Exam Wrappers, Reflection, and Student Performance in Engineering Mechanics

Dr. Ashraf Badir P.E., Florida Gulf Coast University

Dr. Badir is an Associate Professor in the Environmental and Civil Engineering Department at the U.A. Whitaker College of Engineering in Florida Gulf Coast University. He earned his B.Sc. (1982) in Civil Engineering and M.Sc. (1985) in Structural Engineering from Alexandria University, Egypt. He also holds a M.Sc. (1989) and a Ph.D. (1992) in Aerospace Engineering from Georgia Institute of Technology.

Dr. Jiehong Liao, Florida Gulf Coast University

Dr. Jiehong Liao is an Assistant Professor in the Department of Bioengineering at Florida Gulf Coast University (FGCU). She earned a B.S. in Biomedical Engineering from Rensselaer Polytechnic Institute (RPI) in 2004 as a Rensselaer Medalist and as a member of the inaugural class of Gates Gates Millennium Scholars. In 2011, she earned a Ph.D. in Bioengineering from Rice University. Before joining FGCU, she was a visiting Assistant Professor of Biotechnology in the Division of Science and Technology at the United International College (UIC) in Zhuhai China. Since her training with ASCE's Excellence in Civil Engineering Education (ExCEEd) initiative in 2016, she has been exploring and applying evidence-based strategies for instruction. In addition to the scholarship of teaching and learning, her research interests and collaborations are in the areas of biomaterials, tissue engineering, and regenerative medicine.

Dr. Tanya Kunberger, Florida Gulf Coast University

Dr. Kunberger is an Associate Professor in the Department of Environmental and Civil Engineering in the U. A. Whitaker College of Engineering at Florida Gulf Coast University. Dr. Kunberger received her B.C.E. and certificate in Geochemistry from the Georgia Institute of Technology and her M.S. and Ph.D. in Civil Engineering with a minor in Soil Science from North Carolina State University. Her areas of specialization are geotechnical and geo-environmental engineering. Educational areas of interest are self-efficacy and persistence in engineering and development of an interest in STEM topics in K-12 students.

Dr. Galen I. Papkov, Florida Gulf Coast University

Dr. Galen I. Papkov is an Associate Professor of Statistics in the Department of Mathematics at Florida Gulf Coast University. He received his Ph.D. in Statistics from Rice University in 2008, an M.S. in Applied Mathematics from CUNY Hunter College in 2002, and a B.S. in Mathematics and Psychology from SUNY College at Geneseo in 1998. Prior to entering academia, Dr. Papkov was an actuarial analyst in New York City. His primary research focus is in the field of nonparametric density estimation with applications in the areas of classification & discrimination, bump hunting, and change-point analysis. He also has experience and interests in design and analysis of surveys, multivariate analysis, regression modeling, and data mining.

Dr. Long D. Nguyen, Florida Gulf Coast University

Dr. Long Nguyen is an Associate Professor in the Department of Environmental and Civil Engineering at Florida Gulf Coast University (FGCU). Before joining FGCU, he was the deputy director of Tuan Le Construction and a lecturer at Bach Khoa University (BKU). Prior to his tenure at BKU, he worked as a construction consultant at Jax Kneppers Associates, Inc. in Walnut Creek, CA. He is a professional engineer registered in California. He earned his B.Eng. in Civil Engineering from BKU in 1999, M.Eng. in Construction Engineering and Management from Asian Institute of Technology (AIT) in 2003, and M.S. and Ph.D. in Engineering – Civil and Environmental Engineering from the University of California, Berkeley in 2005 and 2007, respectively.

Dr. Robert O'Neill P.E., Florida Gulf Coast University

Dr. ROBERT (BOB) J. O'NEILL is Professor and Chair of the Department of Environmental and Civil Engineering, U.A. Whitaker College of Engineering, Florida Gulf Coast University. He received a B.S. from the United States Military Academy in 1975, an M.S. in Structural Engineering and an M.S. in Geotechnical Engineering from Stanford University in 1984 and a Ph.D. in Structural Engineering from Kansas State University in 1993. Prior to his coming to FGCU he was a Professor of Engineering at Roger Williams University and an Associate Professor and Director of the Civil Engineering Analysis Group at the United States Military Academy. Dr. O' Neill is a retired Lieutenant Colonel, U.S. Army Corps of Engineers. He has been active at the national level with ASCE's Technical Council on Computing and Information Technology (TCCIT), Committee on Faculty Development (CFD) and Excellence in Civil Engineering Education (ExCEEd) initiative. Dr. O'Neill is a licensed Professional Engineer in California, Florida, Nevada and Virginia. He is a civil engineering program evaluator for the Accreditation Board for Engineering and Technology (ABET). He is an American Society of Civil Engineering Fellow (ASCE), a member of the American Society for Engineering Education (ASEE), and Phi Kappa Phi National Honor Society.

