MRS Communications

Prospective

Check for updates

Writing-to-learn in introductory materials science and engineering

L. Marks and **H. Lu**, Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA; Sweetland Center for Writing, University of Michigan, Ann Arbor, MI 48109, USA

T. Chambers, Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA

S. Finkenstaedt-Quinn, Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

R. S. Goldman, Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA; Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

Address all correspondence to R. S. Goldman at **rsgold@umich.edu**

Abstract

We examine the impact of writing-to-learn (WTL) on promoting conceptual understanding of introductory materials science and engineering, including crystal structures, stress–strain behavior, phase diagrams, and corrosion. We use an analysis of writing products in comparison with pre/post concept-inventory-style assessments. For all topics, statistically significant improvements between draft and revision scores are apparent. For the stress–strain and phase diagram WTL assignments that require synthesis of qualitative data into quantitative formats, while emphasizing microstructure-properties correlations, the highest WTL effect sizes and medium-to-high gains on corresponding assessments are observed. We present these findings and suggest strategies for future WTL design and implementation.

Introduction

Writing has long been used to support learning across a range of contexts and disciplines.[1–3] One such writing-based instructional practice, writing-to-learn (WTL), has been incorporated into classrooms in forms spanning reflective writing to long, scaffolded writing assignments. Across disciplines, WTL has been used to support instructional goals such as developing both disciplinary thinking and conceptual learning.[4,5] Within science education, WTL assignments have been used to support the development of scientific argumentation, metacognition, and conceptual understanding by students.[5–7] These goals are further represented in the writing assignments described in the engineering education literature.[8–11] However, only a few studies of WTL in materials science have been reported to date.[10,12] 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.[12] Here, we expand upon our prior work by considering student responses across a comprehensive set of WTL assignments spanning materials classes and functionalities. Our aim is to investigate student gains in comprehension and application of course content across a term and to 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 a prompt conveying an authentic scenario, performing and receiving content-focused peer review, and finally revising their initial response.[13] This three-step WTL process incorporates the key elements for effective WTL assignments identified by meta-analyses of WTL literature, namely clearly defined and interactive writing expectations that incorporate meaning-making tasks and support metacognition.[6,14,15] This WTL process also aligns with cognitive theories of learning such as social constructivism,[1,6,10,14,15] which posits that students learn within their individual social environments by restructuring existing knowledge to incorporate new knowledge.[14,15] Indeed, research has shown that this WTL process has enabled students to constructively engage with the peer review and revision processes,[16,17] thereby supporting them in learning challenging content in a wide range of introductory STEM courses, including biology, chemistry, and statistics.[12,18–22] Our WTL implementation also follows the principles for designing effective "writing to communicate" experiences in engineering, with writing assignments that include an authentic investigation and audience, are tied to the technical course content, and provide useful practice for engineering careers, while not being overly burdensome for the engineering faculty instructor.[23]

The core objective of an introductory materials science and engineering course is to introduce the principles of engineering materials, with an emphasis on understanding 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.[24] Typically, these principles are framed in the context of materials classes (metals, ceramics, polymers, semiconductors, and composites), each with their characteristic chemistry and internal structure. The role of materials processing is described in terms of thermodynamics (via phase diagrams) and kinetics (via diffusion). Finally, materials functionality is introduced, 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 both an authentic scenario, and social elements in the form of peer review and revision, on student understanding of key concepts in introductory materials science and engineering (MSE). With an emphasis on key MSE concepts, including crystal structures, stress–strain behavior, phase diagrams, and corrosion, the 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 MSE learning goals are best supported by the WTL assignment design?

Methods
Implementation
For this study, we use two main sources of data: (1) draft and revision responses to the WTL assignments, numerically scored using a rubric generated by the research team, and (2) responses to an MSE concept-inventory-style assessment external to the required coursework given at both the beginning and the end of the course, which we term the "pre" and "post" assessments, respectively. To examine the contributions of peer review and revision to improvements in students' conceptual descriptions from draft to revision, we employed qualitative and quantitative analyses of student writing. We also used quantitative analysis of the concept-inventory-style assessment responses to compare the learning gains of students who participated in WTL with those who instead participated in a guided-group discussion intended to offer comparable content exposure. In a given term, all students completed at least two of the four WTL assignments.

This study was conducted at a midwestern university in a lower-level MSE course during three separate terms. The course consists of lecture and recitation sections with coursework including traditional problem sets, bi-weekly reflective writing, and WTL assignments. Prerequisite courses include either general chemistry or introductory organic chemistry. The textbook for the course is "Materials Science and Engineering" by Callister and Rethwisch.[25] Across the three terms (Winter 2017, Winter 2018, and Winter 2019), the course participants consisted of 151 students who ranged from sophomores to seniors, as well as two graduate students auditing the course. Most

of the students were enrolled in the College of Engineering, with more than fifty percent intending to major 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 focus on concepts identified to be the most challenging for students of introductory materials science.[26–29] Each WTL assignment consists of an initial written response to a prompt ("draft"), anonymous open-response peer review performed by 2–3 randomly selected students ("peer review"), and a revision of the draft ("revision"). Peer review is guided by content-focused rubrics, as shown in the Appendix. For each WTL assignment, one week is given for the initial response and half a week for both the peer review and revision, respectively. For the draft and peer review, student scores are based upon completion, with a cursory check to verify that all prompt requirements are addressed. For the revision, student scores are based upon alignment of student responses with assignment-specific rubric criteria. Throughout each term, additional support for students is provided by two peer tutors ("Writing Fellows") familiar with introductory materials science through prior enrollment in this course or its equivalent. The writing fellows are trained to help students approach the writing assignments while learning content; they are available to facilitate the peer review process and answer both writing and content questions.

