AC 2008-2537: ASSESSING COGNITIVE REASONING AND LEARNING IN MECHANICS

Chris Papadopoulos, University of Wisconsin - Milwaukee

Chris Papadopoulos earned BS degrees in Civil Engineering and Mathematics from Carnegie Mellon University, and a PhD in Theoretical and Applied Mechanics, Cornell University. He previously served on the faculty of Engineering at the University of Wisconsin-Milwaukee, where he is currently a research associate, grant writer, lecturer, and director of educational programs. His research interests include biomechanics, nonlinear structural mechanics, computational mechanics, engineering education, and engineering ethics. He is an active member of American Society for Engineering Education (ASEE), American Society of Civil Engineers (ASCE), and Association for Practical and Professional Ethics (APPE).

Assessing Cognitive Reasoning and Learning in Mechanics

1. Introduction

Mechanics is the discipline that anchors the engineer's scientific reasoning. While mechanics is properly concerned with practical treatment of the behavior of structures and mechanisms – "what to think", it is equally concerned with the methodology of problem formulation and solution – "how to think". Not coincidentally, all accredited engineering programs require mechanics courses at entry to major. Because mechanics is so centrally situated in the engineer's intellectual training, it lends itself to the study of engineers' thinking, learning, and metacognition.

Perhaps because of these characteristics, a great deal of research has been conducted to assess student learning in mechanics and methods of teaching mechanics. Educators in physics and engineering have developed a clear understanding of misconceptions that conflict with student learning, and the concept inventory has emerged as a powerful tool to identify these misconceptions. I review several results of the literature on misconceptions and use of concept inventories. In the course of this review, I raise the issue of whether concept understanding is sufficient for problem-solving, and I suggest that procedural knowledge must also be emphasized in mechanics instruction.

I also review other methods of assessing student learning, such as student interviews and interactive reviews of student work. I also provide some information on studies that have attempted to evaluate textbooks. I frame the discussion on textbooks within the larger context of promoting problem-solving by the incorporation of systematic procedure and the fostering of sound problem-solving habits.

2. Misconceptions and Concept Inventories in Physics Education

Formal study of student knowledge and learning in mechanics is rooted in research conducted by physics educators since the early 1970s. In 1984, Lillian McDermott[18] published what is perhaps the first review of these efforts. The over-riding themes that had emerged were that (1) students bring misconceptions to the classroom that contradict principles of Newtonian physics; (2) misconceptions are often resistant to change, and often persist after instruction; (3) students can obtain correct answers to problems (e.g., by plugging numbers into formulae) without understanding the underlying concepts; and (4) students have difficulty applying basic concepts to actual physical situations. Corresponding to these general themes were several specific types of misconceptions that were widely held by students:

- *Neglecting "passive" forces*. Students often neglect forces that do not appear to be "active", such as the normal force supplied by a tabletop to support a book that rests on the tabletop[18]. Such forces are termed "passive" in the physics education literature and roughly correspond to "reactions" in engineering.

- *"Motion Implies Force"*. Students often believe that "motion implies force", in which forces are incorrectly neglected to bodies that are stationary, and incorrectly thought to parallel velocities[3,30]. An example would be believing that the net force acting on a ball thrown upwards gradually diminishes as the ball ascends, and/or a total vanishing of all forces (including gravity) when the ball reaches is maximum height. Several researchers have observed that such misconceptions often parallel incorrect "impetus" theories proposed by early scientists[3,25].

- *"Dominance Principle"*. Given a large object contacting a small object, students often assume the that the force that the large object exerts on the small object is greater than the force that the small object exerts on large object. This is referred to as the "dominance principle"[14], and contradicts Newton's Third Law of Action-Reaction.

- *Velocity and Acceleration*. Students often incorrectly judge the flight time of an object based on how fast it is launched or how heavy it is. For example, a student might believe that a bullet fired horizontally will fly for longer time than a bullet that is dropped from the same height[25].

Research has documented that these and other misconceptions persist even after instruction in Newtonian mechanics. For example, in two experiments with several dozen students, Clement[3] found that the "motion implies force" misconception is evident in approximately 90% of students prior to taking introductory physics, and over 70% afterwards. As a result of these and other studies, physics educators began to develop new pedagogies to foster understanding of concepts. These new approaches included direct address of misconceptions and increased interaction and feedback.

