The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #42424

Student Performance on Statics Problems in Deformable Solids: A Look at Pre- and Post-Test Results

Dr. Amie Baisley, University of Florida

I am an Instructional Assistant Professor at the University of Florida teaching primarily 2nd year mechanics courses. My teaching and research interests are alternative pedagogies, mastery-based learning and assessment, student persistence in their first two years, and faculty development.

Prof. Keith D. Hjelmstad, Arizona State University

Keith D. Hjelmstad is President's Professor of Civil Engineering in the School of Sustainable Engineering and the Built Environment at Arizona State University.

Student Performance on Statics Problems in Deformable Solids:
A Look at Pre- and Post-Test Results

Deformable Solids is a course that students generally take the semester immediately following the prerequisite course on *Statics* and builds directly on it. The tacit expectation in *Deformable Solids* is that students have a solid grasp of statics concepts and should be able to solve a variety of statics problems. Over fourteen semesters, we have measured student ability to solve statics problems using a diagnostic test, which is administered in *Deformable Solids* as a pre-test at the start of the semester and a post-test near the end of the semester. The same test is used for both pre- and post-tests.

The diagnostic test consists of eight multiple choice statics questions. The problems are not concept questions, but more like problems they would be asked to solve in *Statics*. The students are given as much time as needed to complete the test and they are encouraged to work through the problems with all necessary steps to get the correct answer, including drawing a free body diagram and performing necessary calculations. In a multiple-choice format, the test is graded only for the final answer and, of course, multiple-choice test-taking strategies can be employed.

The results from the pre-test have been consistent over the years with students on average answering roughly three of eight questions correctly. The post-test results show an improvement of roughly one additional question correct, influenced mostly by the types of problems on the test most similar to the problems encountered in the *Deformable Solids* course. The degree of success is similar to the results of other multiple-choice tests like concept inventories, suggesting that the raw score is not simply an indicator of statics knowledge, but also measures student attitudes about taking multiple-choice tests (e.g., rushing through the test, sloppy calculations, and guessing). Over the past few semesters, we have collected information about each student's approach to the test to help understand the effect of the latter factor on the outcome of the test. The increase in performance at the end of *Deformable Solids* suggests that the development of solid conceptual understanding of the abstract concepts of statics takes students more than a semester or year to achieve.

This paper reports the results of the pre- and post-test from 14 semesters, along with the students' different approaches to the tests. This study gives a detailed account of the testing instrument and testing conditions, it sheds light on pre- and post-testing as a means of understanding student knowledge of statics, and it suggests reasonable expectations on performance given student attitudes toward testing.

Introduction

The desire to understand how students are doing in a course in an easily accessible format has encouraged the continued use of multiple-choice tests in many courses. Specifically, in engineering courses that are heavy on problem-solving, the use of multiple-choice tests still happen as they are a route to offer quick, quantifiable results about *something* in that course. However, exactly what that something is and what the results represent has been debated for many years.

Engineering has used multiple-choice tests in a variety of contexts. These include concept inventory tests, course exams, and diagnostic testing. Multiple-choice tests (MCT), if the answers are well written to include quality distractors, can be a valuable form of assessment in STEM courses [1]. There has been a high correlation between MCT and student written responses to instructor exams if the exams are well written [2]. Even in the context of standardized testing it has been found that student GRE scores compared to student written responses had a high correlation between the results [3]. Multiple-choice tests can be valid assessment instruments if written correctly, which has led to many concept inventories being created in STEM, like the Mechanics Diagnostic Test, Force Concept Inventory, Statics Concept Inventory, Dynamics Concept Inventory, and many others [2, 4, 5].

Often MCT are used as pre-/post-tests to try to identify changes in learning. The quantitative results of these multiple-choice tests provide easy comparison data when looked at from a pre-/post-test analysis, but the scores do not always adequately reflect a learning change because there are also many external factors that must be considered [6]. There have been positive changes in student performance on pre/post-test on the Statics Concept Inventory over the course of a semester but identifying what the students have gained is still a challenge [4]. Even when the testing happens more regularly throughout the semester it has been found that students do not change a lot in their conceptual understanding of physics, but they have small improvements on their answers to the MCT [7]. Another study found that in the initial testing for a dynamics concept inventory test, the students answered around 32% of the questions correctly initially and had improvement when taken again at the end of the semester [5]. Other studies on concept inventories found that fewer than 50% of students answered a statics concept inventory correctly [8]. These results show that on average the scores are low for these tests, and while there is improvement over a semester it is not clear what knowledge the students have gained or if they have simply improved at taking the test. When students are asked to expand on their responses to concept inventories, valuable information becomes clear as to why the students are performing the way they are on the test. For example, in a study using the statics concept inventory found that most students performed poorly because of their lack of skills needed to draw correct free body diagrams [9].

