2021 ASEE ANNUAL CONFERENCE



Development of a New Concept Inventory for Mechanics of Materials

Dr. Stephen N. Kuchnicki, York College of Pennsylvania

Dr. Stephen Kuchnicki has been an Associate Professor of Mechanical Engineering at York College of Pennsylvania since August 2013 and an Assistant Professor from January 2008 until that date. From 2015-2019, he also served as Coordinator of Mechanical Engineering, a position in which he managed the operation of the program at York College of Pennsylvania. Previously, he was a postdoctoral research associate at Rutgers University, specializing in computational modeling of dynamic deformations in solids. His areas of technical expertise include solid mechanics, crystal plasticity, vibration, and fluid-structure interaction. He received his Ph.D. from Rutgers University in 2001.

Paper ID #33110

Dr. Tristan M. Ericson, York College of Pennsylvania

Dr. Tristan Ericson is an Associate Professor at York College of Pennsylvania and has been teaching there since 2013. Prior to this appointment, he was a Visiting Professor at Bucknell University. His teaching interests include solid mechanics, vibrations, materials science, and MATLAB. He advises the YC Racing FormulaSAE team. His technical research interests include vibrations of planetary gear systems, strengthening 3D printed materials, and making things go faster. He enjoys activities that promote STEM fields in local high schools. He received his PhD from Ohio State University in 2012.

Development of a New Concept Inventory for Mechanics of Materials

Abstract: Concept inventories are useful diagnostic tools for courses in which students' knowledge of the fundamentals covered by a course can be tested without calculations. Thus, the students' understanding of the main concepts can be tested without worry regarding computational errors. These inventories are often used as assessment tools (both for ABET purposes and internally for course improvement) since they can illustrate which topics students struggle with and whether improvement in a topic was evident throughout the course. Concept inventories exist for many subjects, including Physics, Statics, Thermodynamics, and Dynamics. At least two groups had begun work on concept inventories for Mechanics of Materials which now appears to have been abandoned. Thus, the authors developed their own concept inventory for use with a Mechanics of Materials course.

This concept inventory was administered to students in the Strength of Materials course at York College of Pennsylvania (YCP) twice in the Spring 2020 semester – once as a pretest in the first week of the course, and then again as a post-test in the last week of the course. The results of these tests were used for analyses of student understanding as well as the performance of the concept test itself. We examined the performance of students on the post-test relative to the pretest and correlated results of both tests to course grades. We also determined diagnostic information about the test questions, such as difficulty index and discrimination index, to find whether the questions on the test were doing the job we wished them to do.

Overall, we found that the student performance improved between the pretest and the post-test. Also, we found a weak positive correlation between student course percentage grade and the post-test performance. Most test questions had difficulty values in an acceptable range; we saw more questions which may have been too hard than those which may have been too easy. Over 80% of the questions were found to have discrimination index values over 0.2, meaning that the questions were able to differentiate between students performing well and students who performed poorly. This instrument seems to provide a foundation for future improvements.

Introduction

Concept inventories are common tools used in several disciplines. These instruments are useful to determine student understanding of basic concepts within a course and to supplement more common course-level metrics such as numerical problem-solving. The utility of these instruments is noted across various disciplines. For example, concept inventories have been established for statistics [1] and calculus [2] concepts. Additionally, concept inventories are broadly available for several physics topics, including energy and momentum [3], waves [4], and basic physics of thermal sciences [5].

These inventories have also been used in engineering courses. Basic engineering mechanics, including statics, forces and dynamics, have well-developed concept inventories [6,7,8,9]. Similar

instruments have been used in courses with advanced topics. There exists work on development of concept inventories for biomechanics [10], materials engineering [11], electronics and signals [12,13], and fluid mechanics [14]. Development of a concept inventory for strength of materials has also been undertaken [15].

It would seem, then, that the extensive work on this topic leaves little room for further development. In reality, however, the most recent work on many of these concept inventories is over a decade old. The concept inventory for strength of materials is a good example – the last paper published on this inventory was produced in 2003.

Instructors in the strength of materials course at YCP wished to determine the conceptual understanding of students in the course. The state of the existing strength of materials concept inventory led these instructors to develop a new instrument to assess this understanding.