Writing analysis: rubrics and statistical significance
We analyzed draft and revision WTL submissions using rubrics designed to evaluate conceptual understanding by probing student ability to describe relevant course content, as shown in the Appendix. For each WTL assignment, a rubric consisting of at least two content-specific criteria, each scored from 0 (lowest) to 4 (highest), was iteratively developed. In addition, the understandability for an audience with minimal scientific background was evaluated as an additional rubric criterion. For the rubric development process, we initially examined ~10% of total submissions in order to gain a general understanding of student responses, as well as to identify common mistakes and patterns in student writing and understanding. Using this analysis to develop initial rubric criteria, 20% of the student drafts and revisions were randomly selected for scoring by 3 experienced graders to determine inter-rater reliability (IRR) via percent agreement. Since a rubric with IRR ≥ 0.75 is considered to be reliable for observational data, we utilized an iterative process of refining the rubrics, scoring a random selection of student submissions, and re-calculation of IRR until 0.75 was achieved.[30] Once all rubrics met this reliability standard, the scoring system was considered finalized. We note that many of the rubric criteria achieved reliability of IRR ≥ 0.85, which is considered very reliable for textual analysis.[31] For subsequent analysis, every individual assignment was scored by an experienced

Table I. Analysis of writing score data sets by topics and rubrics. N is the number of participants. Each data set has been fit to a normal distribution, yielding mean draft and revisions scores ($\overline{x}_d$, $\overline{x}_r$) and standard deviations of the draft and revision scores (σ_d, σ_r) with R^2 values primarily exceeding 0.8.

Topic and rubric	N	$\overline{x}_d$	σ_d	R^2_d	$\overline{x}_r$	σ_r	R^2_r
Crystal structure	140						
Slip stability		2.26	0.94	0.81	2.89	0.75	0.67
Atomic packing		2.00	1.01	0.92	3.21	0.87	0.84
Understandability		1.34	1.14	0.93	2.30	0.80	0.97
Stress–strain	125						
Macro-/micro-load response		2.70	0.89	0.94	3.28	0.70	0.94
Curve before/after recycling		2.84	0.79	0.97	3.35	0.63	0.95
Argument for recycled HDPE		2.71	0.81	0.94	3.47	0.53	0.99
Accurate summary		2.74	0.79	0.99	3.54	0.55	0.92
Understandability		2.49	1.00	0.99	3.68	0.52	0.92
Phase diagrams	120						
Microstructure performance		2.43	0.84	0.41	2.85	0.80	0.90
Melting T		2.31	0.71	0.82	2.89	0.82	0.89
Phase fractions		2.11	0.79	0.99	2.72	0.82	0.94
Discipline-specific terminology		2.18	0.87	0.98	2.75	0.88	0.98
Understandability		1.96	1.16	0.95	2.77	0.83	0.98
Corrosion	113						
Water system upgrades		2.13	0.80	0.86	2.65	0.82	0.98
Corrosion prevention		1.99	0.91	0.99	2.43	0.85	0.94
Corrosion chemistry		1.95	0.85	0.86	2.35	0.77	0.96
Understandability		2.58	0.97	0.96	3.02	0.87	0.96

researcher using the finalized rubric scoring system. Finally, the statistical validity of this "writing score" analysis was examined. For each rubric criteria, the draft and revision scores were fit to a normal distribution, yielding most R^2 values in excess of 0.8,[32] as shown in Table I.

To quantify the statistical significance of improvements in student writing from draft to revision, we performed t-tests with a significance threshold of $\alpha = 0.05$.[33] For those criteria with $p < \alpha$, we then used Cohen's d statistic as a measure of effect size to quantify improvements in student writing. Cohen's d is a measure of the difference in two quantities relative to their variability in the population of interest.[34] For our data, we calculated

$$d = \frac{\overline{x}_r - \overline{x}_d}{\sigma_x} \qquad (1)$$

where $\overline{x}_d$ and $\overline{x}_r$ are the mean draft and revision scores for a given rubric criterion, and

$$\sigma_x = \sqrt{\frac{(\sigma_r^2 + \sigma_d^2)}{2}} \qquad (2)$$

(where σ_r and σ_d are the standard deviations of draft and revision scores for each rubric) is the pooled standard deviation of the scores for that rubric. For each rubric criterion, the value of d is attributed to the combined effects of the peer review and revision processes. For the present analysis, we consider $d \leq 0.5, 0.5 < d \leq 1.0$, and $d > 1.0$ to be small, medium, and large effect sizes, respectively.[34]

Content knowledge assessment

To probe student gains in conceptual understanding, we developed and administered an MSE concept-inventory-style assessment, shown in the Appendix. Similar to the WTL assignments described above, the assessment questions focus on concepts identified to be the most challenging for students of introductory materials science.[27–29] Drawing from the Materials Concept Inventory,[35] the Crystal Spatial Visualization Survey,[36] and other published assessments for introductory materials science,[27] we compiled a set of candidate questions. A team of course instructors and other subject matter experts then selected items with the greatest content validity relative to course and WTL assignment topics.[37] Four of the assessment topics were also represented in WTL assignments (crystal structures, stress–strain, phase diagrams, and corrosion), while one was not (atomic bonding).

The assessment consisted of eleven three-tiered items including a conceptual question, a short answer prompt to explain reasoning, and a confidence self-rating. For the first-tier questions, the format was either multiple choice or "select all that apply." For several items, the first-tier consisted of questions with multiple parts; in our analysis, each part was scored as a separate question, resulting in 19 total conceptual questions. Preliminary analysis of second-tier responses revealed that student explanations were too brief or sporadic to inform our research; thus, these were not included in 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). Finally, during the writing of this manuscript, it was discovered that appropriate answers were

not available for one of the items that consisted of 4 conceptual questions. Therefore, this paper focuses on the analysis of first-tier student responses for 15 conceptual questions.

For the analysis, we considered only the responses from students who completed both the pre- and post-assessments, including 120 students across three terms. The resulting assessment data were then categorized into three groups based on population and topic, as follows:

- WTL Group: students who completed the assessments and the associated WTL assignment.
- Non-WTL Group: students who completed the assessments but did not complete the associated WTL assignment.
- WTL-Free Group: all students who completed the assessments for atomic bonding and the water phase diagram (part of the phase diagrams assessment topic), which do not have an associated WTL assignment.