The course *Introduction to Deformable Solids* (often called *Mechanics of Materials* or some similar name) is the second course in a series of mechanics courses that most students take in mechanical, aerospace, civil and similar engineering curricula. The content scaffolds directly on what students learn in the prerequisite course *Statics*. A student who passes *Statics* should be proficient in the material from that course and should be able to solve any *Statics* problem given the chance. The conceptual knowledge in *Statics* has been documented in an established framework that the subject of deformable solids builds off of [10]. To test the level facility with solving *Statics* problems upon completion of the course, we started administering a multiple-choice diagnostic pre-test on statics concepts, created by the instructor, at the beginning of our course on deformable solids to assess readiness for the course, to identify areas where students would benefit from review, and to find topics that might benefit from reinforcement through additional course materials.

In the first few semesters, we observed that the scores were lower than expected. Realizing that we had no real basis for our performance expectations, we embarked on a broader study that included administering the test again at the end of *Deformable Solids* in an attempt to quantify learning gains in *Statics* that accrue from continuing to practice statics throughout the second course. The results of these tests offer interesting insight into how the students take the test and how to interpret their performance.

Pre-test/post-test environment

The multiple-choice diagnostic test is given to students on the first or second day of the semester in deformable solids. The students are informed that they will be given a statics pre-test through a course announcement several days before the class starts and are encouraged to brush up on their statics skills. The students are given an entire 75-minute class period to do the test and they are encouraged to spend the full amount of time and use as much scratch paper as needed to solve the problems. The answers are submitted through a Google Form, but students are also asked to submit their written work through the LMS as a means of encouraging them to actually solve the problems using the methods they learned in *Statics*.

There are eight questions that cover a broad range of problems students are likely to have seen in *Statics*. The questions are not concept questions, but rather set-up and solve type problems similar to those they did when taking *Statics*—problems that require drawing a free body diagram, writing equations of equilibrium, and solving for unknowns. At a minimum, a free body diagram and a force or moment balance equation should be used to find the correct answer. However, for ease of grading and due to the diagnostic nature, the test is a multiple-choice test with five answer options for each question. The students receive participation points in the course for how well they do on the test and are encouraged to brush up on the material covered if they found it difficult.

In the last week of the class, the students are given the same eight questions as a diagnostic post-test. The students know the post-test is coming at the end of the semester because it is on the course schedule, but they are not told whether it will be the same or different than the pre-test. During the post-test the students have another 75-minute class period to work and are encouraged to work out each problem as much as possible.

For both the pre-test and post-test, there are students that finish very quickly and there are students that take the entire class period. In recent years, the student work for the tests have been collected along with the results from the test to see if their work will provide better indication of what the students know (and to encourage good statics habits, as mentioned previously).

The test problems

The eight multiple choice questions are shown in Fig. 1 along with the multiple-choice options. The range of questions is representative of the types of problems generally encountered in a *Statics* course. The first three are basic reaction calculations, the fourth is a very simple truss problem, the fifth is a hydrostatics problem, the sixth is a beam reaction problem, the seventh is a more complex truss problem, and the eighth is a beam with a distributed load.

All problems have a ‘none of the above’ option to discourage guessing and encourage setting up and solving the problems. The problem statements are very brief, requiring a modest amount of interpretation of standard terminology and sketches. Students are permitted to ask questions about the problem statements to avoid errors due to misinterpretation of the problem. Experience shows that very few students ask questions of clarification.