During the 1980s, David Hestenes formalized the measurement of misconceptions in mechanics by developing the Mechanics Diagnostic Test[12], which was later refined into the Force Concept Inventory (FCI)[14] to test students' knowledge of forces in Newtonian mechanics. A concept inventory, and in particular the FCI, is a multiple-choice test in which a single, correct answer is placed amongst several false "distracters" that reflect commonly held misconceptions. By coding each possible answer, student responses on the concept inventory can be used to diagnose misconceptions. The FCI was designed to test 30 concepts (grouped into 6 categories) through 29 multiple choice questions.

Hestenes et al. used the FCI not only to measure student knowledge, but also to evaluate the effectiveness of teaching methods[14]. The FCI was administered to 1500 high school students and 500 university students, both prior to receiving mechanics instruction (pre-test) and afterwards (post-test). Some students in each group were exposed to "interactive" instruction (computer exercises were integrated into the class), whereas the others received "standard" instruction. Students who received interactive instruction performed better than students receiving standard instruction, both in terms of absolute post-test score and relative gain from pre-test to post-test. As defined by Hake[11], relative gain <g> is the ratio of the increase in test score to the maximum possible increase:

$$< g > = \frac{POST - PRE}{100 - PRE}$$

Although Hestenes et al.[14] reported only raw pre- and post- test scores, according to my manual calculations from this data, students with interactive instruction achieved average relative gains of 49% (range 44%-62%), compared with 28% (range 21%-34%) for students in standard instruction.

Later, Hake[11] administered the FCI to approximately 6500 high school, college, and university students. His principal conclusion was that students at any of these levels who were exposed to instruction featuring "interactive engagement" (essentially any bona fide attempt to supplement traditional lecturing) achieved significantly higher relative gains from pre-test to post-test. Hake reported that students receiving interactive engagement achieved average relative gains of 48% (s.d. = 14%), compared to 23% (s.d. = 4%) for students in standard instruction. He did not directly report post-test scores; I extracted these through manual calculation.

Table 1 summarizes the data collected by Hestenes and Hake, and also includes additional analysis to facilitate direct comparisons between data sets. As can be seen, the two data sets are in close agreement with one another. Because the meaning of "interactive" is broad – many types of teaching methods can qualify as "interactive", provided that they attempt student engagement beyond traditional classroom instruction – these two studies provide compelling evidence that strategic interaction with students fosters good learning. Yet with relative gains averaging below 50% even for the interactively engaged students, these studies also call for further innovations in teaching.

Table 1. Summary of FCI Average Student Performance Data

FCI Administrator	Instructional Mode	N	Post Test (%)	Relative Gain (%)
Hestenes et al. [14]	Interactive	282	67	*49*
Hestenes et al. [14]	Standard	958	51	*28*
Hake [11]	Interactive	4458	*70*	48
Hake [11]	Standard	2084	*56*	23

Data in *italics* are manually calculated from reported raw data.

Hestenes and Hake provide some analysis to suggest that the reliability and validity of the FCI, but few details are available. Hestenes et al. claim that students who truly understand Newtonian reasoning are unlikely to score poorly, although it FCI scores tend somewhat to overestimate student understanding of concepts[14].

3. Misconceptions and Concept Inventories in Engineering Mechanics Education

Following the research in the physics community, engineering educators are aware of misconceptions held by students and are responding with concept inventories in mechanics and several other subject areas. A current initiative is coordinated through the Foundation Coalition[26] to "reinvigorate" the use of concept inventories in the STEM disciplines. An excellent overview of the design, implementation, and uses of concept inventories in several areas of engineering is provided by Richardson[28]. I will summarize engineering educators' efforts in developing and implementing concept inventories in mechanics.

In 2004, Steif[31] proposed concepts that are important for Statics and developed the Statics Concept Inventory (SCI). The SCI consists of 27 multiple choice questions to test 9 concepts. The SCI differs from the FCI by focusing on concepts germane to engineering problem-solving (e.g., understanding internal forces at connections between structural elements). Steif & Dantzler[32] and Steif & Hansen[33] performed rigorous statistical analyses to demonstrate the validity and reliability of the SCI (e.g., that the SCI can reliably diagnose misconceptions, with students tending to choose the same misconceptions repeatedly on questions testing the same concept). Steif & Hansen[33] also advance the practicality of administering the SCI via the web, to facilitate wide dissemination, collection, and analysis of data.