For both the pre- and post-assessments, the fractions of correct responses were calculated individually for each first-tier question (separated into parts, when applicable) and collectively by assessment topic. To quantify the compounded effects on content knowledge of WTL assignments plus instruction vs. instruction alone, we calculated the statistical significance of the differences in mean fraction of correct answers from the pre- and post-assessments and the learning gains from the pre- to post-assessment. Statistical significance was calculated using McNemar's test, which is appropriate 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.

To gauge the efficacy of WTL in promoting conceptual understanding, we also compare the ratio of the average gain achieved by each population (WTL vs. non-WTL groups) to their maximum possible gain, i.e., the normalized gain $\langle g \rangle$[38]:

$$\langle g \rangle = \frac{\overline{x}_f - \overline{x}_i}{1 - \overline{x}_i} \quad (3)$$

where $\overline{x}_i$ and $\overline{x}_f$ are the fractions of correct responses for the pre- and post-assessments, respectively. On most concept inventories, $\langle g \rangle < 0.3$ is generally considered small, $0.3 \leq \langle g \rangle < 0.7$ medium, and $\langle g \rangle \geq 0.7$ large; in traditional lecture-based courses, $\langle g \rangle < 0.3$ is typical.[38]

Mid-term and end-of-term student reflections

As part of the course, students responded to short, reflective writing questions throughout the term. The mid-term and end-of-term 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 self-reported student attitudes about the assignments.[39] In total 252 responses were examined across the three terms.

Results and discussion
Analysis of writing products

In this section, we present the analyses of the writing products for the WTL assignments, emphasizing student performance on drafts and revisions. To identify the content knowledge and skills for which WTL is an effective pedagogy in the context of MSE, we consider the analyses of the writing products both individually and collectively. Figure 1 presents bar charts of average writing scores achieved on rubric criteria for the (a) crystal structures, (b) stress–strain, (c) phase diagram, and (d) corrosion WTL assignments. For all rubric criteria, the increases between draft and revision scores are statistically significant; for most criteria, the revision scores, $\overline{x_r} > 3$ on a scale of 0–4. The consistent improvement in student responses coupled with high revision scores suggests that the peer review and revision processes guide students toward a robust level of content understanding. To further probe the role of the WTL process on student content understanding, we examine the Cohen's effect size, d.[34] As shown in Fig. 1(a)–(d), $d > 0.5$ for all rubric criteria across WTL assignments, indicating that the peer review and revision processes contribute meaningfully to overall student understanding of course content.

For the crystal structure WTL, shown in Fig. 1(a), medium effect sizes in the range $0.75 \leq d \leq 0.96$ are observed for all rubric criteria. The highest effect sizes are apparent for the "atomic packing" and "slip stability" rubric criteria, which emphasize the relationship between tightly packed spheres (oranges) and crystal structures, and the mechanical stability of stacked planes of spheres (oranges), respectively. The "atomic packing" rubric criterion assesses student crystal structure visualization ability, a critical skill for understanding the role of atomistic structure on macroscopic properties. The "slip stability" rubric criterion also assesses crystal structure visualization ability, while further probing student ability to recognize the differences in the distribution of forces for atoms vs. oranges.[40] It is interesting to note that the "understandability" rubric criterion, which probes the accessibility of each writing product to an audience with minimal scientific background, exhibits the lowest effect size for the crystal structure WTL. This rubric criterion differs from the others in that it defines success by quality of communication rather than demonstrated application of technical course content.

For the stress–strain WTL, shown in Fig. 1(b), medium-to-large effect sizes in the range $0.70 \leq d \leq 1.48$ are observed. In this case, the "macro-/micro- load response" and "σ-ε curve before/after recycling" rubric criteria exhibit the highest effect sizes, with $d = 1.48$ and $d = 1.16$, respectively. It is interesting to note that the "macro-/micro- load response" rubric criterion assesses student ability to link polymer macroscopic properties with the microscopic configurations of their constituent molecules. Since learning to connect microscopic and macroscopic phenomena is a primary objective of many introductory materials science courses,[26] these findings motivate further implementation of this WTL assignment in such courses. Meanwhile, the "σ–ε curve before/after recycling" rubric criterion evaluates
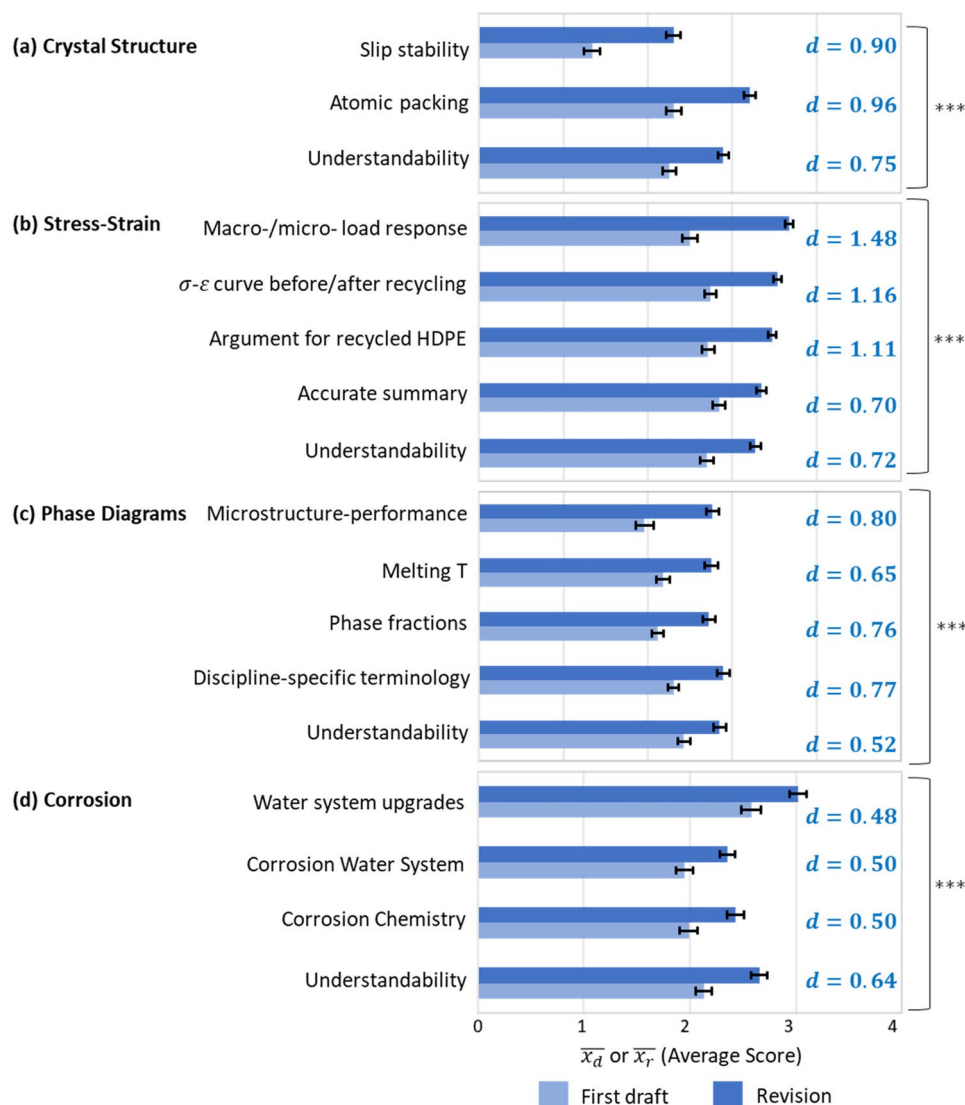
Figure 1. Bar charts of mean writing scores on rubric criteria for the (a) crystal structure ($N = 140$), (b) stress–strain ($N = 123$), (c) phase diagrams ($N = 119$), and (d) corrosion ($N = 114$) WTL prompts. For each rubric criterion, student writing was scored on a scale from 0 (lowest) to 4 (highest). All writing-score data sets were fit to a normal distribution, most with $R^2 > 0.8$, as shown in Table I. To quantify improvements in student writing, Cohen's d values were computed; those d values with a significance threshold of $p < 0.05$ are shown next to each item. Error bars represent the standard deviations of the means. ***indicates $p < 0.001$.