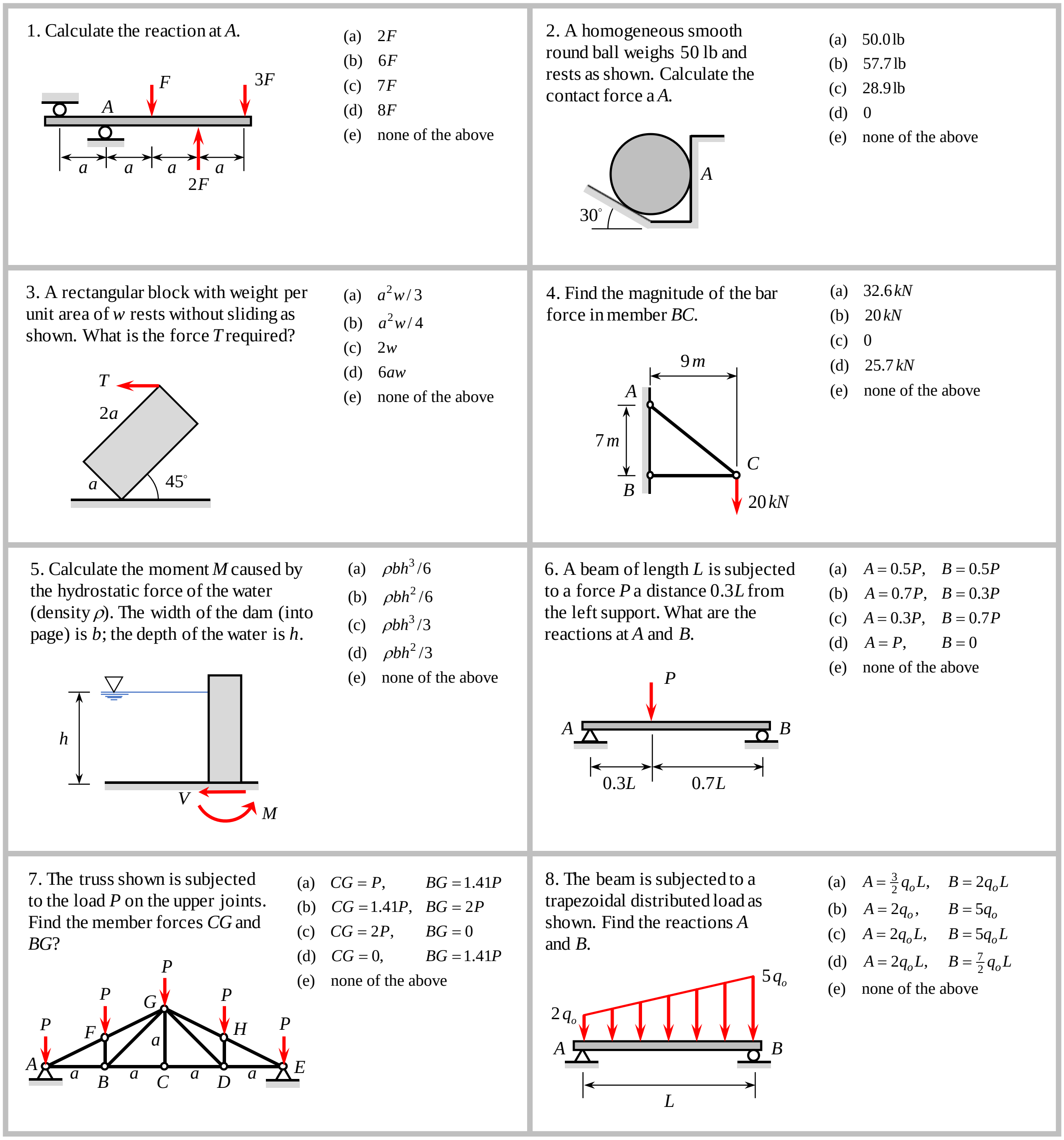

Fig. 1. Eight multiple-choice test questions for a *Statics* pre-test and post-test, administered at the beginning and end of the course *Introduction to Deformable Solids*

The first problem has a very attractive distractor (answer ‘a’) that catches students who ignore the support not called out in the problem statement. It is common for students to select the distractor because they can sum forces in their head to see an unbalance of $2F$. That choice, of course, ignores the fact that there are two reaction forces present, not just the one asked for in the question. The other problems have distractors, but none as attractive as the one in the first problem. A reduced subset of possible answers to almost all the problems can be deduced from a units argument (e.g., options 'c' and 'd' in problem 3 and option 'b' in problem 8 can be excluded because they do not have units of force, and options 'b' and 'd' in problem 5 can be excluded because they do not have units of moment). Option 'd' in problem 7 is the only one that has *CG* as a zero-force-member. Option 'a' in problem 8 is the only one for which the reactions sum to the total load. Conceptual understanding of statics is an advantage, but with option 'e' available, it is not sufficient to completely solve the test. Comments from students suggest that there is a wide range of risk tolerance among test takers, and some are willing to guess when the number of reasonable options is reduced, even it if is possible that the answer is 'e.'