Initial Development

The first question we asked in developing the new concept inventory was perhaps the most natural – What are the important concepts in a strength of materials course? We arrived the following list of concepts for the course:

- Stress due to axial loading
- Stress due to bending loading
- Stress due to torsion loading
- Principal stress states and maximum shear stress
- Deflection due to bending
- Boundary conditions for deflection
- Stress due to combined loading
- Location of maximum stress under an arbitrary loading

This list of concepts helped us create questions by which we could address each of them. The first draft of our concept inventory consists of 21 questions, all of which are multiple choice. Two of these questions had multiple answers. One of these, for example, asked students to select all of the principal stress states out of four possibilities. We will discuss why these questions in particular were less useful than originally hoped in the discussion of our results.

Using 21 questions for eight general concepts means that several concepts are addressed in multiple questions. One challenge in creating questions for a concept inventory is not so much in addressing the concepts but rather in isolating concepts, as noted in [15]. It is easy to create questions that incorporate two or three concepts, but these questions lead to unclear results of the concept inventory; if a question addresses multiple concepts, then it does not illustrate student understanding of any concept. Thus, most of the questions are written in a way that might be considered "simple" or too straightforward. We feel this approach is necessary to isolate concepts and enable the instrument to evaluate concepts independently. Otherwise, the results of this inventory could not identify potential weaknesses in the instruction of a course, for example. Example questions are provided in the Appendix.

Implementation

This concept inventory was implemented in the learning management system, Canvas, used at YCP. A Canvas quiz was constructed using the questions of the concept inventory. This quiz was not timed in order to allow students to feel they had the time they needed to think about the questions. The concept inventory did not count as a grade in the courses where it was offered. Rather, students earned a grade of 100% on a homework assignment every time the concept inventory was offered, regardless of score on the concept inventory. This choice was made to both encourage participation and to discourage collaboration. While we tend to encourage students discussing homework problems with their peers, this is a case in which we wanted a purely individual metric. The quiz was available for several days in Canvas to allow students to take it at their leisure rather than during class time. This also allowed students to take all the time they felt they needed on the concept inventory rather than feeling somewhat rushed by a timed test. Because the initial implementation of the quiz occurred in the Spring 2020 semester, the choice of an online implementation proved fortunate – when YCP transitioned to fully-remote learning in March 2020 (like many other institutions), this concept inventory could still be administered in the same way.

Students were offered this instrument three times during calendar year 2020. The first offering was in January 2020, prior to the beginning of the Strength of Materials course, to assess the initial understanding of students from intuition and gauge what they learned in the prerequisite Statics course. This Statics course at YCP includes basic axial and bending stress computations as well as deflection mainly via tables, so we thought that students would have had a chance to develop some conceptual understanding from that course also.

The second offering came in May 2020, just before the end of the Strength of Materials course. We thought this timing would serve two purposes – one for us, and one for the students. For us, it would show whether there was an increase in conceptual understanding from the start of the course to the end. We also presented it to students as a final exam study tool – as the concepts in this instrument were covered on the comprehensive final exam, why not use this as a free chance to review those concepts? The student population for both of these two administrations of the inventory was the sophomore Civil Engineering and Mechanical Engineering students.

The third administration of the concept inventory occurred in August 2020 as a pretest before our Machine Design course. This course uses the Strength of Materials course as a prerequisite and it expands upon topics covered in that course, so this was a sensible place to use the instrument. This test then served as a gauge of student understanding of the fundamentals (which were reviewed at the beginning of the course) and a student self-assessment of what topics might need some extra review. This implementation of the concept inventory has a different population – this third assessment only captured the Mechanical Engineering students, as the Civil Engineering students do not take this course.

We note again that the implementation of this concept inventory did not change among the three offerings. It was a Canvas quiz all three times and it consisted of the same multiple-choice

questions appearing randomly. The only material changes among the three offerings are the change in population noted above and small changes in the order of answers to a question.

