with coursework including traditional problem sets, bi-weekly reflective writing, and WTL assignments. The prerequisite was either general chemistry or introductory organic chemistry. The textbook for the course was “Materials Science and Engineering” by Callister and Rethwisch.[22] Across the four semesters, the course participants consisted of 151 students who ranged from sophomores to seniors, as well as two graduate students auditing the course. The students were primarily affiliated with the College of Engineering, over half with intended majors in Biomedical Engineering. Amongst the 120 students who completed the WTL assignments and the pre and post external assessments, 38 self-identified as female, 23 as non-US born, and 13 as first generation college students.

The WTL assignments focused on aspects of course content known to be conceptually challenging for students.[23,24] Each WTL assignment consisted of a written response to a prompt, anonymous open response peer review performed by the students, and revision of their original draft. Peer review was guided by a content-focused rubric and students did not score one another’s writing. The writing assignment prompts are included in the supporting information (SI - Writing Assignments). Students had one week to write their initial response and half a week for both peer review and revision, respectively. For the initial draft and peer review, scores were based upon completion, with a cursory check to verify consideration of all the prompt requirements . For the final draft, scores were based upon alignment of student responses with rubric criteria. Additional support for students was provided by two peer tutors (“Writing Fellows”) who had previously taken the course. The writing fellows were trained to help students approach the writing assignments and learn content. Throughout the semester, writing fellows were available to facilitate peer review and answer both content and writing questions regarding the WTL assignments. Our WTL implementation also follows the five guidelines for designing effective “writing to communicate” experiences in engineering classes, proposed by Dr. Suzan Lord.[25]

Writing Analysis

We analyzed students’ draft and revision submissions of four WTL assignments using carefully-designed rubrics. Each assignment had its own unique rubric, designed to evaluate students’ conceptual understanding by probing their ability to describe course content relating to the assignment’s topic. Each rubric had at least three criteria, with the first criterion assessing students’ ability to write at a level that understandable to an audience with minimal scientific background. Characteristics of scores 0 to 4 were defined for each rubric criterion, with 0 being the lowest and 4 the highest. Note that only whole number scores were awarded. Scores were designed to represent what students should be including to address the assignment description and to account for common mistakes and patterns in students’ writing and understanding identified by researchers. About 10% of the total assignments were chosen at random for review by the researchers during the iterative rubric development process. After the initial version of the rubric criteria was developed, 20% of the students’ drafts and revision submissions were randomly chosen and scored by 3 experienced graders to determine inter-rater reliability (IRR) via percent agreement. A rubric with $IRR \geq 0.75$ is considered to be reliable for observational data. For rubrics with $IRR \leq 0.75$, an iterative process of refining the rubrics, scoring a random selection of student submissions, and calculation of IRR was conducted. Once all rubrics met our reliability standard, the scoring system was considered finalized (we note that many of the rubric criteria achieved reliability of $IRR \geq 0.85$, which is considered very reliable for textual analysis).[26] For subsequent analysis, every individual assignment was scored by an experienced researcher using the finalized rubric scoring system.

To quantify whether the improvements in student writing from draft to revision were statistically significant we performed t-tests with a statistical significance threshold of $\alpha = 0.05$. To identify whether the changes were meaningful, we used Cohen's d statistic as a measure of effect size. Cohen's d is a measure of the difference in two quantities relative to their variability in the population of interest[27]. For our data, we calculated

$$d = \frac{\bar{x}_f - \bar{x}_i}{\sigma_x}$$

where $\bar{x}_f$ and $\bar{x}_i$ are the revision and draft mean scores for the class on a given rubric, and σ_x is the pooled standard deviation of the scores on that rubric. The value of d for each rubric represents the combined effect of the peer review and revision processes and course instruction occurring during this timeframe on the attribute of student writing described by that rubric. For the present analysis, we consider $d \leq 0.5$ to be a small effect, $0.5 < d \leq 1.0$ a medium effect, and $d > 1.0$ a large effect.

Assessment of Content Knowledge

A concept-inventory-style assessment was developed to probe student gains in conceptual understanding. This assessment was administered in-class at the beginning and end of the semester; students were informed prior to administration that course credit would be given for completion of the assessment, and that correctness on assessment items would have no impact on course grades. The full assessment is provided in the Supplemental Materials.