The pre-/post-test has been given over 14 semesters with results from SP17 through FA23 (the pre-test was given for a few semesters prior to that without a post-test). The number of students varied each semester with the most registered students in the spring terms and about half the number of registered students in the fall terms. The number of students in the fall each semester ranged from 31 to 65 students and in the spring semester the number ranged from 45 to 94 students. There was a total of 814 students that took both the pre-test and the post-test during this timeframe (we eliminated students who took only one due to absence at the beginning of the semester of dropping the course before the end of the semester). The semesters affected by COVID-19 still took the same pre-test and post-test but it was given over Zoom (similar to all other exams) with the same 75-minute class period to complete it.

Test results on average

The tests were scored for correct answers, and the result was tracked for each test each semester. Figure 2 and Table 1 give the results for the average number of correct answers for the pre- and post-test for all 14 semesters. Table 1 gives a breakdown of the average number of correct answers for each semester.

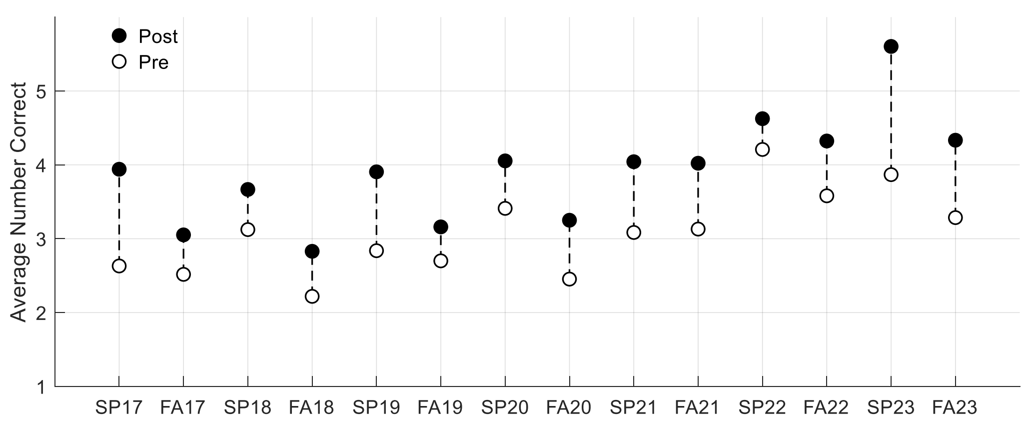

Fig. 2. Gains from pre- to post-test by semester

Table 1. Average number correct on eight pre- and post-test questions

Semester	# Students	Average Number Correct		Percent Increase
		Pre-test	Post-test	
SP17	84	2.63	3.94	49.8
FA17	58	2.52	3.05	21.2
SP18	81	3.12	3.67	17.4
FA18	41	2.22	2.83	27.5
SP19	74	2.84	3.91	37.6
FA19	50	2.70	3.16	17.0
SP20	56	3.41	4.05	18.8
FA20	64	2.45	3.25	32.5
SP21	71	3.08	4.04	31.1
FA21	46	3.13	4.02	28.5
SP22	48	4.21	4.62	9.9
FA22	31	3.58	4.32	20.7
SP23	68	3.87	5.60	44.9
FA23	42	3.29	4.33	31.9
All Semesters	814	3.05	3.92	28.7

On average, students answer about 3 questions correctly on the pre-test and almost 4 questions correctly on the post-test. There was approximately a full question gain from the pre- to the post-test (a gain of about 30%). The graph and table show that there was improvement from the pre-test to the post-test for every semester, with some having larger improvements than others. The performance average of between 40 and 50 percent is consistent with observation on other MCT testing instruments [5,8]. The histogram of all students in the sample is shown in Fig. 3.

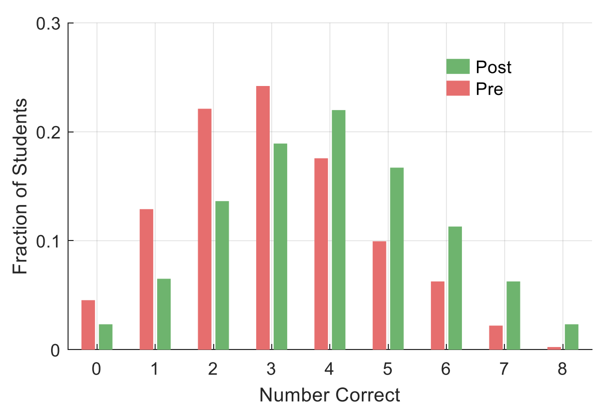

Fig. 3. Histogram of the number of the fraction of students vs. the number of questions answered correctly for the pre- and post-test ($N = 814$).