Using data from a sample of approximately 100 students at a single institution, Steif & Dantzler[32] report pre-test (prior to Statics instruction) scores averaging 39% (10.4/27) and post-test scores (after Statics instruction) averaging 75% (20.3/27), equating to an average relative gain of 59%. However, they point out that even after instruction, incorrect answers still dominate on some questions, and certain wrong answers occur very frequently. However, in a study of over 1100 students from 14 different institutions, Steif & Hansen[33] reported that pre-test scores corresponded to random guessing (~20%) and back-calculation from their data indicated that post-test scores averaged 47%, corresponding to an average relative gain of 32%. Neither of the SCI studies attempted to distinguish student performance based on mode of instruction (e.g., interactive instruction).

A multi-member team consisting of Gray, Costanza, Cornwell, and Self has developed the Dynamics Concept Inventory (DCI)[9]. The DCI was developed through a Delphi process in which 25 expert dynamics instructors proposed concepts that they viewed to be both important and frequently misunderstood in dynamics. A preliminary DCI was developed and tested with student focus groups, some of which took the test in expository rather than multiple-choice form to draw out problem statements, misconceptions, and distractors in the final multiple-choice version. The DCI contains 30 questions to test 11 concepts and has been statistically analyzed for validity and reliability, and has been shown to be a good diagnostic tool for misconceptions in dynamics.

DCI data has been collected from 1200 students in 5 groups: 2 pre-test groups and 3 post-test groups[10]. Students in 2 of the post-test groups were directly exposed to concept questions about every 1-2 weeks during class. The average scores of the pre-test groups and the post-test group that did not receive direct concept-questioning ranged from 31% - 35%. Scores of the post-test groups that did receive direct concept questioning were markedly higher, ranging from 56% - 64%.

Table 2 summarizes data from the SCI and DCI. The data cannot easily be compared because they come from different levels of students under different conditions. As more data is collected, perhaps baseline averages can be developed to enable direct interpretation of post-test scores. The adoption of uniform measures of reporting CI data would further facilitate the study and interpretation of CI data.

Table 2. Summary of SCI and DCI Average Student Performance Data				
FCI Administrator	Instructional Mode	N	Post Test (%)	Relative Gain (%)
Steif & Dantzler [32]	N/A	105	*75*	*59*
Steif & Hansen [33]	N/A	1164	*47*	*32*
Gray et al. [10]	Direct concept-teaching	476	59	*39*
Gray et al. [10]	No direct concept teaching	147	N/A	Negligible
Data in *italics* are manually calculated from reported raw data.				

Finally, in conjunction with the Foundation Coalition effort, Richardson, Morgan, and Steif are developing a Strength of Materials concept inventory (which I will abbreviate SOM-CI)[27]. Less data from the SOM-CI appears to be available. Brown et al. has used this SOM-CI to give a preliminary evaluation of the effectiveness of using concept tutorials given to an experimental group of 50 mechanics of materials students during the study week prior to the final exam[2]. A control group of 36 students in the same class did not receive concept tutorials. Each group was then tested using 4 questions from the SOM-CI dealing with beam bending. The data showed a uniformly inverse relationship between concept teaching and SOM-CI score: the control group outscored the experimental group on all four questions. The experimental group did outperform the control group on the final exam, a fact that was not noted by the authors.

Sweeney et al.[35] have developed what appears to be an independent Mechanics of Materials concept inventory (MOM-CI). This MOM-CI focuses on concepts rooted in Hooke's Law and states of stress and strain that are relevant to professional engineering design. No data appears to be available regarding results using this MOM-CI.

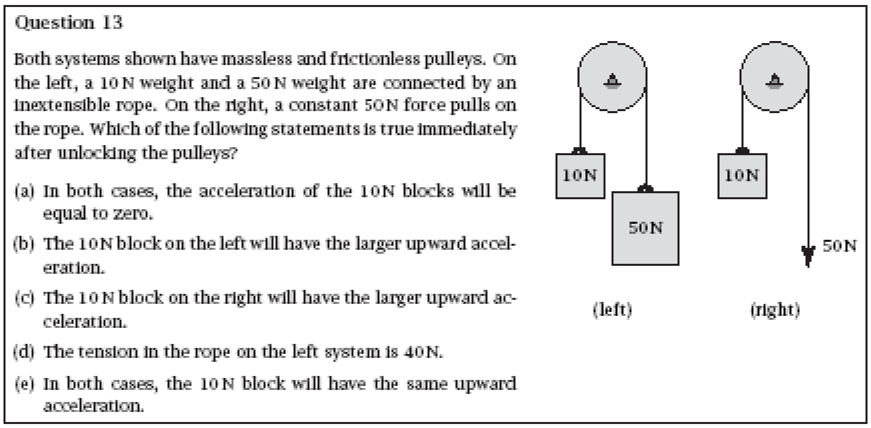
4. Conceptual Understanding vs. Procedural Knowledge