The assessment consisted of eleven three-tiered items, with the first tier being a conceptual question, the second a short answer prompt for students to explain the reasoning behind their answer, and the third a prompt for students to rate their confidence in their answer. First-tier questions were all either multiple choice or “select all that apply” formats. Note that for some items the first tier consisted of questions with multiple parts; in our analysis, each part is scored as a separate question. Preliminary analysis of second tier responses revealed that student explanations were too brief or sporadic to inform our research, and so were excluded from the data set. In the third tier, students were prompted to report their confidence on a 1-5 Likert scale (with 1 corresponding to the lowest confidence, and 5, the highest). This paper focuses on analysis of students’ patterns of correct and incorrect responses from the first tier of each item.

Each of the eleven items (comprising 19 conceptual question parts in total) fell into one of six major topics: binary phase diagrams, stress-strain behavior, corrosion, crystal structures, atomic bonding, and the water phase diagram. In constructing the assessment, a set of candidate questions was compiled from existing question banks including a previously published assessment, exam questions, textbook problems, and homework sets. A team of course instructors and other subject matter experts then selected items on these topics with the greatest face validity relative to course content and content of WTL assignments. Four of the topics covered on the assessment were represented in WTL assignments (binary phase diagrams, stress-strain, corrosion, and crystal structures), and two were not (bonding and the water phase diagram).

Student assessment responses were included in the analysis only if both pre- and post-questions were completed, with responses collected from a population of 120 students across three semesters. For this study, assessment data was categorized into three groups based on population and topic.:

- WTL Group. Students who completed the associated WTL assignment
- Non-WTL Group. Students who did not complete the associated WTL assignment
- WTL-Free Group. All students, for question topics that did not have an affiliated WTL assignment

Students moved between the WTL Group and non-WTL Group for topics with an associated WTL assignment depending on whether they completed the assignment. In an attempt to offer content exposure comparable to that gained from completing a WTL assignment, students in the non-WTL group attended a discussion section covering concepts from the WTL assignment. All students are represented in the WTL-free Group. Furthermore, of the two assessment topics without an associated WTL assignment, bonding was taught in lectures while the water phase diagram was not directly covered.

For both the pre- and post-assessment, the fractions of correct responses were calculated individually for each first-tier question (separated into parts, if applicable) and in total for each topic. To quantify the compounded effect on content knowledge of WTL assignments plus instruction vs. instruction alone, we calculated the statistical significance and effect size of changes in student responses on the assessment. Statistical significance was calculated using the McNemar's test, which is an appropriate test for paired dichotomous data such as before-and-after responses categorized as either correct or incorrect. The test statistic has a χ^2 distribution with one degree of freedom, enabling determination of p-values using a χ^2 table. Effect size was quantified by calculating the normalized gain $\langle g \rangle$ of correctness fraction by question and by topic for both groups:

$$\langle g \rangle = \frac{\bar{x}_f - \bar{x}_i}{1 - \bar{x}_i} \quad \text{(Eq. 1.)}$$

where $\bar{x}_i$ and $\bar{x}_f$ are the draft and revision fraction correct, respectively. On most concept inventories normalized gains less than 0.3 are generally considered small, 0.3 to 0.6 medium, and 0.6 or above large; in traditional lecture-based courses, consistent normalized gains above 0.3 are very rare.[28]

Student Feedback on the WTL assignments

As part of the course, students responded to short, reflective writing questions throughout the course of the semester. The mid-term and end-of-semester reflective writing questions solicited feedback on the structure of the course, including the WTL assignments. The portion of the responses specifically about the WTL assignments were examined thematically to characterize students' self-reported attitudes about the assignments.[29] In total 252 responses were examined across the four semesters.

Results and Discussion

Writing Analysis:

To quantify the effect of the WTL process on students' conceptual understanding, writing prompt submissions were analyzed and scored, as described above. Average scores on each rubric criterion for the draft and revision reveal that students as a whole were able to improve their responses and successfully describe target content by the time of their revision. This can be seen in the increase between draft and revision scores on all criteria, and in the fact that most criteria have revision scores above 3 on a 0-4 scale. A t-test indicated that all writing score gains in this study were statistically significant. The consistent improvement and high scores observed among revisions indicates the ability of WTL to guide students toward a robust level of content understanding. However, it is insufficient to only focus on raw scores and raw improvement. To determine if engaging in the WTL process enhances student ability to apply course content to the topics covered in WTL assignments, we must look closely at the normalized magnitude of score improvement for each rubric criterion in order to identify the types of knowledge and skills for which WTL may be a highly effective pedagogy.

This normalized magnitude of score improvement is represented by Cohen's *d* effect size, as provided in Figure 1 for each scoring rubric criterion. The full set of rubrics can be found in the supplemental materials.