Exam Wrappers, Reflection and Student Performance in Engineering Mechanics

Abstract

This paper presents the authors' continuing study in investigating the efficacy of quizzes and homework in an engineering mechanics (Statics & Dynamics) course, starting from academic year 2013-2014. With these longitudinal data, our previous papers on this study concluded that homework did not significantly correlate with student performance on exams and that variations in homework methods had little impact on student performance in the class. Time surveys confirm that students often spend their time studying for the next quiz or exam. However, mistakes made on previous quizzes and exams are often repeated on later exams; suggesting students may not spend sufficient time in reflecting and improving their mastering of topics covered in previous quizzes and exams.

As such, the authors have implemented exam wrappers to obtain students' reflection on their quiz and early exam performance in current course offerings. Wrappers formalize the process of reviewing quiz and exam performance by asking students to formally record where mistakes may have occurred. Additionally, exam wrappers encourage reflection not only on exam performance, but also on how individuals prepared for the exam. Our premise is that the behavioral change from students, such as more reflection on the results of the formative assessment and clearer awareness of where mistakes occurred, may improve their performance on subsequent exams. In this paper, the authors will explore the relationship between wrappers and exam scores. The effectiveness of the wrappers will also be qualitatively assessed from a survey of students' perceptions at the end of each semester.

Introduction

Over the past 4 years, the authors have been collecting and reporting data on homework, quiz, and exam performance, as well as survey data on students' perceptions of learning and opinions on the methods used in the course delivery. Thus far, we have concluded that the use of traditional hand-written homework, frequent assessment via quizzes [1], or the Pearson Mastering Engineering [2] software for formative assessment did not have a significant impact on students' performance on exams. It was also observed that neither traditional nor online homework scores correlated well with exam scores; however, in-class quizzes did correlate with final exam scores. Most recently, using the Mastering Engineering Online system, specifically the inclusion of the Adaptive Follow-Up modules [3], it was observed that this also lacked any impact on overall student performance. In fact, adaptive follow-up in the Mastering Engineering system was seen as punitive by some of the students rather than as a resource to encourage mastery of the material [4].

In an attempt to take a different approach to improving student mastery of the course material since there has been no qualitative improvement with previous efforts, the instructors in the course have turned to posttest analysis. In contrast with previous interventions undertaken by this group of researchers, the current approach considers post exam activities in lieu of pre-exam modifications. Students were asked to reflect on their exam performance in conjunction with their study approach in order to engage in a metacognitive manner with their work even after it had been graded. The act of metacognition, or thinking about thinking, has been identified as a key approach to teaching and learning [5] and encourages students to adopt a growth mindset [6] by prompting them to identify personal actions that contribute to their performance on class assignments [7]. A formalized approach to this has been termed as a "test autopsy," "test postmortem," or (as the authors' prefer) an "exam wrapper" [8]. Developed by Lovett [9] an exam wrapper compels students to look more closely at their returned assignment by asking questions about how they prepared, where and why they earned or lost credit, and what they plan to do differently for future assignments.

We hope that by encouraging reflection after each assessment, students will gain awareness of their mistakes and come to realize how their attitude towards learning affects their performance. Furthermore, we hope that highlighting areas of improvement will help students make adjustments to optimize their performance, though it remains up to students to proactively take the steps for adjustment themselves. The overall theoretical framework for this study is illustrated in Figure 1.