Although concept inventories – especially those that have been rigorously analyzed for reliability and validity – are effective tools for diagnosing student misunderstandings. Indeed, a student who answers, say, all three questions related to a given concept is likely to really understand the concept as is presumably drawing upon more than rote knowledge. Nevertheless, partly because of their multiple-choice nature, concept inventories are somewhat limited in their ability to determine whether students can apply or transfer their knowledge beyond the idealized confines of the concept questions.

Shedding some light on this is the account by Evans et al.[6] regarding the development of one of the questions on the DCI related to angular momentum balance. Initially the developers proposed a question that essentially asked why a car seems to "nose-dive" if the brakes are applied suddenly. But testing in student focus groups revealed that most students could not answer this question directly. The developers replaced this question which a much more idealized problem that asked students to predict the acceleration of the center of mass of a planar box that is loaded with a non-centroidal force (even though the box will have an angular acceleration, the center of mass will accelerate parallel to the applied force). This version of the question did draw correct responses from about a quarter of the students.

I contend that there are two basic modes by which to approach a problem, including "concept questions". The first is to have direct intuition; the second is to have knowledge and confidence in a set of general principles and procedure, even in the absence of direct intuition. In fact, the power of procedural methods lies in their ability to solve problems when intuition fails.

To illustrate the difference in the two approaches, consider a problem regarding Atwood's machine that appears on the DCI and in other literature in physics education. Figure 1 is a reproduction this question[10].


Question 13

Both systems shown have massless and frictionless pulleys. On the left, a 10 N weight and a 50 N weight are connected by an inextensible rope. On the right, a constant 50 N force pulls on the rope. Which of the following statements is true immediately after unlocking the pulleys?

(a) In both cases, the acceleration of the 10 N blocks will be equal to zero.
(b) The 10 N block on the left will have the larger upward acceleration.
(c) The 10 N block on the right will have the larger upward acceleration.
(d) The tension in the rope on the left system is 40 N.
(e) In both cases, the 10 N block will have the same upward acceleration.

Figure 1. Question regarding Atwood's Machine in the DCI.

I remember having difficulty solving this problem in high school and into my undergraduate studies. I at first thought that the systems should have the same behavior, entirely circumventing any analysis. After realizing that this was incorrect, I slugged my way through drawing the free body diagrams of the separate elements of the system, writing the system of equations, and arriving at the correct answer. But something still nagged me – *why* did the systems behave differently? After some thought and discussion, I either realized or it was suggested to me that since the scenario with the 50N weight had more mass than the one with the 50N force, it should accelerate more slowly. This seemed satisfying, and it is this sense of deeper understanding that I challenge my students to attain.

But even here there is a price. Many students, including myself, believe that appealing to the deeper understanding is *the* expected approach and shun long-hand procedure as a crutch. Even during my graduate studies, I clung to a notion that I needed to intuit the elegant solution directly, but it was not until I embraced the use of procedure that I really developed what I now consider to be my expertise in mechanics. I suspect that for reasons ranging from laziness to sincere sense of requirement and curiosity, students circumvent the application of procedure and attempt to identify a simple concept (which often corresponds to a simple formula) to solve a problem.

I also remember very clearly a poignant incident that further illustrates the distinction between procedure and concept[22]. Two former students once asked me about the design of a crane hook for a project in another course. Their initial question boiled down to the determination of the reactions of a simply supported beam with a uniform load. I have no doubt that the students could analyze an idealized simply-supported beam in isolation; their problem was that they could not "see" this idealization embedded in their problem. Had they taken the effort to draw a free body diagram as their natural starting point, rather than "trying to find the underlying secret", they might have through procedure been able to answer their own question.