Evaluation

Our evaluation of this concept inventory looks at several aspects that we saw as indicators of its usefulness. We first describe how well inventory performance correlates to student success in the Strength of Materials course overall. Then, we discuss how this concept inventory relates to student learning of topics over the duration of the course. Next, we talk about the effectiveness of individual questions via their difficulty index and discrimination index values. We relate the performance of the mechanical engineering students to retention of topics across an intervening summer term, and then we finish by correlating conceptual understanding as evidenced by concept inventory performance to success within the Strength of Materials course. As we analyze these indicators of student success, we consider student concept inventory performance, course letter grades, and course percentage grades.

Correlation to Student Success

The concept inventory covers topics that our students would have seen in both the prerequisite Statics course and the Strength of Materials course. This makes a comparison between the pretest concept inventory score and student Statics grades relevant (see Figure 1). Note that these grades are listed numerically; 4.0 is an "A" grade, 3.5 is a "B" grade, and so forth to 2.0 as a "C" grade. Because the Strength of Materials course prerequisite requires that students pass Statics with a "C" grade, there are no grades lower than 2.0 on the horizontal axis of Figure 1.



Figure 1 - Pre-test scores versus Statics letter grade

This figure shows a weak positive correlation between the Statics grade and the pre-test performance. This is a reasonable expectation because the concept inventory was not composed of entirely new material, but rather included some material students would see as review. The mean score is 9.9 out of 21 with a standard deviation of 2.24.

We also compare these pre-test scores to Strength of Materials course percentage grade in Figure 2. Note that the correlation here is essentially zero. We expected this result because the pre-test was administered before the students learned the Strength of Materials course content. Considering this initial null relationship, any positive correlation observed later – at the time of the post-test – would indicate learning of concepts within the course and demonstrate the link between conceptual understanding and course grades.



Figure 2 - Pre-test score versus Strength of Materials course percentage

Figure 3 plots the post-test scores versus student grade percentage in Strength of Materials. Note that we see a larger positive correlation than we did in Figures 1 or 2 for the pre-test. The slope is a small numerical value because the total number of test questions is one fifth a perfect course percentage. The value 0.1107 implies that each additional correct post-test question correlates to a 9% increase in course percentage. Further, we note that the Machine Design course at YCP uses a grade of at least "C" (70%) in Strength of Materials as a prerequisite. No student who earned below 70% in Strength of Materials earned a higher score than 13/21 (62%) on the concept inventory. The concept inventory is then similar to the overall course percentage in terms of determining readiness for any follow-on courses.



Figure 3 - Post-test score versus Strength of Materials course percentage

Analysis of Student Learning Within the Course

This concept inventory also helps determine whether students learned the key concepts throughout the course. Figures 2 and 3 above demonstrate this. Student grades in the Strength of Materials course correlate more strongly with post-test scores than pre-test scores. This implies that students have learned important concepts in the course. Further, these results show that students who demonstrated better knowledge on homework, exams, and the other graded events in the course have a stronger understanding of the course concepts.

We also consider student performance improvement between the pre-test and the post-test. If students learn the course concepts as a result of taking the class, then we expect their performance to improve on the concept inventory from the pre-test to the post-test. Figure 4 tracks this improvement. Positive values correspond to students who did better on the post-test than on the pre-test. The thick black line approximates the mean improvement of 3.5 points; the median improvement was 3.54 points. This sample is limited to those who took both the pre-test and the post-test; some students only took one or the other. This gives us 65 data points with 3 students who stayed the same between the tests and eight whose scores went down on the post-test. The remaining 54 students improved their performance on the post-test compared to the pre-test.



Figure 4 - Change in score from pre-test to post-test

A metric that other authors (see especially [13] and [16]) have used to quantify increased student knowledge is *normalized gain*. This quantity is the fraction of the possible improvement students attained between administrations of a test. Using g in the below as the normalized gain, we can define this as:

$$g = \frac{\{T_2 - T_1\}}{\{21 - T_1\}}$$

where T_2 is the student post-test score and T_1 is the student pre-test score. This equation is written to use our 21-question concept inventory raw score; other authors have used percentages instead. The normalized gain for all students is shown in Figure 5 – the mean value is 0.31. The red line corresponds to equal performance on the pre- and post-tests (no normalized gain). The dashed line represents an average normalized gain, which is 6.3 more questions correct. This dashed line necessarily approaches zero gain for a perfect pre-test score as no gain would be possible in that instance.