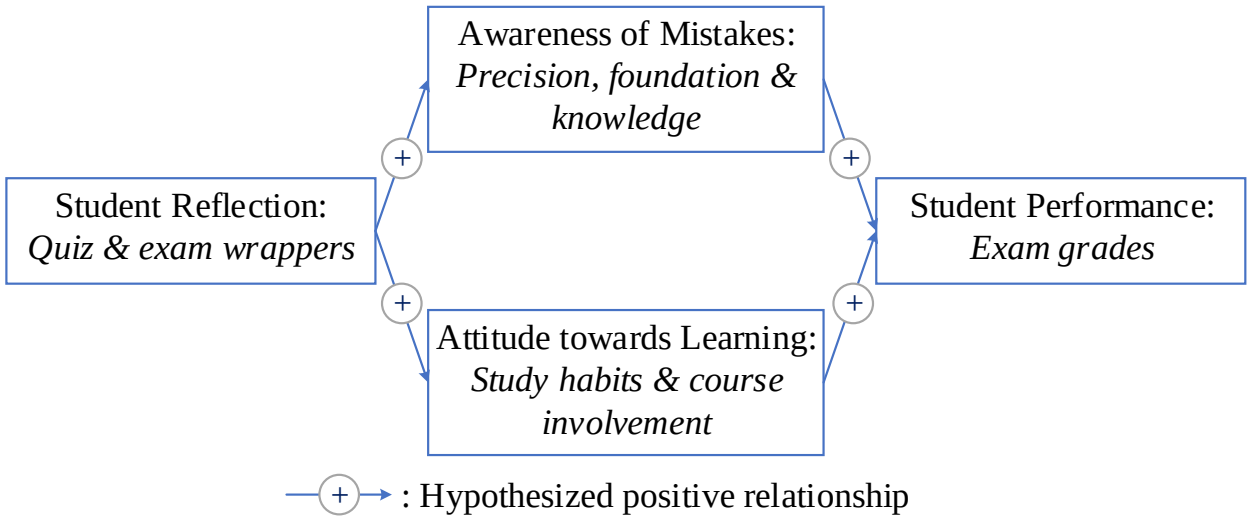

Figure 1. Theoretical framework for this study. Student reflection increases awareness of mistakes and serves to highlight attitudes that may need improvement. Making adjustments in these areas should lead to better performance.

Methods

Quiz and Exam Wrappers

Students in two sections of an Engineering Mechanics (Statics and Dynamics) course in Fall 2017 were given Wrappers to fill out with each graded assessment as a means to reflect on their mistakes, as well as attitudes and activities, leading up to the assessment. Quiz and Exam Wrapper questions are summarized in the Appendix (Tables A1 and A2, respectively). Quiz Wrappers were designed primarily to encourage student reflection, while Exam Wrappers further explored quantitative patterns in point loss, study habits, and course involvement. Students tallied their individual point losses in specific areas, shown in Table 1, and point losses pooled into each principal reason of point loss were used for analyses. Students also ranked their study activities and rated their course involvement as given in Table 2.

Table 1. Description of Principle Reasons for Point Loss

Principle Reasons	Specific Areas of Point Loss
Precision	Significant figures, lack of units, unit conversions, careless computation error, calculator issues, error or incomplete answer or format
Foundation	Algebraic substitution, use of simultaneous equations, issues with geometry or trigonometry
Knowledge	Unclear about terminology, error in constructing a free body diagram, uncertainty on how to approach the problem

Note: Students tallied their individual point losses in specific areas. Point losses pooled into each principle reason of point loss were used for analyses.

Table 2. Description of Attitude towards Learning

Inventory Categories	Components
Study Activities	Reading the textbook, reviewing lesson notes, looking over worksheet solutions, looking over homework solutions, reworking worksheets and homework, working new problems.
Course Involvement	Preparation for class, participation during class, engagement after class, seeking help outside of class.

Note: Students ranked their study activities from the most used to the least used (ranking from 6 to 1). Students rated their course involvement from excellent to poor (rating from 4 to 1).

Exit Survey

An Exit Survey was given at the end of the semester, in which students were asked whether they used the Quiz and Exam Wrappers to adjust their study habits, and if not, comment on why no adjustments were made. Reponses were tallied and converted to relative frequencies. Reasons for not making adjustments were analyzed for common themes.

Statistical Analysis

Data from Exam Wrappers for Statics Exam 1 and Statics Exam 2 were analyzed from two sections of Fall 2017. Overall, 93% of the students completed the Wrapper after Statics Exam 1, and 83% completed the Wrapper after Statics Exam 2. For each statistical question, Wrappers with incomplete responses were excluded from the analysis. In addition, mean exam scores from two sections in this current study of Fall 2017 with 83 students were compared with mean exam scores from two sections of Spring 2017 with 75 students and two sections of Fall 2106 with 75 students to investigate the effect of reflection via Quiz and Exam Wrappers on overall exam performance.

All data analyses were performed in the R statistical system [10] at the 5% level of significance. A mixture of parametric and nonparametric tests were utilized. In particular, the following tests were performed: Analysis of Variance, Fisher's Exact, Kruskal-Wallis, Permutation F, and two-sample t-tests. When appropriate, Bonferroni-adjusted Wilcoxon Rank Sum tests were conducted for post-hoc analyses.