For these and other reasons, I emphasize the teaching of procedure in mechanics, above the use of intuition. I have developed a one-page "problem-solving strategy" [see Appendix A] that I include in the syllabus for my mechanics classes. The procedure contains 4 elements (Kinematics, Free Body Diagrams, Writing Equations, and Solving & Checking). I use the word "element" rather than "step" to communicate that in various problems, the order of some of the elements might depend on how the problem is framed.

Truly, a number of engineering educators and researchers believe that teaching procedural skills is important. Gray et al.[8] describe a 5-step procedure for statics and use this as the basis for their upcoming textbook. This procedure shares many common features with mine, particularly in distinguishing the writing and solving as separate steps. Undoubtedly, many mechanics instructors employ similar approaches in their classes. Taraban et al.[35,36] cite the importance of procedural knowledge in engineering education and have developed an interviewing procedure to capture some dimension of students' abilities to use procedural knowledge. And a recent paper by Taraban identifies several procedural methods that have been developed for statics, and codes them according to levels of cognition[37].

5. Measuring Procedural Knowledge and Its Impact

Having made the point that teaching how to think systematically and procedurally is essential in mechanics, how can the learning and facility with procedural knowledge be measured? I will discuss the role of interviews and direct review of student exposition as possible approaches.

Interviews. Interviewing students is a well-developed method in educational research, including in engineering mechanics, that enables a detailed understanding of student thought process to be observed, recorded, and analyzed. One interviewing technique is to directly tape or transcribe student comments and writings as they solve problems. Another technique is to ask directed questions of students about why or how they solved a problem. I make no attempt to provide a systematic review of studies employing interviewing here.

Interviews require time-intensive efforts, both at the data collection and analysis, and typically involve dozens of students or less. Interviews seem to be used primarily to assess student thought, although as with concept inventories, they can also be used to assess effectiveness of different modes of teaching. Taraban et al.[35,36] recently developed a careful interview process to record student thinking in introductory thermodynamics, identify student thinking against a hierarchical scale of cognitive level, and measure levels of cognition evoked by different levels of interactive instruction delivered by computer. A primary result of their work is that students express higher-level cognition in the presence of interactive questions and quizzes than in modes of text and text-graphics. Another effort to study cognition through interviews is Kowalski et al.[16].

Hestenes et al.[14] used interviews in conjunction with the FCI to substantiate the ability of the FCI to diagnose misconceptions. Clement also used interviews to probe student misconceptions[4]. Coupling concept inventories with interviews to better understand student misconceptions is now emerging in engineering as well[1,19].

A few years ago, colleague Adeeb Rahman, student Josh Bostwick, and I conducted a series of student interviews[20]. We did not publish the results, which I summarize here. We selected three students who had completed Dynamics whose final grades were A, B, and D (the D-student was viewed to be an under-achiever). Each student was asked to solve four homework-type problems (one per week) from Statics and Dynamics in a structured interview format in our presence. Working under the premise that students learn, in part, through being allowed (and even encouraged) to pursue their own erroneous reasoning, students were asked to solve each problem in a series 10-minute stages (up to 4): at Stage 0, students read the problem and have the chance to ask questions (which might not be answered by the interviewer); during Stages 1-4, students solve the problem by hand, showing all reasoning in writing, but without dialogue. At the end of each stage, the interviewer interrupts the student to review his or her progress. During these breaks, the interviewer would invite questions from the students and identify areas of concern, sometimes affirming correct reasoning, but never offering corrections. At the end of the problem, students were asked to articulate what they learned from the process.

Despite the small sample size, some trends emerged that a more rigorous study could potentially verify. First, during this process, each student was able to arrive at the correct answer and produce the requisite reasoning to support the answer (the problems were designed such that absence of reasoning would be unlikely to allow student to reach the correct answer, and student responses were scrutinized to ensure that correct answers did not arise coincidentally). This is encouraging because it provides plausibility that students are capable of following and articulating the essential steps of fundamental procedure with targeted assistance. We did not demonstrate whether such a small set of exercises (coupled with the students' prior instruction) is sufficient to ensure that the students will be able to solve problems independently.

Second, we observed that all students were to some degree impeded by their lack of fluency with certain fundamental concepts. For example, two of the problems contained structural elements that are best modeled as two-force members, but each student at least once represented the end forces on such members as independent orthogonal force components. Also, each student expressed hesitancy of how to incorporate rotary inertia or the masslessness of a rigid body. These results seem to match some of the findings that went into the development of the DCI.