Figure 3 shows that the data are very close to being normally distributed, with a fairly broad spread. It is interesting to note that there are students who take the test and get zero correct. That is true even for the post-test. It is also worth noting that the incidence of students getting a perfect score of 8 is very rare—almost zero for the pre-test but still very low for the post-test.

Figure 2 and Table 1 show two interesting trends. First, it appears that many of the fall semesters start with a lower average number of correct answers and/or have smaller number of gains than many of the spring semesters. An explanation for this could be due to the spring semesters having the students that are typically on-track in their curriculum progress. *Deformable Solids* is typically scheduled in the spring semester following *Statics* in the fall semester for students that are following the 4-year plan for the program. The *Statics* class in the spring includes both students who are behind, usually due to needing a remedial math course, and those who are ahead, usually those who come to college with advance placement credit. Also, the fall semester follows a longer break (summer) than the spring semester does.

In the written comments, students often say that they have forgotten how to do some kinds of problems (e.g., truss problems or hydrostatics problems). This observation supports the hypothesis that student knowledge of statics after the semester they take the course is still fragile. That is probably especially true for students who passed statics but were at the lower end of the mastery spectrum. Another interesting trend is that the results have been improving in the most recent semesters. One reason for this might be general improvements in the learning environment over time. The post-test results for the SP23 semester seem unusually high and should probably be considered an outlier. The outcomes from the past are used to inform teaching strategies going forward.

Test results broken down by question

The problems on the test cover many different *Statics* topics with each one having a varying level of difficulty. The percentage of students that got each question correct was tracked for each semester. The results are shown in Fig. 4.

The results for each problem are generally in the same range for all semesters showing that student responses are consistent across the years. Problems 3, 5, and 7 are consistently low on correct responses, while problem 2 and 6 tend to have many more students answer it correctly. Problem 1 and 8 show an alternating pattern that varies between the spring and fall semesters which is a similar trend to what was seen across the entire test previously in Fig. 2. These results can be explored further by looking at the exact problems provided in Fig. 1 to understand what students are answering correctly or not.

The problems that have the most incorrect answers are:

- Problem 3, a block problem with distributed weight, which requires moment equilibrium to answer,
- Problem 5, a hydrostatics problem, also requiring moment equilibrium, and
- Problem 7, a truss problem where knowledge of zero-force members aids recognizing the correct answer.

Problems 3 and 5 are challenging problems that require students to find the resultant of a distributed load and apply moment equilibrium correctly. The only other problem with a distributed load is problem 8, which had varying results across all semesters. This could imply that students are not comfortable with distributed loads after statics or that students did not take the necessary steps to answer these questions on the test. The low results on Problem 7 reveal that students do not fully understand zero-force members for trusses or are less willing to engage in a method-of-sections calculation.