Results

Diagnosing the Principle Reasons of Point Loss

In an effort to diagnose the underlying reasons for point loss, and hence, exam performance, individual student responses were analyzed from the Statics Exam 1 Wrapper. As students reviewed their graded exam, they tallied their total point loss in specific areas listed in a chart. Specific areas of point loss were categorized into the principle reasons of Precision, Foundation, and Knowledge for analyses (Table 1). A Kruskal-Wallis test confirmed our expectation that students lose more points in certain areas than others do (X^2=23.10, df=2, $p<0.001$). Pairwise comparisons via Wilcoxon Rank Sum tests, adjusted with a Bonferroni correction, showed that points tend to be lost more due to both the lack of Knowledge ($p<0.001$) and Precision ($p<0.001$) than due to the lack of Foundation. Figure 2 provides a visual representation of these differences with error bars corresponding to 95% confidence intervals. The results of this diagnostic suggest that interventions to help students improve on exam performance should be targeted towards, not only facilitating the mastery of content knowledge, but to also more rigorous training on precision and attention to detail.

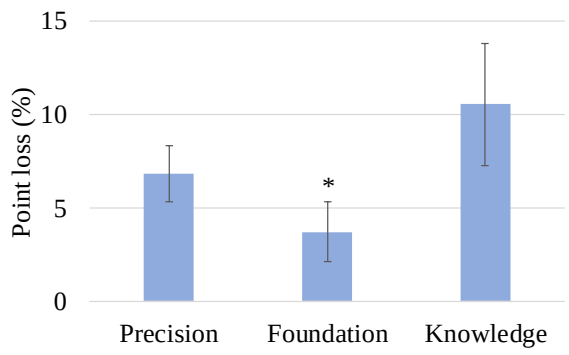

Figure 2. Principle reasons for point loss. Students lost more points due to a lack of precision and knowledge than a lack of foundation, as noted with (*).

Effect of Study Habits on Exam Performance

As part of the Statics Exam 2 Wrapper, students were asked whether they made adjustments in their study habits from Statics Exam 1 to Statics Exam 2. A Fisher's Exact test revealed a significant difference in the distribution of the perceived change in performance from Exam 1 to Exam 2 between those that adjusted their study habits and those that did not ($p=0.03$), as shown in Figure 3. Of the students who adjusted their study habits, over one-third (37.5%) reported that they performed much better on Exam 2, whereas this positive change in performance was not reported at all by those who did not make adjustments. Another noteworthy difference is that a higher proportion of students who did not make adjustments reported that they did much worse on Exam 2 than those who adjusted their study habits. This difference in distribution suggests that if students are presented with the opportunity to reflect on their performance, those that proactively take the steps for adjustment tend to perceive better outcomes. Future Exam Wrappers designed to track students from Exam 1 to Exam 2 could serve to quantitatively confirm this perceived improvement in exam performance.

We further delved into the relationship between study activities and exam performance using the Statics Exam 2 Wrapper by asking students to rank their study activities from the most used to the least used (Table 2). Although a Permutation F test showed no significant difference in rankings across study activities between those who reported better performance on Exam 2 than those who reported worse performance ($p=0.26$), there was an overall significant difference in rankings across study activities ($p<0.001$). Pairwise comparisons via Wilcoxon Rank Sum tests, adjusted with a Bonferroni correction, showed that all other study activities were used more than reading the textbook ($p<0.002$). Furthermore, reworking worksheets and homework problems as well as looking over worksheet solutions were employed more than reviewing lesson notes ($p<0.002$). Reworking problems is also more preferred than looking over homework solutions ($p<0.001$). Figure 4 provides a visual representation of these differences with error bars corresponding to 95% confidence intervals. The results of this inventory show a pattern of study habits consistent with our expectations for students to perform well in this course. However, since the textbook does not seem to be a popular resource, other less costly references should be considered.