student ability to synthesize literature data into quantitative formats, namely to construct stress–strain curves based upon numerical values of key physical parameters. Since the construction of stress–strain curves has been identified as a difficult skill for students at the introductory level,[26] these findings reveal the broad benefit of this WTL assignment to early career undergraduates. Finally, for the stress–strain WTL, the "understandability" rubric criterion once again exhibits the lowest effect size. Across WTL assignments, the emerging trend of comparatively lower effect sizes for the "understandability" rubric criterion reveals that student growth in conceptual knowledge exceeds student growth in writing ability during the revision process.

For the phase diagrams WTL, shown in Fig. 1(c), medium effect sizes in the range $0.52 \leq d \leq 0.80$ are apparent. The

"discipline-specific terminology" and "microstructure-performance relationship" rubric criteria exhibit the highest effect sizes, with $d = 0.77$ and $d = 0.80$, respectively. The medium effect size for the "discipline-specific terminology" rubric criterion, which analyzes student ability to accurately incorporate discipline-specific terminology into writing products, reveals the benefits of applying verbal reasoning to the WTL process. Indeed, requiring students to use and explain expert-like language helps them to establish familiarity and fluency with relevant terms and concepts, lending to the development of a robust discipline-specific vocabulary. Similar to the large effect size for the "macro-/micro- load response" rubric criterion for the stress–strain WTL described above, the medium effect size for the "microstructure-performance relationship" rubric criterion,

which targets student ability to relate the microstructure of solder to its macroscopic performance, further supports the implementation of WTL in introductory materials science courses. Due to these medium-to-high effect sizes, we hypothesize that the WTL process facilitates student growth in connecting microscopic structure to macroscopic properties and performance. Finally, for phase diagrams WTL, the lowest effect size is observed for the "understandability" rubric criterion, consistent with the lower student growth in writing ability during the revision process discussed above.

For the corrosion WTL, shown in Fig. 1(d), low-to-medium effect sizes in the range $0.48 \leq d \leq 0.64$ are observed. In this case, the "understandability" rubric criterion exhibits the highest effect size amongst all WTL assignments. Furthermore, for the content-focused rubric criteria, "corrosion chemistry," "water system corrosion," and "water system upgrades," minimal variation between draft and revision scores are apparent, resulting in effect sizes of $d = 0.50, 0.50$, and 0.48, respectively. In comparison with the other WTL assignments, the relatively high effect size for the "understandability" criterion, coupled with the relatively low effect sizes for the content-focused rubric criteria, is likely due to the emphasis on declarative knowledge rather than quantitative problem-solving and micro-/macroscopic linkage. As will be discussed below, we suggest that future iterations of the Corrosion WTL assignment include opportunities for students to make quantitative correlations between microstructure and properties.

For most WTL assignments, the lower effect sizes for the "understandability" rubric criterion provides evidence that the primary learning outcome from the WTL process is in conceptual learning and discipline-specific thinking rather than refined prose. Indeed, by committing to a concrete verbalization of thoughts during the writing and editing of drafts, and by interacting with peers during the peer review process, students are led to metacognitively engage with course content through evaluation and revision of writing products. This process enables solidification of student comprehension by enabling them to identify and address their mistakes, in alignment with the overall goals of WTL as a form of pedagogy and curriculum. To optimize these outcomes, we suggest that future WTL assignments mimic the structure of the prompts that demonstrated high effect sizes on content-focused rubric criteria (i.e., the crystal structures, stress–strain, and phase diagrams prompts) by providing guidance within assignments to help students link concepts and quantitative reasoning.

Concept-inventory style content assessment

We now present the analysis of the concept-inventory-style content assessment that was administered at the beginning and end of each term. To individually and collectively examine the normalized gains in conceptual knowledge, we consider 15 conceptual assessment items grouped by topic. As shown in Fig. 2, the assessment topics include (a) atomic bonding, (b) crystal structures, (c) stress–strain, (d) phase diagrams, and (e)

corrosion. To distinguish the datasets based upon population, the bar charts are color-coded to indicate their response group, i.e., WTL Group, Non-WTL Group, and WTL-Free Group, as described above. The comparison of scores and normalized gains, $\langle g \rangle$, across the WTL, Non-WTL, and WTL-Free Groups provides information about the relative roles of the WTL revision process vs. other course learning opportunities in enhancing conceptual learning.