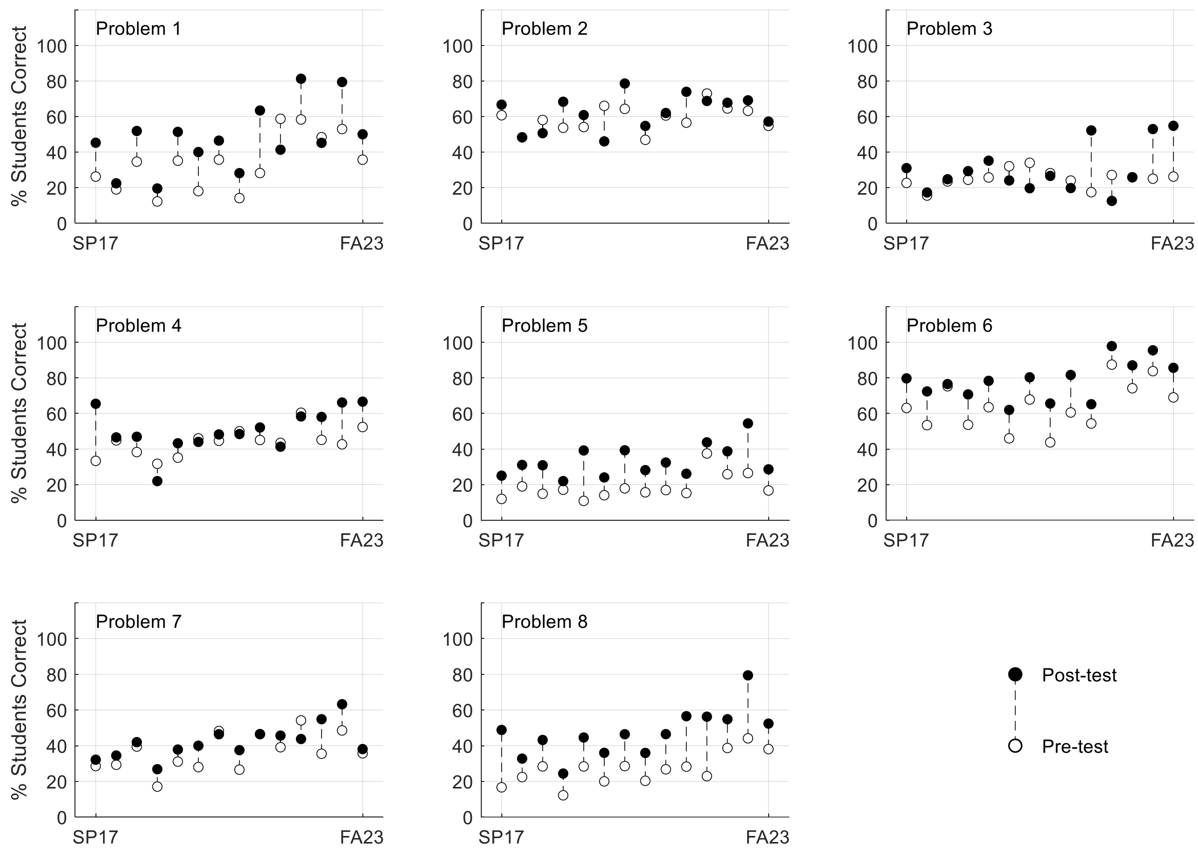

Fig. 4. Gains for pre- and post-tests over each semester by organized by problem.

The problems with the highest number of correct answers include problem 2, the ball problem, and problem 6, a simply supported beam problem. Problem 2 is similar to problems that the students may have seen in physics, so it is more familiar, and they understand how to deal with smooth contact surfaces since it has been a topic of multiple semesters. The simply supported beam requires moment equilibrium to solve it, but it is a common statics problem that students have a lot of experience with. The students are more successful on these problems because both problem 2 and problem 6 are repeated topics across multiple courses that students likely feel more comfortable with than problems that include individual course topics like trusses or hydrostatic pressure.

The gains for each problem from pre-test to post-test were also variable. For problem 3, 4, and 7 the gains were very small for many of the semesters. This is likely the result that deformable solids problems did not involve many truss or block problems, so the students did not have any more experience solving these during the semester. However, the gains for problem 1, 6, and 8 are much larger from the pre- to post-test. These three problems are beam problems which are like many of the problems solved in deformable solids. The students have had more practice with these problem types, and that is reflected in their improvement in answering those questions.

The problem-specific results indicate that students are likely to improve on problems that they have more practice on over multiple courses. If the topic is seen in multiple courses or one that they practice in deformable solids, the results reflected that. Also, the more computationally intensive problems, like problems with distributed loads, tend to have poorer performance. This does not necessarily indicate a lack of ability to solve these problems but that students likely did not want to work out the more complex problems on this test. It is likely that a test like this one measures, at least in part, willingness to take the test. In recent semesters, student work has been collected to try to identify some of these trends better.

Use of free body diagrams on the pre- and post-test

In recent semesters students, were also asked to comment, among other things, on whether they drew a free body diagram (FBD) to answer the questions. They were asked to identify whether they drew an FBD for all questions, some questions, or none of the questions, and the responses identify that many students did not spend the time to draw an FBD, and therefore did not go through that important step of the equilibrium process to answer the questions. Those who report not drawing free body diagrams either guessed, used multiple-choice test logic to answer the questions, or both. Figure 5 gives the percentage of students that drew an FBD to solve all, some, or none of the problems.