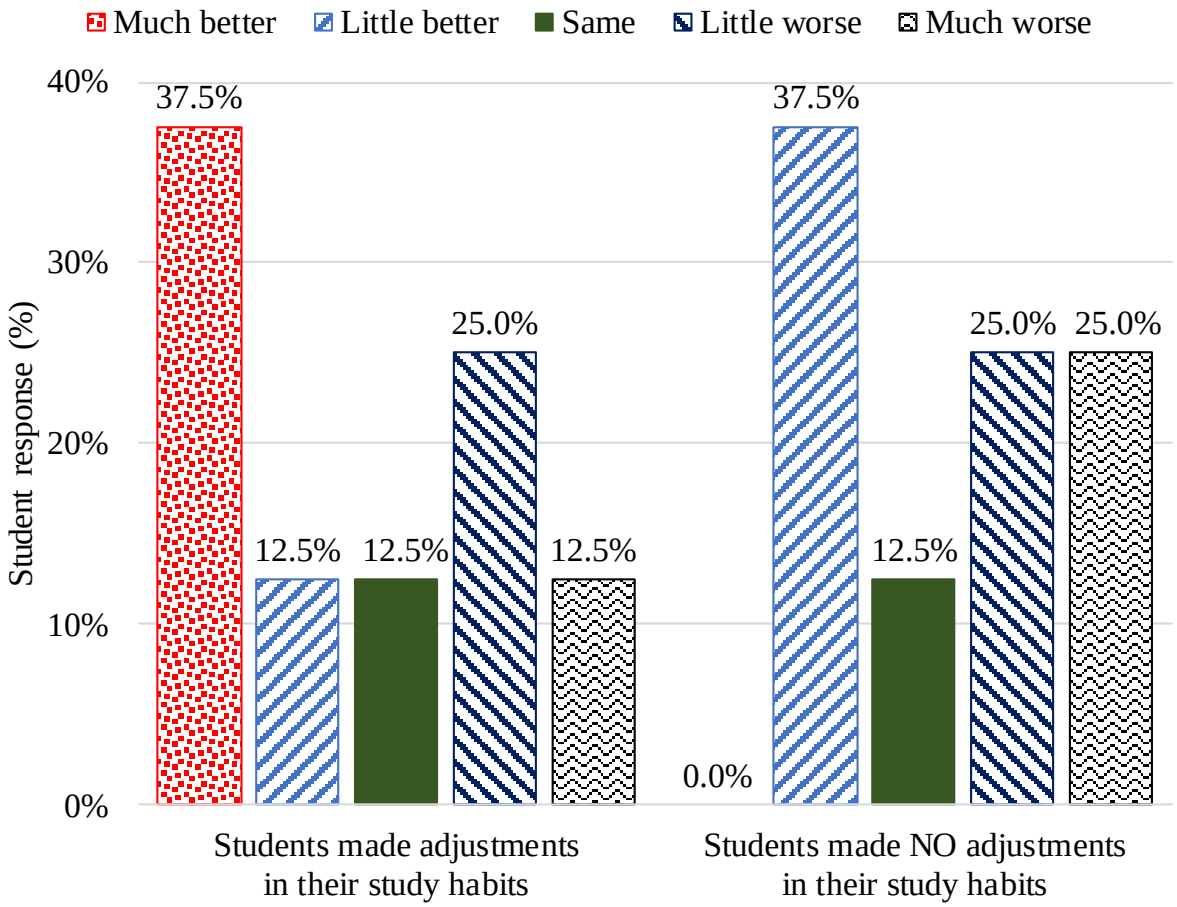

Figure 3. Student perception of exam performance. Students who made adjustments in their study habits reported much better exam performance.

Interestingly, a T test revealed that students who reported a positive change in performance from Statics Exam 1 to Statics Exam 2, recorded a significantly higher proportion of time (71.43%) studying on their own individually than those who reported worse exam performance ($t=2.026$, $df=33$, $p=0.025$). Another noteworthy finding, also through a T test, showed that students who adjusted their study habits from Exam 1 to Exam 2, recorded a significantly lower proportion of time (28.63%) studying within the 24 hours prior to the exam than those who did not make adjustments in their study habits ($t=-1.748$, $df=38.9$, $p=0.044$). Taken together, these differences in proportions of study time suggest better performance outcomes when studying individually outside of class (rather than in a study group), and that those who proactively take the steps for adjustment tend to "cram" less in the day before the exam (hence study ahead of time).