We first consider the atomic bonding topic, shown in Fig. 2(a), which is not currently represented by a corresponding WTL assignment. For this topic, the relationships between the shape of potential energy curves and materials properties are probed. As shown in the bar chart in Fig. 2(a), small-to-medium statistically significant gains of $\langle g \rangle \geq 0.20$ are apparent for all four bonding items. The highest gains were obtained for "melting point," "elastic modulus," and "lattice parameter" items with $\langle g \rangle = 0.51, 0.44$, and 0.40, respectively, exceeding gains of $\langle g \rangle = 0.30$ anticipated for traditional lecture-based courses.[38]

For the crystal structures topic, all students completed the associated WTL assignment. As shown in the bar chart in Fig. 2(b), for three of the four crystal structure items, small-to-medium gains of $\langle g \rangle \geq 0.29$ are apparent. The highest gains were obtained for the "identify (100) FCC plane" and "identify (110) FCC plane" items, with $\langle g \rangle = 0.59$ and 0.38, respectively, both exceeding expectations for traditional lecture-based courses. On the other hand, lower gains—albeit statistically significant—were obtained for the "selection of close-packed plane" item, namely $\langle g \rangle = 0.29$, within the expected gains for a traditional lecture-based course. Finally, for the "identifying (111) FCC plane" item, gains of $\langle g \rangle = 0.09$ were observed. These findings are consistent with earlier reports revealing student misconceptions about the atomic configurations within the (111) FCC plane.[41] Interestingly, in an earlier study of learning gains in a traditional introductory MSE course (without WTL), the "identifying (111) FCC plane" item yielded $\langle g \rangle = 0.02$, approximately one-fourth of our gain.[36] While our data set is not large enough to establish statistical significance for the difference in these gains, the comparison suggests that WTL is effective in helping students overcome stubborn misconceptions about planes in FCC structures.

Next, we consider the stress–strain topic, with "polymer stress–strain" and "metal yield strength" shown in the bar chart in Fig. 2(c). For the stress–strain topic, small-to medium gains of $\langle g \rangle \geq 0.29$ are apparent for the WTL group, while statistically insignificant gains of $\langle g \rangle \leq 0.13$ are observed for the non-WTL group. As discussed above, the stress–strain WTL requires students to link polymer macroscopic properties with the microscopic behavior of their constituent molecules. The stress–strain WTL also requires students to distill tabulated stress–strain data into quantitative stress–strain plots. Since the ability to interpret quantitative and qualitative stress–strain data was critical for the "polymer stress–strain" and "metal yield strength" items, the reinforcement of these skills through the stress–strain WTL has evidently led to higher scores for the WTL group. As will be further illustrated for the corrosion
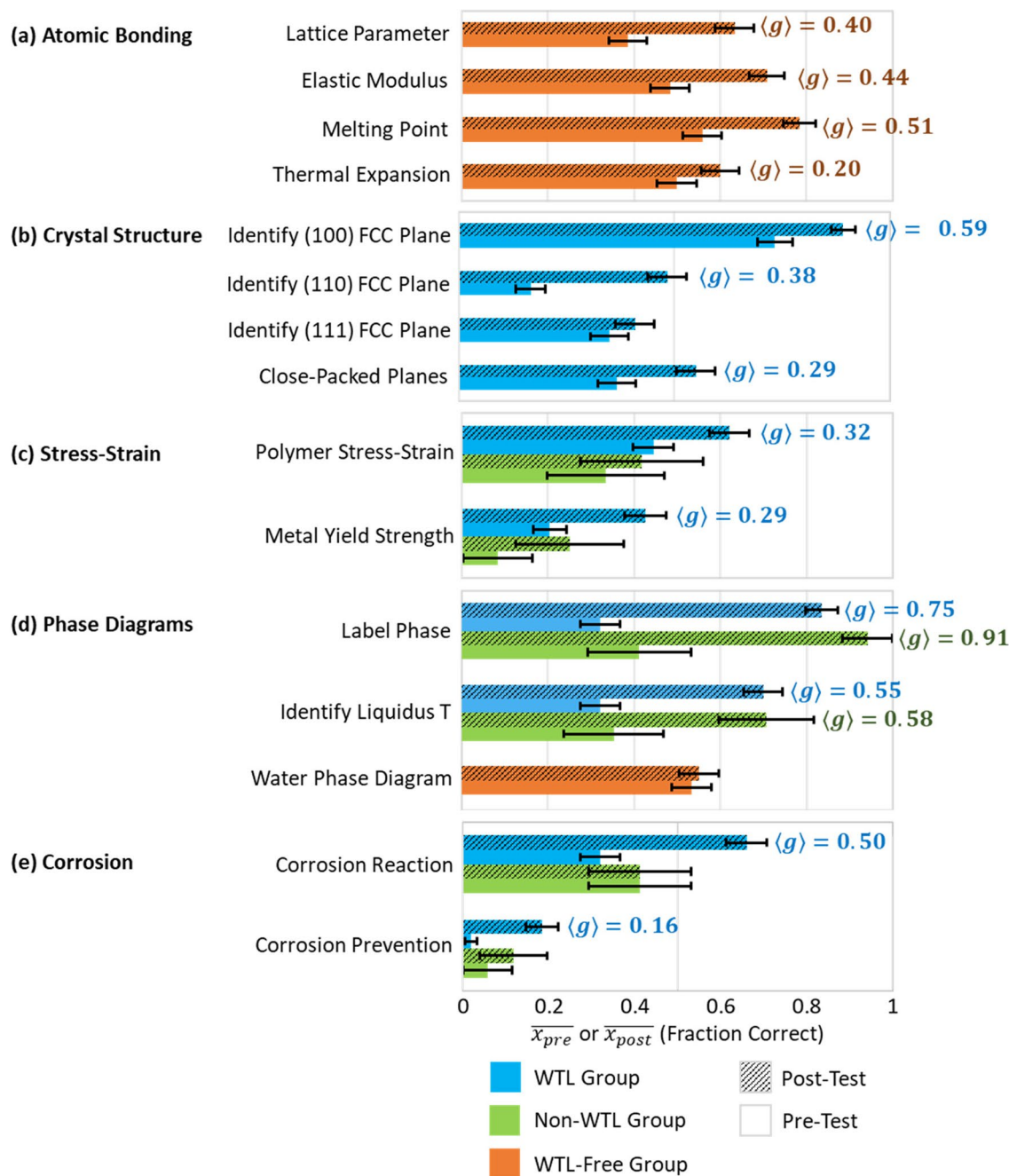
Figure 2. Pre- and post-assessment average item scores grouped by student group (WTL vs. non-WTL vs. WTL-free) and by topic: (a) atomic bonding, (b) crystal structure, (c) stress-strain, (d) phase diagrams, and (e) corrosion. Statistically significant ($p < 0.05$) normalized gain values, $\langle g \rangle$, i.e., the ratio of the average gain achieved by a population to their maximum possible gain, were used to gauge the efficacy of a WTL in promoting conceptual understanding. Error bars represent standard error in scores. More details on the population and distribution of scores are available in Table II.