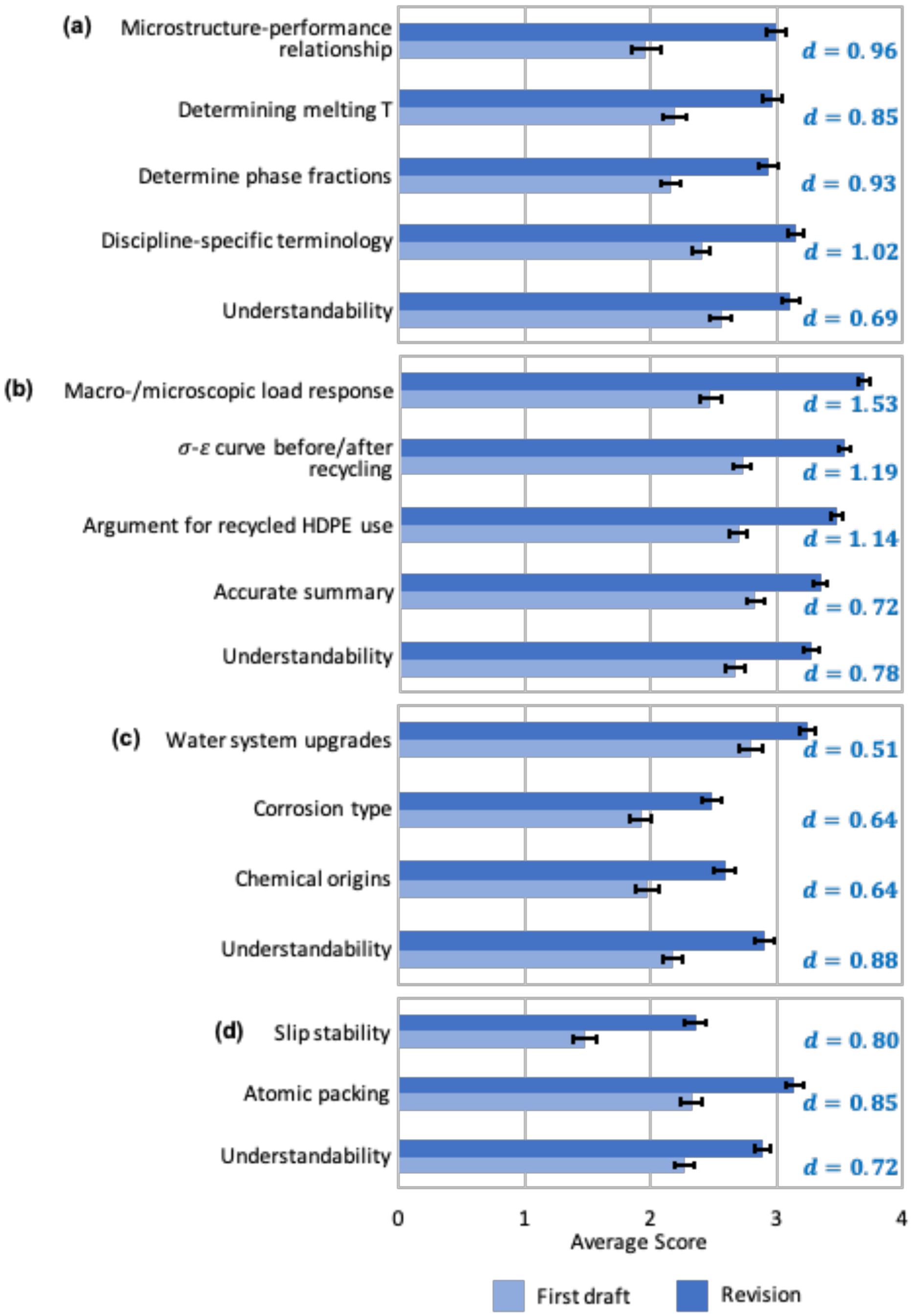

Figure 1. Bar charts of average writing scores achieved on each rubric criterion for the (a) binary phase diagram (N = 119), (b) stress-strain (N = 123), (c) corrosion (N = 114), and (d) crystal structures WTL prompts (N = 140). Student writing was scored on a 0-4 scale for each rubric criterion, with 0 indicating that a criterion's requirements were fully unaddressed in a student's response. Error bars represent standard error in scores.

Notably, Table I reveals that all Cohen's *d* effect sizes are considered medium or large, as defined above. This finding suggests that overall, the WTL process can contribute meaningfully to students' understanding of course content. These results can help identify knowledge and skills gained from completing WTL.