Figure 5 - Normalized gain for all students who took pre- and post-tests (N=65)

The concept inventory in [13] reported a normalized gain of 0.24 with a standard deviation of 0.08. Our mean here is 0.31 with a standard deviation of 0.26. Our data shows a mean gain comparable with other concept inventories but with a larger standard deviation. This standard deviation may be due to a smaller sample drawn from a single class or may indicate a need to examine the questions themselves more carefully.

These two metrics – scores versus course grade and improvement between pre- and post-test – lead us to conclude that our concept inventory results demonstrate increased student conceptual understanding as a result of taking the course.

Analysis of Concept Inventory Questions

Student performance on the concept inventory is one part of the overall question. We must also consider metrics that help determine the quality of questions on the concept inventory. We use the difficulty index and discrimination index, both computed by Canvas, to determine the quality of the questions on our concept inventory.

Figure 6 shows the difficulty index values for the questions at the pre-test delivery. The difficulty index from Canvas ranges from 0 to 1, with lower values corresponding to more difficult questions. This metric is the fraction of students who answered each question correctly. Difficulty index values under 0.2 correspond to especially difficult questions; those over 0.8 denote easier questions. These areas are shaded appropriately in the figure.

Figure 6 - Pre-test difficulty index values from Canvas

Note that only 19 questions appear in Figure 5 rather than the 21 on the concept inventory. This is because we have two questions on our concept inventory that had multiple answers (i.e., "Choose all of the principal stress states below"). Canvas does not provide statistics for questions that have multiple answers. Of the 19 questions represented, five fall into the "Hard" range (with a sixth on the edge). Five more are in the "Easy" range.

We consider the content of the five questions deemed easy by this pre-test. Question 1 tested stress due to axial loading. Question 6 covered the deformed shape of a beam under a point load. Question 7 asked for the best cross-section to resist a bending moment. Question 13 covered construction of a shear diagram from two point loads, and question 15 asked students to determine the correct moment diagram given a shear diagram. All of these concepts appear in the Statics class at YCP – thus, this result shows retention of concepts from that course.

The discrimination index rates how well each individual question identifies high scorers on the test versus low scorers. The higher the discrimination index value, the more likely that someone who got that question correct scored high on the exam overall. This value ranges from -1 to 1; a value of greater than 0.24 corresponds to a question that is considered a good discriminator [17]. Discrimination index values for the pre-test are shown in Figure 7. Nine of these questions did not discriminate well on the pre-test. While the initial indicators here are not good for these questions, the post-test shows improvement – see Figure 8.

Figure 7 - Discrimination index values for the pre-test

Figure 8 - Discrimination index values for the post-test

We see that all post-test questions aside from 3, 6, and 7 exceed the benchmark discrimination index of 0.20. This result shows that most of these questions do an acceptable job of differentiating between students who do well on the test overall and those who do not. Also note that Questions 6 and 7 were noted above as being particularly easy. Thus, they seem not to discriminate well because most students got them correct.

Student Retention of Information

We chose to administer this concept inventory a third time, in August 2020, to the mechanical engineering students only as part of the Machine Design course. This course uses Strength of Materials knowledge as a prerequisite. The concept inventory could then gauge student retention

of this course material across an intervening summer. It also allowed students to see which concepts they may have needed to review as they began the Machine Design course.

Figure 9 plots the performance on the Machine Design pretest versus performance on the Strength of Materials post-test. Because the student populations are similar between the two administrations of the test and because there was no formal instruction between these tests we anticipated a strong correlation between these datasets. Figure 8 shows that we did not get this correlation.

Figure 9 - Machine Design pre-test scores versus Strength of Materials post-test scores

The results in Figure 9 are surprising. Our expectation was that student performance would be consistent between these two administrations of the test. However, we see almost no correlation between these two test events. The implication here is that the intervening summer had a negative effect on student retention of concepts. Normally, the sophomore class at YCP would be participating in a co-op experience for this intervening summer. Thus, students could still be exposed to some of the concepts from the Strength of Materials course as part of their work. Due to the ongoing COVID pandemic, however, students did not uniformly participate in co-op over this intervening summer. An interesting comparison point would be to repeat this exercise for the current sophomore class to determine whether the co-op over the summer assists with retention of concepts.