topic below, the difference in gains between the WTL and non-WTL groups indicate that WTL assignments have enhanced student abilities 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.

Figure 2(d) presents the average scores for the phase diagrams topic, including the "label phase," "identify liquidus," and "water phase diagram" items. For the "water phase diagram" item, which is not represented by a corresponding WTL assignment, negligible gains are observed. Although pressure–volume phase diagrams are briefly discussed in lectures and readings, they are not included in conventional homework;

Table II. Assessment data by topic and item for the WTL, non-WTL, and WTL-free groups. N is the number of participants for each topic, $\bar{x}_i$ and $\bar{x}_f$ are the mean fraction of correct answers for each item from the pre- and post-assessments, χ^2 is the McNemar statistic, p indicates the statistical significance of changes from pre- to post-scores, and $\langle g \rangle$ is the normalized gain from pre- to post-scores.

Topic and item	N	$\bar{x}_i$	$\bar{x}_f$	χ^2	p	$\langle g \rangle$
WTL group						
Crystal structure	120					
Identify (100) FCC plane		0.73	0.89	12.44	<0.05	0.59
Identify (110) FCC plane		0.17	0.48	28.88	<0.05	0.38
Identify (111) FCC plane		0.35	0.41	1.00	0.32	0.09
Close-packed planes		0.37	0.55	7.12	<0.05	0.29
Stress–strain	108					
Polymer stress–strain		0.45	0.62	7.37	<0.05	0.32
Metal yield strength		0.20	0.43	16.00	<0.05	0.29
Phase diagrams	102					
Label phase		0.32	0.83	46.00	<0.05	0.75
Identify liquidus T		0.32	0.70	27.65	<0.05	0.55
Corrosion	102					
Corrosion reaction		0.31	0.66	22.27	<0.05	0.50
Corrosion prevention		0.20	0.18	13.76	<0.05	0.16
Non-WTL group						
Stress–strain	12					
Polymer stress–strain		0.27	0.36	0.14	0.71	0.13
Metal yield strength		0.09	0.18	1.00	0.32	0.10
Phase diagrams	18					
Label phase		0.39	0.94	9.00	<0.05	0.91
Identify liquidus T		0.33	0.72	3.00	<0.05	0.58
Corrosion	18					
Corrosion reaction		0.44	0.44	0.00	1.00	0.00
Corrosion prevention		0.06	0.17	0.33	0.56	0.18
WTL-free group						
Atomic bonding	12					
Lattice parameter		0.39	0.63	14.25	<0.05	0.40
Elastic modulus		0.48	0.71	11.95	<0.05	0.44
Melting point		0.56	0.78	12.36	<0.05	0.51
Thermal expansion		0.50	0.60	2.48	<0.05	0.20
Phase diagrams	18					
Water phase diagram		0.53	0.55	2.00	0.16	0.04

therefore, the negligible gains are not surprising. For the "label phase diagram" and "identify liquidus" items, both the WTL and non-WTL groups achieved statistically significant normalized gains, all with $\langle g \rangle > 0.55$, as shown in Fig. 2(d). We attribute this result to robust classroom instruction and rigorous (non-writing) homework assignments on binary phase diagrams. These high gains may also be explained by the extended emphasis on binary phase diagrams, with an allocation of approximately 1.5 weeks during a 14-week term. The similarity of these gains may also be attributed to the content of the two binary phase diagrams assessment items, which both emphasize declarative knowledge rather than deep microstructural analysis skills developed throughout the WTL process. Therefore, for future assessment iterations, we suggest an added emphasis on microstructure-focused analysis of binary phase diagrams in order to better understand the efficacy of the binary phase diagrams WTL assignment in enhancing conceptual learning.

Finally, for the corrosion topic, the average scores of the "corrosion reaction" and "corrosion prevention" items are shown in the bar charts in Fig. 2(e). In this case, small-to-medium gains are observed for the WTL group, while negligible gains are apparent for the non-WTL group. For the "corrosion reaction" item, which emphasizes equation recognition and manipulation, the WTL group achieved a medium gain of $\langle g \rangle = 0.50$, while for the "corrosion prevention" item, which emphasizes declarative knowledge, the WTL group achieved a small gain of $\langle g \rangle = 0.16$. The relatively low pre-test score for the "corrosion reaction" item ($\bar{x}_i = 0.31$), suggests that students had limited pre-course familiarity with the relationships between oxidation, reduction, and corrosion. Subsequent student exposure to oxidation and reduction reactions from lectures, readings, and conventional homework combined with the interpretation of the reactions in the corrosion WTL assignment, likely provided sufficient time-on-task for recall. Finally,

the lower gains for the "corrosion prevention" item reveals that further student growth may require the corrosion WTL assignment to facilitate more extensive connections between microstructure and macroscopic properties.