Especially high effect sizes corresponding to specific binary phase diagram and stress-strain rubric criterion indicate that the WTL process can facilitate growth in students' abilities to connect microscopic behavior and macroscopic properties, and to qualitatively analyze quantitative data (Figure 1). The largest effect size was achieved on the stress-strain rubric criterion assessing students' ability to link the macroscopic properties of polymers to the microscopic behavior of their constituent molecules $(d = 1.53)$, where the average revision score was 3.69 and indicates that students were successfully able to make this connection. A high average revision score of 2.99 and strong gains $(d = 0.96)$ were also achieved on the binary phase diagram rubric criterion, which targeted quality of macroscopic-microscopic connections in the context of how microconstituents dictate solder performance. Since learning to connect microscopic and macroscopic phenomena is a primary objective of many introductory materials science courses, these findings motivate implementation of WTL in such courses.[30]

Evidence that the WTL revision process can enhance the ability to process and qualitatively interpret quantitative data can be seen in the large effect sizes achieved on the binary phase diagram and stress-strain rubric criteria that evaluate the ability to accurately synthesize literature data into quantitative formats (lever rule calculations and stress-strain curve generation respectively). Success on the stress-strain rubric criterion for this skill is promising, since it has been shown that students at the introductory level often struggle to construct accurate stress-strain curves.[31] We believe strong growth in translating between representations is partly facilitated by the incorporation of writing into the data analysis process. By committing to a concrete verbalization of their thoughts while both writing/editing their drafts and interacting with peers during the peer review process, students are led to metacognitively engage with course content through evaluation and correction of their work and their peers' work. This process solidifies students' comprehension of course content by developing their ability to identify and address their mistakes.

Additional benefits of applying verbal reasoning are seen in the large gains for the binary phase diagram rubric criterion analyzing student ability to accurately incorporate discipline-specific terminology into their writing. Requiring students to use and explain expert-like language can help establish familiarity and fluency with relevant terms and concepts, lending to the development of a robust discipline-specific vocabulary.

While even our lower effect size values all fall into either the medium or large categories for this study, it is valuable to compare effect sizes across assignments to find evidence of factors that made some WTL assignments especially effective. Identifying differences between rubric criteria with higher and lower gains allows us to identify which capabilities were targeted effectively in WTL assignments, and which can benefit from enhanced or modified intervention in the future.

For all four prompts, the rubric criterion assessing whether the "memo is understandable to a person with minimal scientific background" demonstrated student gains of medium effect size ($0.69 < d < 0.88$). This rubric criterion differs from the others in that it defines success by quality of communication rather than demonstrated application of technical course content. Thus, the trend of comparatively lower Rubric 1 effect sizes seen in Table I reveals that students in general grow less in writing ability than in conceptual knowledge during the WTL revision process. The effect size distribution provides evidence that while students do make meaningful gains in writing ability, the primary learning outcomes are in conceptual learning and discipline-specific thinking ability, in alignment with the overall goals of WTL as a form of pedagogy and curriculum.

Students demonstrated moderate growth on the corrosion assignment, with comparatively smaller effect sizes observed in three of the corrosion rubric criteria. This outcome may arise from differences in the corrosion prompt structure compared to the stress-strain and binary phase diagram prompts: while the latter two prompts explicitly instruct students to leverage provided data and course content to produce a well-connected and robust response, the corrosion prompt effectively offers a checklist of facts to cover. Student responses often simply restated facts from course resources, with little variation between draft and revision, rather than generating qualitative/quantitative and macroscopic/microscopic connections. We find that the corrosion prompt led students to engage in recitation of declarative knowledge rather than rigorous analysis and application, as they likely lacked the expert-like understanding necessary to identify deeper connections between seemingly-unrelated factual components of the prompt.

Average draft scores for two criteria on this assignment were among the lowest draft rubric scores overall, which could reduce the efficacy of the peer revision process, as students may not have entered the WTL process with enough knowledge on the topic to effectively help each other improve via peer revision. The exclusion of contextual, quantitative problem-solving in the corrosion assignment may also have inhibited students' development of problem-solving strategies and deeper conceptual reasoning around "checklist" declarative knowledge items.

However, the third corrosion rubric criterion, on students' application of knowledge of corrosion mechanisms in designing improvements to public water infrastructure, demonstrates the second highest draft score on the assignment and a correspondingly high revision score. This result indicates that this WTL assignment does lead students to engage in some critical thinking on corrosion, especially in areas where they are better prepared before the draft writing process.