Importance of Conceptual Understanding to Course Success

The concept inventory allows us to evaluate how well we teach and emphasize these concepts within our course. If we expect students will need to understand the concepts well to excel in the

Strength of Materials course, then it follows that improvement on the concept inventory should correlate to success in the course. A part of this was shown earlier in Figure 3, where our trend showed a positive slope of 0.1107 when we plotted concept inventory post-test score versus course percentage. This slope then implies that one extra correct question on the concept inventory corresponds to roughly a 9% increase in the course percentage.

We can also look at the higher level of course letter grades. We plot the post-test performance versus Strength of Materials letter grade in Figure 10. This figure uses the same scale as Figure 1 -4.0 is an "A" grade, 3.5 is a "B+" grade, down to 0 as an "F" grade.

Figure 10 - Post-test score versus Strength of Materials course grade

The positive slope here implies that an increase of one letter grade in the course corresponds to approximately 1 $\frac{1}{2}$ more questions correct on the post-test – about 7% better performance. Alternately, we can view this as on more correct question being worth about 0.7% of a letter grade – enough to move from a C+ to a B or from a B+ to an A. These results indicate that the concept inventory measures understanding of the course material to an extent.

We say "to some extent" here because we have outlier scores at higher course grades that are lower than students who earned lower course grades. There are several reasons why this might be so. First, while the overall course grade may be increased by understanding of these concepts, the course grade is not based exclusively on them. A class exam problem may combine multiple concepts. Thus, a student may be able to earn a passing score on a class exam problem without understanding all of the concepts within the problem to the degree asked on the concept inventory.

A combined loading problem is a good example of this case. On the exam, a student could evaluate the axial, bending, and torsional stresses on a body properly but then look for the maximum combined stress at the incorrect point within the cross-section. This student would likely earn a good score for the class exam – demonstrating understanding of three concepts! – but would be completely incorrect on a concept inventory problem.

Another potential issue is the incentive offered for the concept inventory. For this first run of the inventory, we offered students a grade of 100% on a homework assignment for completing the concept inventory, regardless of student score on the concept inventory. This may have led some students to regard the concept test less seriously than an exam, for which the grade earned would be a major factor in the course grade, rather than mere completion.

Conclusions

We have presented the results of a concept inventory for Strength of Materials. This concept inventory shows a weak correlation between student course performance and scores on the concept inventory when it is administered as a post-test, but no correlation when it is administered as a pretest. A 5% improvement in correct answers on the concept inventory (one more correct answer) correlates to a 9% improvement in the course grade, which highlights the importance of conceptual understanding in academic success. We also examined the normalized gain from the pre-test to post-test and found an average gain of 0.31, or six more questions correct.

The individual questions on the concept inventory were also rated for difficulty and discrimination. We found that most questions had a difficulty index above 0.20 (hard), with 9 of 19 questions between 0.20 and 0.80. These results imply that we should give this concept inventory another iteration to determine whether some questions need to be made difficult.

The discrimination index of most questions (16 of 19) was found to be at least low-acceptable (greater than 0.25). Three questions did not discriminate well between high and low scorers. While we feel these questions are important, we will continue to monitor the results and to consider possible ways to address the same concepts in a different manner.

Our plan for the next iteration is to revise the two questions we had on the concept inventory that had multiple answers to have a single answer instead. This change will allow us to determine difficulty and discrimination indices for these questions and to draw informed conclusions about their efficacy.

References

- 1. K. Allen, "The statistics concept inventory: development and analysis of a cognitive assessment instrument in statistics" Ph.D. Dissertation, Univ. of Oklahoma, Norman, Oklahoma, 2006.
- J. Epstein, Development and validation of the Calculus Concept Inventory", In 9th Int. Conf. on Math. Educ. in a Global Community, Charlotte, NC, September 2007, pp. 165-170.
- 3. C. Singh and D. Rosengrant, "Multiple-choice test of energy and momentum concepts," *Amer. J. of Phys.* vol. 71, no. 6, pp. 607-617, 2003.
- 4. T. R. Thoads and R. J. Roedel, "The wave concept inventory-a cognitive instrument based on Bloom's taxonomy", In 29th Annual Frontiers in Educ. Conf., Nov. 1999, pp.13C1-14.
- 5. S. Yeo and M. Zadnik, "Introductory thermal concept evaluation: Assessing students' understanding", *The Phys. Teacher*, vol. 39, no. 8, 496-504, 2001.