Across assessment topics, the normalized gains for the WTL group exceed those of the non-WTL group. This difference in gains implies the presence of an additional source of learning beyond that of the traditional course components including lectures, recitations, homework, and exams. Given that this course employs traditional didactic pedagogy outside of the WTL component, we attribute the large gains to the learning acquired by engaging with the WTL process. In particular, we hypothesize that student ability to extrapolate and critically apply course content has been enhanced by participation in the entire WTL process, including draft writing, peer review, and revision.

Student perceptions of writing-to-learn

To examine student perceptions of WTL, we gathered feedback about all course elements, including the WTL assignments, at the midpoint and end of the term. For this analysis, we characterized self-reported textual input on learning and attitudes. Approximately half of the students report that the WTL assignments enhanced their learning, ranging from gaining a better understanding of the content to developing their writing ability. Over a third of the students discussed the benefits of the WTL assignments 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.

Many students discussed the role of the authentic scenarios of the WTL assignments in supporting their learning. In particular, many students mentioned that the authentic scenarios allowed them to apply the concepts they were learning in class, which in turn supported 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, student feedback responses provide additional evidence that the WTL assignments supported student learning. Broadly, students reported perceived learning gains from the WTL assignments, including the draft, peer review, and revision components. The authentic context successfully engaged students and led them to think more deeply about the targeted content. Additionally, for a subset of students, peer review and revision were perceived to be beneficial. These findings align with prior research on WTL in organic chemistry,[17,42] for which students similarly reported perceived learning benefits from engaging in writing, in particular discussing the beneficial roles of authentic contexts and peer review in supporting their learning.

Summary and outlook

In summary, we have evaluated the effectiveness of WTL assignments and their impact on student learning of foundational concepts in introductory MSE, especially those identified to be the most challenging. Using analyses of writing products in comparison with concept-inventory-style assessments, we addressed the following research questions: (1) do student descriptions of the targeted concepts improve from WTL assignment draft to revision?, (2) do students develop more robust understanding of those concepts?, and (3) which learning goals are best supported by the WTL approach? For all WTL topics, student concept descriptions improved from draft to revision, while students also developed more robust understanding of those concepts. To identify the learning goals that are best supported by the WTL approach, we compare the WTL effect sizes and assessment normalized gains across topics.

For the stress-strain and phase diagram WTL assignments that require students to synthesize qualitative data into quantitative formats, while emphasizing the connection between microscopic structure and macroscopic performance, the highest WTL effect sizes and medium-to-high gains on the corresponding assessment topics are observed. On the other hand, for the crystal structure and corrosion WTL assignments that emphasize crystal structure visualization and declarative knowledge, respectively, medium WTL effect sizes and low-to-medium gains on the corresponding assessment topics, are apparent. Since crystal structure visualization is a critical skill for understanding the influence of atomistic structure on macroscopic properties and the microstructure of materials plays a central role in corrosion mitigation, we suggest that future iterations of these and other WTL assignments resemble the stress–strain and phase diagram WTL assignments by including opportunities for students to identify meaningful structure–property correlations while providing a rigorous problem-solving scaffolding for processing and contextualizing quantitative data.

Our findings suggest that WTL pedagogies enhance student learning of concepts across multiple length-scales, including correlations between microscopic structure and macroscopic performance. Such multi-scale structure–property correlations are critical for several MSE-related fields, including chemistry, physics, mechanical engineering, and civil engineering. Our findings also indicate that WTL pedagogies enhance the ability of students to synthesize qualitative data into quantitative formats, a critical skill for STEM-related fields and beyond. Taken together, these findings suggest that WTL

pedagogies are likely to enhance student learning in STEM-related fields and beyond.

Finally, during recent course offerings (Winter 2020, Spring 2020, Winter 2021, and Spring 2021), not included in this study, an additional intervention from the Writing Fellows was added. For all draft submissions (following the peer review stage), Writing Fellows provide rigorous written feedback to the students. This additional intervention provides a scaffolded review process in which students receive directed feedback and reinforcement on how to better align with assignment expectations and goals. The effects of this additional intervention on student learning are currently under investigation.

Acknowledgments

Declarations

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1557/s43579-021-00114-z.

References

1. P.D. Klein, P. Boscolo, Trends in research on writing as learning activity. Journal of Writing Research **7**, 311 (2016)
2. J. Emig, Writing as a mode of learning. Coll. Compos. Commun. **28**, 122 (1977)
3. D. Galbraith, Conditions for discovery through writing. Instr. Sci. **21**, 45 (1992)
4. P. D. Klein, N. Arcon, S. Baker, in *Writing to learn*, ed. by C.A. MacArthur, S. Graham, J. Fitzgerald. Handbook of Writing Research, 2nd edn. (The Guilford Press, New York, 2015)
5. L.P. Rivard, A review of writing to learn in science: implications for practice and research. J. Res. Sci. Teach. **31**, 969 (1994)
6. A.R. Gere, N. Limlamai, E. Wilson, K. MacDougall Saylor, R. Pugh, Writing and conceptual learning in science: an analysis of assignments. Writ. Commun. **36**, 99 (2019)
7. J.A. Reynolds, C. Thaiss, W. Katkin, R.J. Thompson, Writing-to-learn in undergraduate science education: a community-based, conceptually driven approach. CBE Life Sci. Educ. **11**, 17 (2012)
8. J.H. Hanson, J.M. Williams, Using writing assignments to improve self-assessment and communication skills in an engineering statics course. J. Eng. Educ. **97**, 515 (2008)
9. E. Wheeler, R.L. McDonald, Writing in engineering courses. J. Eng. Educ. **89**, 481 (2000)
10. D. Ridgway, V. L. Young, and K. J. Sampson, Applying a writing to learn strategy in a traditional materials science and engineering course. *Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition*, pp. 10.211.1–10.211.7 (2005)
11. J. Bylander, M. Gustafsson, Improved content mastery and written communication through a lab-report assignment with peer review: an example from a quantum engineering course. Eur. Phys. Soc. **42**, 025701 (2021)
12. 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. J. Chem. Educ. **94**, 1610 (2017)
13. 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. J. Chem. Educ. **98**, 1548 (2021). https://doi.org/10.1021/acs.jchemed.0c01482
14. 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) (Pearson Prentice Hall, Upper Saddle River, 2007)
15. G.M. Bodner, Constructivism: a theory of knowledge. J. Chem. Educ. **63**, 873 (1986)
16. S.A. Finkenstaedt-Quinn, N. Polakowski, B. Gunderson, G.V. Shultz, A.R. Gere, Utilizing peer review and revision in STEM to support the development of conceptual knowledge through writing. Writ. Commun. **38**, 351 (2021). https://doi.org/10.1177/07410883211006038
17. T. Gupte, F.M. Watts, J.A. Schmidt-McCormack, I. Zaimi, A.R. Gere, G.V. Shultz, Students' meaningful learning experiences from participating in organic chemistry writing-to-learn activities. Chem. Educ. Res. Pract. **22**, 396–414 (2021)
18. 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. Chem. Educ. Res. Pract. **21**, 922 (2020)
19. A. Moon, E. Zotos, S. Finkenstaedt-Quinn, A.R. Gere, G. Shultz, Investigation of the role of writing-to-learn in promoting student understanding of light–matter interactions. Chem. Educ. Res. Pract. **19**, 807 (2018)
20. 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. Chem. Educ. Res. Pract. **20**, 383 (2019)
21. S.A. Finkenstaedt-Quinn, E.P. Snyder-White, M.C. Connor, A.R. Gere, G.V. Shultz, Characterizing peer review comments and revision from a writing-to-learn assignment focused on Lewis structures. J. Chem. Educ. **96**, 227 (2019)
22. F.M. Watts, J.A. Schmidt-McCormack, C.A. Wilhelm, A. Karlin, A. Sattar, B.C. Thompson, A.R. Gere, 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. Chem. Educ. Res. Pract. **21**, 1148 (2020)