As illustrated in this analysis of the corrosion data, we can use differences in gains on rubric criteria within and between assignments to inform improvements in the WTL curriculum overall such as more consistently eliciting critical thinking, providing structure within assignments to help students link concepts and quantitative reasoning, and more effectively scaffolding the peer revision process.

Conceptual Assessment:

Gains in conceptual understanding and ability to transfer knowledge were quantified using results from an external assessment given to students at the beginning and end of the course (Figure 1). Supplementary Table S1 shows a further breakdown of these data. Assessment results from three semesters (Winter 2017, Winter 2018, and Winter 2019) were used in this analysis. Results were assigned to the three groups as described above in the methods.

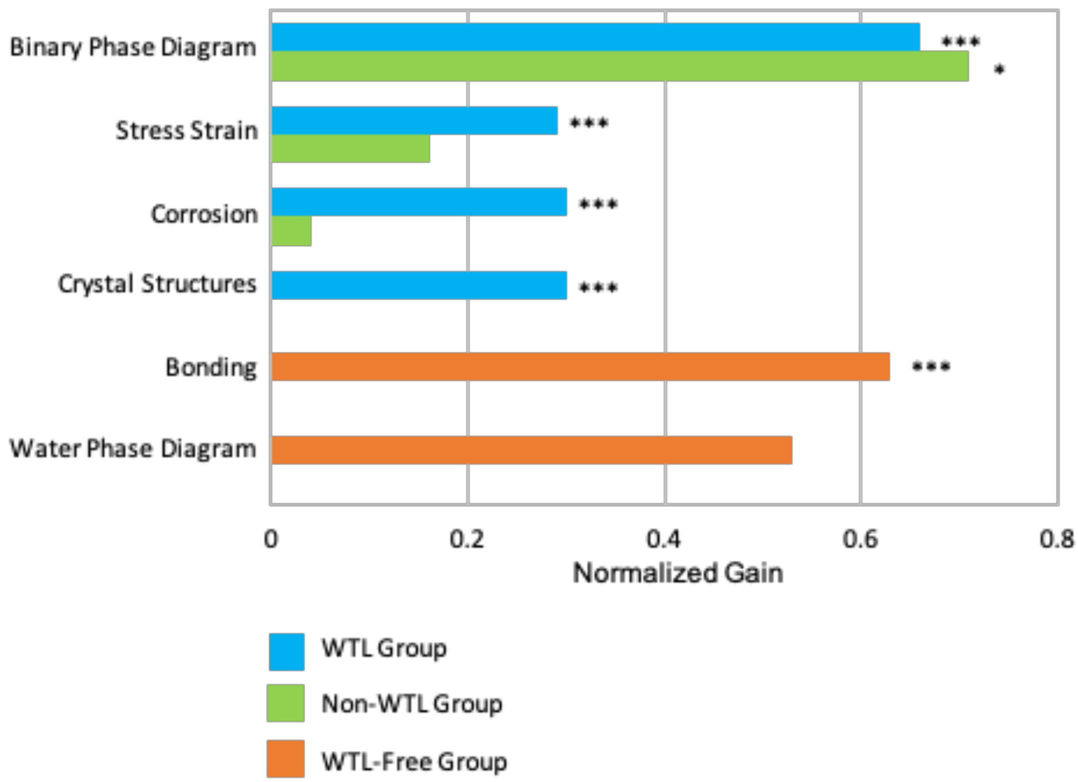

Figure 2. Comparison of pre- and post-assessment gains by topic. Topics represent content covered on each of the four WTL assignments, as well as two which are not covered on WTL assignments. Note that while bonding is covered explicitly within the course, the water phase diagram is not. * indicates $p < 0.05$, and *** indicates $p < 0.001$ for assessment scores.

While WTL score growth directly reflects improvement in students' writing from draft to revision, the assessment score gains in Figure 2 represent growth in conceptual knowledge attributed to both the WTL revision process and other components of the overall course instruction. Contrasting assessment gains between different groups can therefore provide information about the efficacy of WTL assignments per se for enhancing conceptual learning in the course.

Evidence of the positive impact of WTL can be found in comparing the performance of the WTL group to the non-WTL group on assessment topics that were represented in WTL assignments. Supplementary Figure S1 shows that the WTL group achieved statistically significant gains on all binary phase diagram, stress-strain, and corrosion questions, while the gains in the non-WTL group were only statistically significant for the binary phase diagram questions. Furthermore, Table I indicates that the average score of the WTL group for both the stress-strain and corrosion topics exceeded that for the non-WTL group. The difference in gains between the WTL and non-WTL groups indicate that WTL assignments enhanced students' ability to extrapolate and critically apply course content beyond the capabilities acquired from a traditional lecture-based course in isolation. This is likely not simply a function of time-on-task, as non-WTL groups covered the same content in discussion sections instead of completing WTL assignments.