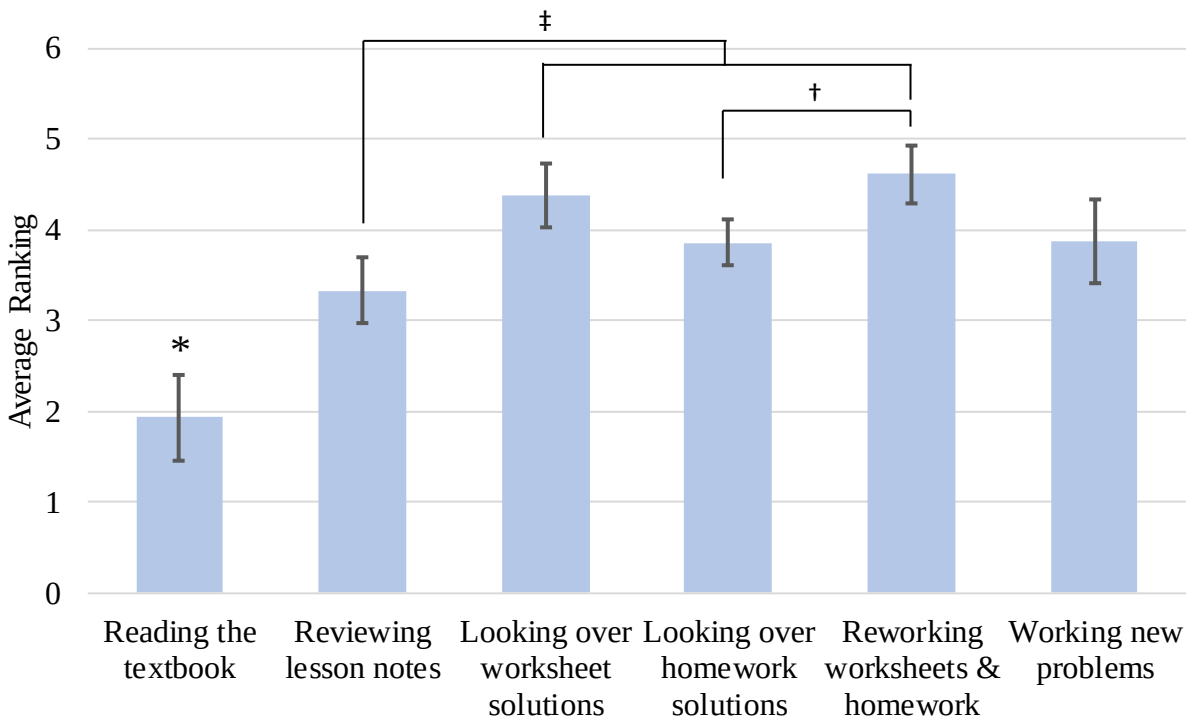

Figure 4. Average ranking of study activities. Rankings from the least used (1) to the most used (6). Reading the textbook was the least used study activity, as noted with (*). Study activities grouped in brackets are significantly different from each other. Looking over worksheet solutions and reworking worksheets and homework were used more as study activities than reviewing lesson notes (‡). Reworking worksheets and homework was also more used than looking over homework solutions (†).

Effect of Course Involvement on Exam Performance

We investigated the relationship between course involvement and exam performance using the Statics Exam 2 Wrapper by asking students to rate their level of involvement before, during, and after class (Table 2). Although a Permutation F test showed no significant difference in rankings across course involvement between those who reported better performance on Exam 2 than those who reported worse performance ($p=0.258$), there was an overall significant difference in levels across course involvement ($p<0.001$). Pairwise comparisons via Wilcoxon Rank Sum tests, adjusted with a Bonferroni correction, showed that students reported higher levels of preparation for class, participation in class, and engagement in course material after class than seeking extra help outside of class ($p<0.001$). Furthermore, students rated their level of preparation for class less than their level of participation in class ($p=0.0025$). Figure 5 provides a visual representation of these differences with error bars corresponding to 95% confidence intervals. The results of this inventory show a pattern of course involvement that could use improvement. Specifically, encouraging students to seek help outside of class as well as providing incentives for students to

come to class better prepared for the lessons would potentially help students perform better in this course.