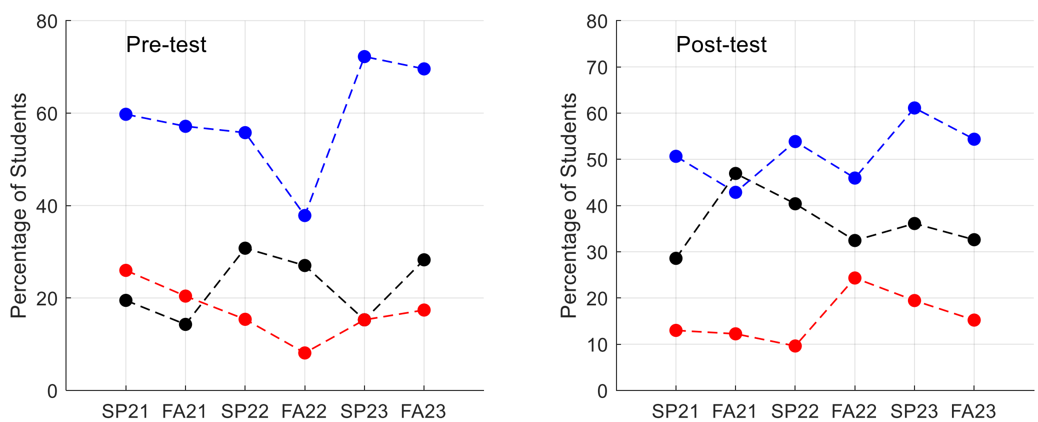

Fig. 5. Percentage of students that report drawing an FBD to solve the problems over six semesters (SP21 through FA23). The black dots represent the response "I drew a solid FBD for *all* problems," the blue dots represent the response "I drew a solid FBD for *some* problems," and the red dots represent the response "I tried to solve the problems without explicitly drawing FBDs."

Figure 4 shows that the majority of students drew FBDs for some or all of the problems. However, a significant fraction attempted to complete the test without drawing free body diagrams. While there is some variation, between 15 and 25 percent of students fall into this category. This number gives insight into the fraction of students who do not take the test seriously. On the other side, between 20 and 30 percent of the students claim to draw free body diagrams for *all* the problems for the pre-test and between 30 and 40 percent claim it for the post-test. This observation suggests that more students are convinced of the need to draw free body diagrams for statics problems after another semester of using them to solve problems in *Deformable Solids*. The number of students who claim to use free body diagrams for *some* of the problems goes down as the number who claim it for all problems goes up. However, the number claiming to draw free body diagrams for some problem is higher than the other two categories for both the pre- and post-test.

One explanation could be that some of the problems might be solved using the process-of-elimination arguments mentioned earlier and, therefore, can be answered with some level of confidence without a free body diagram. It is not clear that the responses ‘all problems’ and ‘some problems’ were mutually exclusive (they could select both), but the results show that the three categories sum to nearly one hundred percent for most of the semesters. There are interesting exceptions, though. In the FA21, FA22, and FA23 semesters, the percentages sum to 92%, 73%, and 115% respectively for the pretest. In the SP21 and SP23 semesters, the percentages sum to 92% and 117%, respectively, for the post-test. The cases where the sum exceeds one hundred percent are likely caused by student selecting both ‘all’ and ‘some.’ The semesters where the sum is less than one hundred percent indicate that students did not respond to the questions about free body diagrams. An improvement to the data collection instrument (Google Forms) would be to require an answer to the question and to make the answers mutually exclusive.

The percentage of students that drew an FBD for every problem generally increased each semester on the post-test, demonstrating that more students found that to be a valuable part of the process to answer the questions. The instructions on the test say:

> “This test is a diagnostic exam to determine your ability to solve statics problems. Please solve all problems (i.e., actually draw free body diagrams and work out the equilibrium equations).”

We also provide verbal encouragement to draw free body diagrams before the test begins. The amount and nature of the encouragement might have varied slightly across the semesters presented. Another factor causing a difference between the pre- and post-tests is that the students are more aware of the incentive (i.e., earning engagement points) when they take the post-test than they are when they take the pre-test. There were two semesters where the percentage of students that did not draw an FBD increased on the post-test while the other semesters had a decrease. This could indicate that those two semesters the students felt more comfortable conceptually answering these questions on the post-test after a semester of practice on similar problems in deformable solids. Overall, these results show that students took the test as a conceptual MCT instead of working through every problem.