23. S.M. Lord, Integrating effective 'writing to communicate" experiences in engineering courses: guidelines and examples.' Int. J. Eng. Ed. **25**, 196 (2009)

24. S. Krause, J. Kelly, A. Tasooji, J. Corkins, D. Baker, S. Purzer, Effect of pedagogy on conceptual change in an introductory materials science course. Int. J. Eng. Ed. **26**, 869 (2010)

25. W.D. Callister, D.G. Rethwisch, *Materials Science and Engineering: An Introduction*, 9th edn. (Wiley, Hoboken, 2014)

26. S. Krause, J.C. Decker, J. Niska, T. Alford, Identifying student misconceptions in introductory materials engineering classes. 2003 American Society for Engineering Education Annual Conference & Exposition pp. 8.648.1–8.648.8

27. A.F. Heckler, R. Rosenblatt, Student difficulties with basic concepts in introductory materials science engineering. 2011 Frontiers in Education (FIE) Conference, pp. S2H-1 (2011). https://doi.org/10.1109/FIE.2011.6142989

28. A.F. Heckler, R. Rosenblatt, Student understanding of atomic bonds and their relation to mechanical properties of metals in an introductory materials science engineering course. in 2020 Proceedings of the Annual Conference of the American Society of Engineering Education, pp. 15.1124.1–15.1124.11.

29. R. Rosenblatt, A.F. Heckler, Student understanding of the mechanical properties of metals in an introductory materials science engineering course. in 2010 Proceedings of the Annual Conference of the American Society of Engineering Education, pp. 15.1126.1–15.1126.13

30. D.V. Cicchetti, Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. **6**, 284 (1994)

31. K.A. Hallgren, Computing inter-rater reliability for observational data: an overview and tutorial. Tutor. Quant. Methods Psychol. **8**, 23 (2012)

32. W.W. Chin, B.L. Marcolin, P.R. Newsted, A partial least squares latent variable modeling approach for measuring interaction effects: results from a monte carlo simulation study and an electronic-mail emotion/adoption study. Inf. Syst. Res. **14**, 189 (2003)

33. C.H. Brase, C.P. Brase, Hypothesis testing. in *Understandable Statistics: Concepts and Methods*, 10th edn (Brooks/Cole Cengage Learning, Boston, 2012)

34. J. Cohen, in *The effect size index: d.* Statistical Power Analysis for the Behavioral Sciences (2nd ed.) (Lawrence Erlbaum Associates, New York, 1988).

35. J. Corkins, J. Kelly, D. Baker, S. Robinson Kurpius, A. Tasooji, S. Krause, Determining the factor structure of the materials-concept inventory. in *2009 American Society for Engineering Education Annual Conference & Exposition*, pp. 14.436.1–14.436.19

36. S.J. Krause, A. Sterling, J.E. Kelly, D. Stehlik, O. Isaacs-Sodeye, D.R. Baker, Development of a crystal spatial visualization survey for introductory materials classes. in 2012 American Society for Engineering Education Annual Conference & Exposition, pp. 25.449.1–25.449.10

37. S.N. Haynes, D.C.S. Richard, E.S. Kubany, Content validity in psychological assessments: a functional approach to concepts and methods. Psychol. Assess. **7**, 238 (1995)

38. R.R. Hake, Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. Am. J. Phys. **66**, 64 (1998)

39. V. Braun, V. Clarke, Using thematic analysis in psychology. Qual Res. Psychol. **3**, 77 (2006)

40. S. Heitkam, W. Drenckhan, J. Forhlich, Packing spheres tightly: influence of mechanical stability on close-packed sphere structures. Phys. Rev. Lett. **108**, 148302 (2012)

41. J. Kelley, K. Heinert, J. Triplett, D. Baker, S. Krause, Uncovering and repairing atomic bonding misconceptions with multimodal assessment of student understanding in an introductory materials course. in 2010 American Society for Engineering Education Annual Conference & Exposition, pp. 15.1289.1–15.1289.16

42. M.N. Petterson, S.A. Finkenstaedt-Quinn, A.R. Gere, G.V. Shultz, The role of authentic contexts and social elements in supporting organic chemistry students. Interactions with Writing-to-Learn Assignments, ASAP (2021)