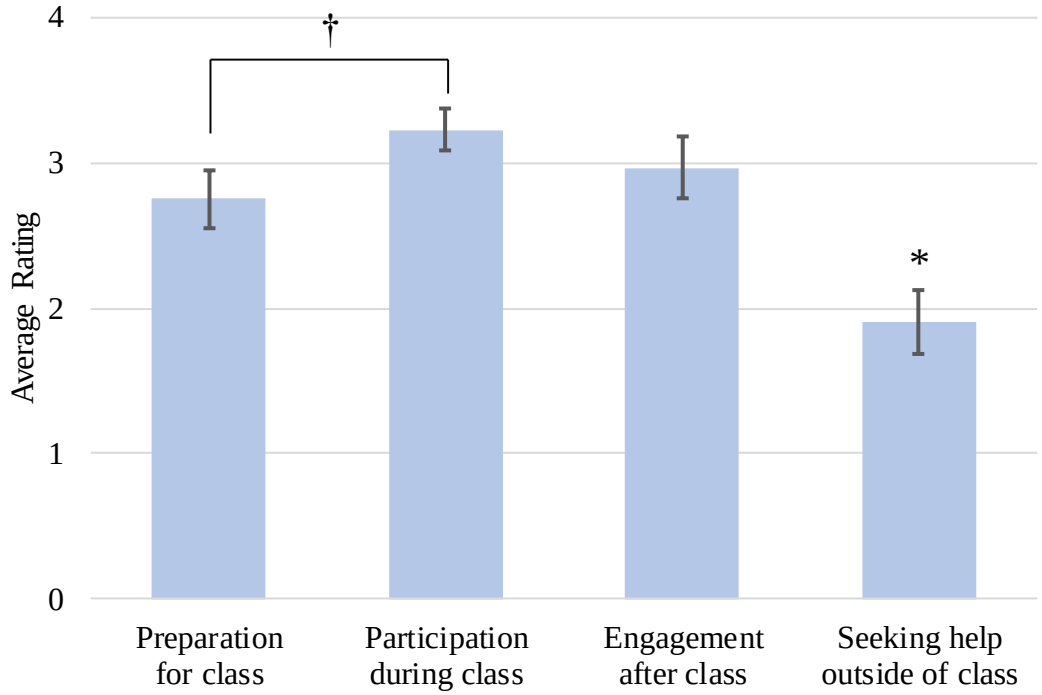

Figure 5. Average rating of course involvement. Ratings from poor (1), fair (2), good (3), to excellent (4). Seeking help outside of class was rated significantly lower than all other aspects of involvement, as noted with (*). Student participation in class was rated significantly higher than preparation for class, as noted with (†).

Reflection via Quiz/Exam Wrappers to Make Adjustments

In an optional Exit Survey given at the end of the semester, 38 respondents across both sections commented on their use of Quiz and Exam Wrappers to adjust their study habits. Survey results showed that 60.5% of respondents did use Wrappers to make adjustments and commented on how the Wrappers highlighted areas in which they could improve. Although 15.8% did not make adjustments in their study habits, these students commented that they were satisfied with their performance and felt that adjustments were not necessary. Reasons for why other students did not make adjustments varied or were not given. Overall, having students fill out Quiz and Exam Wrappers did seem to foster reflection and adjustment in most participants.

Mean scores on Statics Exam 1 were compared across three semesters of cohorts to establish a baseline for comparison. An Analysis of Variance on Exam 1 scores showed no statistical difference between students entering this course (X^2=3.884, df=2, p=0.1434). However, having students reflect on their course involvement, study habits, and reasons for point loss, did not

result in higher scores on Statics Exam 2 (X^2=1.801, df=2, p=0.4063), nor an improvement in scores from Statics Exam 1 to Exam 2 (F=0.785, df_N=2, df_D=179, p=0.4577). Therefore, encouraging students to reflect on their involvement, habits, and sources of error, did not have a significant impact on overall exam performance in this current study.

Conclusions

In an effort to have students consider their performance on quizzes and exams more robustly than simply the earned grade, this study implemented a metacognitive exercise called exam wrappers in the sophomore level Engineering Mechanics course. Results from the first semester of implementation of this study showed that exam wrappers did not have a significant impact on students' final course grades. Although exam wrappers did not seem to increase exam scores and performance; overall, having students fill out quiz and exam wrappers did seem to foster reflection and adjustment in most participants. Over one-third of the students, who adjusted their study habits, reported that they performed much better in subsequent exams. Students who did not make adjustments have not reported this positive change in performance. This difference suggests that if students are presented with the opportunity to reflect on their performance, those that proactively take the steps for adjustment tend to perceive better outcomes.

It should be noted that in this study, faculty members did not spend much formal (in class) or informal (office hours) time emphasizing the importance of specific study habits and/or ensuring the awareness of the underlying sources of exam/quizzes mistakes. The lack of referring the students back to earlier reflection, either on a recommended or required basis, might have restricted the achievement of the desired effect and is a potential area for future research modifications.

References

1. Lura D.J., Badir A., and O'Neill R.J., "Homework Methods in Engineering Mechanics," the 122nd American Society for Engineering Education (ASEE) Annual Conference & Exposition, Seattle WA, June 14-17, 2015.
2. O'Neill R.J., Badir A., Nguyen L.D., and Lura D.J., "Homework Methods in Engineering Mechanics: Part Two," the 123rd American Society for Engineering Education (ASEE) Annual Conference & Exposition, New Orleans LA, June 26-29, 2016.
3. Pearson Inc. Overview: Adaptive Follow-Up assignments, Online: https://help.pearsoncmg.com/mastering/instructor/ccng/Topics/afu_overview.htm Accessed February 8, 2017.
4. Lura, D.J., O'Neill, R.J., Badir, A. and Nguyen, L.D., "Homework Methods in Engineering Mechanics: Part Three," 124th American Society for Engineering Education (ASEE) Annual Conference & Exposition, Columbus, OH, June 25 – 28, 2017.
5. Ambrose, S.A., Bridges, M.W., DiPietro, M., Lovett, M.C. and Norman, M.K. (2010). *How Learning Works: Seven Research-Based Principles for Smart Teaching*. San Francisco, CA: Jossey Bass.