It should be noted that the non-WTL group achieved higher gains on binary phase diagram questions than the WTL group. However, the sample size of the non-WTL group for this topic is only N=18 and the topical gain value exceeds that for the WTL group by a marginal amount (0.09). The non-WTL group's higher pre-test average (non-WTL group fraction correct = 0.36 vs. WTL group fraction correct = 0.32) may have also contributed to this result, as is has been shown that the normalized gain statistic favors larger starting values.[32] Furthermore, the gains of the WTL group for this topic are still high in an absolute sense. We therefore do not believe the marginally higher gains of the non-WTL group on this single topic to be evidence against the efficacy of WTL overall nor this assignment in particular.

Further corroboration of the impact of WTL can be found in examining the topical gain values collectively: While normalized gains of $\langle g \rangle \geq 0.30$ are rarely-encountered in traditional lecture-based courses, Table I demonstrates that the students in the WTL group achieved $\langle g \rangle \geq 0.29$ on all WTL topics.[33] The success of this group therefore implies the presence of an additional factor helping them achieve growth beyond that seen in traditionally-taught courses. Given that this was a traditionally-taught course outside of the WTL component, we attribute these large gains to the learning acquired by engaging with the WTL process.

Assessment results may also provide insight on WTL prompt design limitations, such as in the gains achieved by the WTL group on individual corrosion assessment questions. Evidence of bias in the wording of the corrosion prompt toward factual knowledge over analytical thinking may be found in comparing performance on items 6 ($\langle g \rangle = 0.50$) and 7 ($\langle g \rangle = 0.16$, the smallest statistically-significant gain in this study). Problem 6 asks students to identify a correct corrosion chemical reaction formula, while problem 7 provides a list of corrosion prevention techniques from which students are required to "select all that apply." These problems differ in expectation, with problem 6 favoring fact retrieval, and problem 7, application of underlying details and consequences of corrosion in order to determine how it could be prevented.

The fact that the WTL group achieved large gains for problem 6 is not surprising—on the pre-assessment, students had likely never seen a corrosion reaction. By the post-assessment, however, they had encountered corrosion reactions in class that were reinforced on the WTL assignment. The small gains achieved for problem 7, however, do correlate with our above analysis of the writing scores indicating ways to improve the corrosion WTL prompt for promoting growth in conceptual understanding.

We also propose enhanced WTL intervention for concepts corresponding to assessment questions on which the WTL group demonstrated gains or post-assessment scores that were on the lower end of this study. As indicated in supplementary Figure S1, we observed significant gains for all crystal structures questions except for the one labeled Identify Crystal Plane (c). This question asks students to identify a cross-sectional drawing of the (1 1 1) plane for an FCC crystal.

Notably, significant gains were achieved for the Identify Crystal Plane (a) and (b) questions, in which the students were tested on their ability to identify the (1 0 0) and (1 1 0) FCC planes. The (1 1 1) plane is arguably the most complex of the three planes, and the fact that students struggled to improve significantly in their ability to identify it suggests that their spatial awareness as related to crystal planes is not fully developed. While one may argue that proficiency in defining the (1 0 0) and (1 1 0) planes negates this claim, it should be noted that these planes are simpler and likely taught explicitly in class. Students may therefore benefit from the incorporation of rigorous closest-packed plane identification/explanation into future iterations of the crystal structures WTL assignment.

Student perceptions of the assignments

Beyond looking at gains in student knowledge from engaging in the WTL assignments we also examined student perceptions about the assignments. To do so, we gathered feedback from the students about the course elements, including the WTL assignments, during the middle and end of the semester. We examined students' responses pertaining to the WTL assignments to characterize self-reported learning and attitudes towards the assignments. Approximately half of the students reported that the WTL assignments enhanced their learning, ranging from a better understanding the content to developing their writing ability. Over a third of the students discussed the benefits of the WTL in supporting their conceptual understanding. While most of the responses were general, some students specified that the assignments reinforced, solidified, or deepened their understanding of both fundamental and complex concepts. Students identified that having to explain the targeted concepts allowed them to assess their own understanding and think more deeply about the content. Another subset of students mentioned that the authentic scenarios incorporated into each assignment supported their learning. These responses focused on how the authentic scenarios allowed them to apply the concepts they were learning in class, which students identified as supporting both their understanding of the material and its importance. Additionally, the authentic scenarios may play a role in the affective aspects of learning, as evidenced by some students who discussed how it made the content more interesting and made them feel like engineers. A small subset of the responses touched upon incorporating writing into a materials science course. In these responses, students demonstrated mixed attitudes towards writing, but the majority discussed appreciating the opportunity to develop their writing skills in the context of the WTL assignments. Similarly, student responses were mixed with respect to the peer review and revision elements of the assignments where some students identified them as helpful components and others did not. Overall, students' feedback responses provide additional evidence that the elements of the WTL assignments functioned as intended.