- G. L. Gray, F. Costanzo, D. Evans, P. Cornwell, B. Self, and J. L. Lane, "The dynamics concept inventory assessment test: A progress report and some results," In *Amer. Soc. for Eng. Educ. Annu. Conf. and Expo.*, Jun. 2005. Available: https://sites.esm.psu.edu/dci/.
- 7. D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," *The Phys. Teacher*, vol. 30, no. 3, pp. 141-158, 1992.
- 8. P. Nieminen, A. Savinainen, and J. Viiri, (2010) "Force concept inventory-based multiplechoice test for investigating students' representational consistency," *Phys. Rev. Special Topics-Phys. Educ. Res.*, vol. 6, no. 2, 2010.
- 9. P.S. Steif and J. A. Dantzler, "A statics concept inventory: Development and psychometric analysis," *J. of Eng. Educ.*, vol. 94, no. 4, pp. 363-371, 2005
- D. Knudson, G. Noffal, J. Bauer, P. McGinnis, M. Bird, J. Chow, and J. Abendroth-Smith, "Development and evaluation of a biomechanics concept inventory," *Sports Biomechanics*, vol. 2, no. 2, pp. 267-277, 2003.
- 11. S. Krause, J. C. Decker, and R. Griffin, "Using a materials concept inventory to assess conceptual gain in introductory materials engineering courses," In *33rd Annu. Frontiers in Educ. Conf*, Nov. 2003, vol. 1, pp. T3D-7.
- 12. M. F. Simoni, M. E. Herniter, and B.A. Ferguson, (2004, June). Concepts to questions: Creating an electronics concept inventory exam. In *Proc. of the 2004 Amer. Soc. for Eng. Educ. Annu. Conf. and Expo.*, Jun. 2004.
- 13. K. E. Wage, J. R. Buck, T. B. Welch, and C. H. Wright, "The signals and systems concept inventory", In *Proc. of the 2002 Amer. Soc. for Eng. Educ. Annu. Conf. and Expo.*, Jun. 2002.
- 14. J. Martin, J. Mitchell, and T. Newell, "Development of a concept inventory for fluid mechanics," In *33rd Annu. Frontiers in Educ. Conf.*, Nov. 2003, vol. 1, pp. T3D-23.
- 15. J. Richardson, P. Steif, J. Morgan, and J. Dantzler, "Development of a concept inventory for strength of materials," In *33rd Annu. Frontiers in Educ. Conf*, Nov. 2003, vol. 1, pp. T3D-29.
- 16. R.R. Hake, "Interactive-engagement vs. traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses," *Am. J. Phys.*, vol. 66, pp. 64-74, 1998.
- 17. Instructure, Inc. "Canvas Quiz Item Analysis." <u>https://s3.amazonaws.com/tr-learncanvas/docs/CanvasQuizItemAnalysis.pdf Accessed 3 March 2021</u>.

Appendix – Example Concept Inventory Questions

The following are several of the 21 questions on the concept inventory test. All questions were multiple choice.

Question 1:

Which of the three bars below has the smallest stress?

A. A B. B C. C D. Both A and B E. All bars have the same stress

Question 4:

Given the bar loaded below, which point has the largest shear stress?

A. A B. B C. C D. A and B are equal E. All points have the same shear stress

Question 9:

Given the bar shown with an angled cut and forces applied as shown, which value of angle θ gives the largest normal stress?

A. 0 degrees B. 30 degrees C. 45 degrees D. 60 degrees E. 90 degrees

Question 15:

Choose the moment diagram which corresponds to the shear diagram given below.

D

Question 17:

The beam below is supported by a fixed support at left and a roller at right. What are the appropriate compatibility conditions? (Point C is at the midpoint from A to B.)