6. Dweck, C.S. (2006). *Mindset: The New Psychology of Success*. New York, NY: Random House.

7. McGuire, S.Y., and McGuire, S. (2015). *Teach Students How to Learn*. Sterling, VA: Stylus.

8. Nilson, L.B. (2013). *Creating Self-Regulated Learners*. Sterling, VA: Stylus.

9. Lovett, M.C. (2013). Making Exams Worth More Than the Grade. In M. Kaplan, N. Silver, D. LaVaque-Manty and D. Meizlish (eds). *Using Reflection and Metacognition to Improve Student Learning: Across the Disciplines, Across the Academy*.pp. 18 – 48. Sterling, VA: Stylus.

10. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

APPENDIX

Table A1. Quiz Wrapper Questions.

1. Approximately how many hours did you spend in total preparing for this quiz?
2. What percentage of this time was in the 24 hours prior to the quiz?
3. What did you consider your level of preparation for the quiz:
    a. Excellent (participated in class, completed homework, solved / re-solved problems)
    b. Good (attended class, completed homework, looked over worksheet solutions)
    c. Fair (attended class, looked over work)
    d. Poor (took quiz)
4. How many points did you lose in total on this quiz?
5. What would you consider the primary reason you lost points?

Table A2. Exam Wrapper Questions.

1. Approximately how many hours did you spend in total preparing for this exam?
2. What percentage of this time was in the 24 hours prior to the exam?
3. What percentage of your preparation time was individual rather than group?
4. Rank the following activities used while studying (6 – most used, 1 – least used, N/A – not used): reading the textbook, reviewing lesson notes, looking over worksheet solutions, looking over homework solutions, reworking worksheets and homework, working new problems, other (please explain)
5. How productive do you feel the majority of your study time was?
    a. Highly productive (reading text, completing homework, solving problems)
    b. Productive (reviewing notes, looking over solutions, identifying problem approaches)
    c. Non-productive (locating items, chatting with group, "spinning wheels" on problems)
6. How many classes did you miss for lessons covered on this exam?
7. How many homework sets did you complete for material covered on this exam?
8. How often did you seek help from the instructor out of class? (often, sometimes, rarely, never)
9. What do you consider your typical level of preparation for class:
    a. Excellent (read Canvas page; studied textbook; printed worksheets before class)
    b. Good (skimmed Canvas page; looked over worksheets before class)
    c. Fair (aware of lesson topic before class)
    d. Poor (arrived in class)
10. What do you consider your typical level of participation in class:
    a. Excellent (copied board notes and added your own side notes, involved in class discussions by asking and answering questions, took the lead in group work)
    b. Good (copied board notes, listened to class discussions, helped with group work)
    c. Fair (copied board notes, watched group work)
    d. Poor (copied board notes)

11. What do you consider your typical level of engagement after class:
    a. Excellent (reviewed board notes and added your own side notes, studied textbook, reworked in-class problems on your own)
    b. Good (skimmed board notes, looked over in-class problems)
    c. Fair (skimmed board notes)
    d. Poor (did not review)
12. How many points did you lose in total on this exam? Of those points, how many were due to the following (must add to total points lost):
    a. Units, answer format, significant figures, etc.
    b. Careless computational error, calculator issues
    c. Other mathematical errors (e.g. unit conversions)
    d. Substitution or the use of simultaneous equations
    e. Geometry and/or trigonometry
    f. Terminology or constructing a correct free body diagram
    g. Uncertainty on how to approach the problem
    h. Other (please describe)
13. What would you consider the three most valuable activities that contributed to a strong performance in the exam? (i.e. What are the best activities to help you prepare for the final exam?)

Follow-up questions on Statics Exam 2 Wrapper

14. How was your performance on the second exam compared to the first exam? (significantly better, little better, about the same, little worse, significantly worse)
15. Did you make any changes to your study habits from the first exam to the second exam?
    a. If YES – What primary changes did you make in your approach to studying for this exam?
    b. If NO – Why did you not make changes to your approach to studying for this exam?
16. Did you make any changes to your involvement with class material in general (examples provided in questions 6 – 9) from the first exam to the second exam?
    a. If YES – What primary changes did you make in your class involvement?
    b. If NO – Why did you not make changes to your class involvement?