Conclusion

We have investigated the efficacy of WTL in an introductory Materials Science and Engineering course for facilitating significant gains in conceptual knowledge, and probed student reception of and feedback on these WTL assignments. We found that students' descriptions of content targeted by WTL assignments collectively improved on all areas measured in this study. The greatest improvement was observed on the ability to make connections between microscopic properties and macroscopic behavior, and to interpret quantitative data qualitatively, which are key learning goals of the course. Students demonstrated an enhanced grasp of the topics covered on WTL assignments by achieving statistically significant gains on corresponding assessment questions.

Implications for research and teaching:

We have found significant evidence that WTL can be an effective pedagogical tool for improving students' conceptual understanding in materials science. Combined analysis of writing scores and concept assessment scores indicate that when the WTL assignments provide a rigorous problem-solving scaffold for processing and contextualizing quantitative data, students grow significantly in their ability to identify meaningful macroscopic-microscopic connections and apply course content to qualify results. However, students resorted to disconnected fact recall, inducing less growth in conceptual understanding and application, when this assignment structure was absent. We therefore recommend to any instructor wishing to develop and/or implement WTL activities in their own classes that assignments (1.) incorporate an element of quantitative data interpretation and processing and (2.) provide a robust scaffolding for contextualizing these results in realistic scenarios.

We also note that while student comments about WTL as a teaching and learning activity were broadly positive, negative student comments around WTL primarily concerned the logistics of how the assignments were incorporated into the course—e.g., expectations, grading, and timing. While some logistical issues will arise in any course, their consideration could inform future implementation of WTL and we encourage instructors adopting WTL to be clear and thorough in communicating the purpose of WTL to their students as well as their expectations for how students should engage with WTL assignments.

Acknowledgements

We would like to acknowledge Anne Gere and Ginger Shultz for developing and running the MWrite program, which supports the implementation of WTL in the course discussed herein.

References

[1] P. D. Klein and P. Boscolo: Trends in Research on Writing as a Learning Activity. *Journal of Writing Research,* 7, 311 (2016).

[2] J. Emig: Writing as a Mode of Learning. *College Composition Communication,* 28, 122 (1977).

[3] D. Galbraith: Conditions for discovery through writing. *Instructional Science,* 21, 45 (1992).

[4] P. D. Klein, N. Arcon, and S. Baker: Writing to Learn. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (Second ed.). New York: The Guilford Press (2015).

[5] L. P. Rivard: A review of writing to learn in science: implications for practice and research. *Journal of Research in Science Teaching,* 31, 969 (1994).

[6] A. R. Gere, N. Limlamai, E. Wilson, K. MacDougall Saylor, and R. Pugh: Writing and Conceptual Learning in Science: An Analysis of Assignments. *Written Communication,* 36, 99 (2019).

[7] J. A. Reynolds, C. Thaiss, W. Katkin, and R. J. Thompson: Writing-to-learn in undergraduate science education: a community-based, conceptually driven approach. *CBE life sciences education,* 11, 17 (2012).

[8] J. H. Hanson and J. M. Williams: Using Writing Assignments to Improve Self-Assessment and Communication Skills in an Engineering Statics Course. *Journal of Engineering Education,* 97, 515 (2008).

[9] E. Wheeler and R. L. McDonald: Writing in Engineering Courses. *Journal of Engineering Education,* 89, 481 (2000).

[10] D. Ridgway, V. L. Young, and K. J. Sampson: Applying A Writing To Learn Strategy In A Traditional Material Science And Engineering Course. *Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition*, 2005, pp. 10.211.1 - 10.211.7

[11] J. Bylander and M. Gustafsson: Improved content mastery and written communication through a lab-report assignment with peer review: an example from a quantum engineering course. *European Physical Society,* 42, 025701 (2021).