Conclusions

The same *Statics* pre-test and post-test has been given to students in *Deformable Solids* for many semesters, providing an opportunity to examine how mastery of statics concepts develops over time. The questions are based on topics from the direct prerequisite *Statics* course. The students answer around 3 questions correctly on the pre-test, improving to almost 4 questions correct on the post-test. The improvement on the post-test is primarily due to improving on the types of problems that they see regularly throughout their semester in deformable solids. These results support the notion that mastery of concepts like those represented by *Statics* develop over time and that expectation of mastery after one course may not be reasonable.

The results of this study suggest that the quantitative results do not provide a completely accurate assessment the students' knowledge or facility with statics but are also likely to reflect the students' willingness to commit effort to the test. Even though students were given adequate time to work these problems out fully and answer the questions, the multiple-choice test format makes it difficult to draw this behavior from the students. Students often use MCT strategies to answer the questions instead. As a result, these tests are not reliable summative measures of statics knowledge, and hence should not be used for placement or requirements for remedial actions.

The main limitation of the testing strategy used is the MCT format. A clearer understanding of student performance on statics concepts could be obtained by changing the format to written response, at the expense of greater difficulty in processing the results. One of our main motivations for doing a pre-test was to get students thinking about what types of problems they should feel comfortable solving after taking *Statics* and which areas might benefit from some review. The pre-test accomplishes that goal well and we have taken strides to give direct feedback on the results promptly to support that goal. The post-test serves as a reminder to students of the importance of prerequisite strings in engineering curricula and, in most cases, to show them that they are making progress toward mastery of important engineering concepts.

At the start, we had expected the results of the pre-test to shed light on what changes might be needed in the *Statics* course itself. The gut reaction to the initial low scores on the pre-test was to lay blame on the prerequisite courses. However, the consistent results from the post-test suggest that our initial performance expectations were off, and what we see in student performance across the mechanics courses, as reflected in the post-test, is that statics skill is fragile after one course and that it is normal for confidence and mastery to develop fully over several courses. Thus, we found that the pre-test results, as presently administered, did not suggest any direct modifications to the prerequisite *Statics* course. These results do, on the other hand, provide a cautionary tale about trying to shorten the learning process for mechanics (e.g., by combining courses are giving shorter, reduced credit, offerings).

References

[1] J. V. Buckle *et al.*, "Validity evidence for STEM problem solving multiple choice questions: Developing an engineering admissions test," *IEEE Glob. Eng. Educ. Conf. EDUCON*, vol. 2018-April, pp. 364–368, 2018.

[2] J. L. Docktor and J. P. Mestre, "Synthesis of discipline-based education research in physics," *Phys. Rev. Spec. Top. - Phys. Educ. Res.*, vol. 10, no. 2, pp. 1–58, 2014.

[3] B. Bridgeman, "A Comparison of Quantitative Questions in Open-Ended and Multiple-Choice Formats Author (s): Brent Bridgeman Published by : National Council on Measurement in Education Stable URL : http://www.jstor.com/stable/1435138 REFERENCES Linked references are av," vol. 29, no. 3, pp. 253–271, 1992.

[4] P. S. Steif and J. A. Dantzler, "A statics concept inventory: Development and psychometric analysis," *J. Eng. Educ.*, vol. 94, no. 4, pp. 363–371, 2005.

[5] G. L. Gray, F. Costanzo, D. Evans, P. Cornwall, B. Self, and J. L. Lane, "The dynamics concept inventory assessment test: A progress report and some results," *ASEE Annu. Conf. Expo. Conf. Proc.*, pp. 4819–4833, 2005.

[6] A. F. Heckler and E. C. Sayre, "What happens between pre- and post-tests: Multiple measurements of student understanding during an introductory physics course," *Am. J. Phys.*, vol. 78, no. 7, pp. 768–777, 2010.

[7] E. C. Sayre and A. F. Heckler, "Peaks and decays of student knowledge in an introductory E&M course," *Phys. Rev. Spec. Top. - Phys. Educ. Res.*, vol. 5, no. 1, pp. 1–5, 2009.

[8] P. Steif and M. Hansen, "Feeding back results from a statics concept inventory to improve instruction," *ASEE Annu. Conf. Expo. Conf. Proc.*, 2006.

[9] R. Echempati, "Statics concepts inventory results at Kettering University," *ASEE Annu. Conf. Expo. Conf. Proc.*, 2014.

[10] P. S. Steif, "An articulation of the concepts and skills which underlie engineering Statics," *Proc. - Front. Educ. Conf. FIE*, vol. 2, pp. F1F-5, 2004.