[12] S. A. Finkenstaedt-Quinn, M. N. Petterson, A. Gere, G. Shultz: The Praxis of Writing-to-Learn: A Model for the Design and Propagation of Writing-to-Learn in STEM. *Journal of Chemical Education*, doi: https://doi.org/10.1021/acs.jchemed.0c01482.

[13] R. L. Ferguson: Constructivism and Social Constructivism. In G. M. Bodner, & M. Orgill (Eds.), *Theoretical Frameworks for Research in Chemistry/Science Education,* (pp. 28-49, Prentice Hall Series in Educational Innovation). Upper Saddle River, NJ: Pearson Prentice Hall (2007).

[14] G. M. Bodner: Constructivism: A theory of knowledge. *Journal of Chemical Education,* 63, 873 (1986).

[15] 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. *CBE - Life Sciences,* 17, ar28, doi:10.1187/cbe.17-10-0212 (2018).

[16] S. A. Finkenstaedt-Quinn, A. S. Halim, T. G. Chambers, A. Moon, R. S. Goldman, A. R. Gere, G. V. Shultz: Investigation of the Influence of a Writing-to-Learn Assignment on Student Understanding of Polymer Properties. *Journal of Chemical Education,* 94, 1610 (2017).

[17] S. A. Finkenstaedt-Quinn, A. S. Halim, G. Kasner, C. A. Wilhelm, A. Moon, A. R. Gere, G. V. Shultz: Capturing student conceptions of thermodynamics and kinetics using writing. *Chemistry Education Research and Practice,* 21, 922 (2020).

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

[19] J. A. Schmidt-McCormack, J. A. Judge, K. Spahr, E. Yang, R. Pugh, A. Karlin, A. R. Gere, G. V. Shultz: Analysis of the role of a writing-to-learn assignment in student understanding of organic acid–base concepts. *Chemistry Education Research and Practice,* 20, 383 (2019).

[20] S. A. Finkenstaedt-Quinn, E. P. Snyder-White, M. C. Connor, A. R. Gere, and G. V. Shultz: Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *Journal of Chemical Education,* 96, 227 (2019).

[21] F. M. Watts, J. A. Schmidt-McCormack, C. A. Wilhelm, A. Karlin, A. Sattar, B. C. Thompson, A. R. Gere, and G. V. Shultz: What students write about when students write about mechanisms: analysis of features present in students' written descriptions of an organic reaction mechanism. *Chemistry Education Research and Practice,* 21, 1148 (2020).

[22] William D. Callister, J.; Rethwisch, D. G., Materials Science and Engineering: An Introduction. 8th ed.; John Wiley and Sons: Hoboken, NJ, (2009).

[23] Montfort, D.; Brown, S.; Pollock, D., An Investigation of Students' Conceptual Understanding in Related Sophomore to Graduate-level Engineering and Mechanics Courses. *Journal of Engineering Education,* 98, 111 (2009).

[24] A. F. Heckler, R. Rosenblatt: Student Difficulties with Basic Concepts in Introductory Materials Science Engineering. *2011 Frontiers in Education Conference (FIE)*, pp. S2H-1, DOI: 10.1109/FIE.2011.6142989.

[25] S. M. Lord: Integrating Effective 'Writing to Communicate' Experiences in Engineering Courses: Guidelines and Examples, *Int. J. Engng Ed.* 25, (1), 196 (2009).

[26] K. A. Hallgren: Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, 8, 23 (2012).

[27] J. Cohen: The effect size index: d. *Statistical power analysis for the behavioral sciences* (Second ed.), New York: Lawrence Erlbaum Associates, Publishers (1988).

[28] R. R. Hake: Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics* 66, 65 (1998).

[29] V. Braun and V. Clarke: Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 77 (2006).

[30] S. Krause, J. C. Decker, J. Niska, and T. Alford: Identifying Student Misconceptions in Introductory Materials Engineering Classes. *Proceedings of the 2003 American Society for Engineering Education Annual Conference & Exposition*. American Society for Engineering Education, 2003, pp. 8.648.1-8.648.8.

[31] D. Montfort, S. Brown, and D. Pollock: An Investigation of Students' Conceptual Understanding in Related Sophomore to Graduate-level Engineering and Mechanics Courses. *Journal of Engineering Education 98*, 111 (2009).

[32] R. R. Hake: Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics,* 66, 64 (1998).

[33] J. A. Schmidt-McCormack, J. A. Judge, K. Spahr, E. Yang, R. Pugh, A. Karlin, A. R. Gere, G. V. Shultz: Analysis of the role of a writing-to-learn assignment in student understanding of organic acid–base concepts. *Chemistry Education Research and Practice,* 20, 383 (2019).