Review of Students Work. Another method that can be used to assess procedural knowledge and cognition is the careful review of student work, either by the instructor or peers. An example would be to incorporate a process of feedback and revision in homework exercises, a process that is common in the humanities, but which is relatively absent as a standard teaching practice in engineering[21]. I briefly review some recently published results in engineering that incorporate some method of direct review of student work.

Rahman, Bostwick, and I undertook a study in which we carefully critiqued written responses on homework papers from Dynamics and Strength of Materials[22]. The over-riding question that we sought to address was: "what is the relationship between the student's answer and the process of reasoning that they undertook to solve the problem?", and our goal was to collect data to illustrate the frequencies with which students get the "right answer for the right reason", "right answer for the wrong reason", "wrong answer despite good reasoning", or "wrong answer due to wrong reasoning". To this end, we carefully devised and implemented a scoring system that could be used to measure the quality of an individual student's response, and which would enable the aggregate tabulation of data aligned with the goals that we sought. We selected three topics to examine: Vectors, Coordinates and Sign Conventions ("VCS"; Kinematics), Free Body Diagrams ("FBD"; Kinetics and Constitution), and Units ("UNITS"; Kinetics and Constitution). These topics were selected based on our experience with issues that impede student problem-solving.

We examined a total of 105 papers. In this sample, we discovered that students arrived at the correct answer 67% of the time. But among the students arriving at the correct answer, only 27% expressed what we considered to be complete and correct supporting reasoning. Among the students who did not arrive at the correct answer, only 2% exhibited complete and correct reasoning within the framework that we established (other significant errors that we did not directly measure were committed).

Another approach to directly reviewing student work is develop a process of peer review. Hamilton[13] describes a procedure that he has developed in his courses to mimic engineering professional practice by requiring students to review the work of their peers and provide critiques at a draft stage. Out of this process, final work is improved, and communications skills are also developed. Cloete[5] emphasizes the importance of self-reflection in fostering critical thinking. Although these efforts did not attempt direct measures of student cognition, perhaps in the future, data can be collected to evaluate student cognition in peer review exercises.

Longer-term Assessment. To complement direct observation of student reasoning and cognition during problem-solving exercises, how can we determine if students are retaining knowledge and problem-solving skills into the future? Most methods of assessing student cognition and teaching methods – concept inventories, interviews, and detailed reviews of work – are usually employed during short time frames (e.g., most pre-test/post-test evaluations occur over the duration of a single semester). While these results are meaningful, additional research efforts should be devised to understand student learning over longer time periods, such as longitudinal studies[38].

However, a principal difficulty with longitudinal studies is that experimental and control groups decay and become corrupted over time. Unless undue restrictions are placed on students (e.g. constraining them to follow a specialized course sequence at the pleasure of the researcher), it is becomes more and more difficult to associate learning outcomes with specific teaching strategies that were delivered in prior years.

A good next step to pursue, would be to extend assessment to evaluating student performance in sequel courses. For example, various learning outcomes in Statics could possibly be measured by student performance in Dynamics or Mechanics of Materials.

Froyd et al.[7] conducted a study of student performance in a combined Statics/Dynamics class as a function of student understanding of related concepts introduced in a prior freshman design seminar. They found statistically significant differences in student peformance on statics problems that were similar to statics design problems in the seminar; they found no significant difference in student performance in dynamics problems, despite having some dynamics-related projects in the seminar.

With Bostwick and Dressel, I conducted retrospective analyses of student performance in Fluid Mechanics as a function of their instruction in Dynamics, and student performance in Structural Analysis as a function of their instruction in Statics[26]. To hint at the possible lasting impact of understanding procedure, I compared outcomes in the later courses (Fluids, Structural Analysis) of students who were exposed to a strong philosophy of using systematic procedure in my mechanics (Dynamics, Statics) courses with the outcomes of students who had taken mechanics with other instructors. I discovered some evidence suggesting that the procedure-based teaching that I delivered had positive impact in the Dynamics → Fluid Mechanics association, but a slight negative impact in the Statics → Structural Analysis association. This data is inconclusive, not only because the results are mixed, but because there was not a proper control or experimental group available. However, I believe that in principle, this type of assessment should be conducted by researchers in order to ascertain to what degree student learning persists.

6. Role of Textbooks in Student Learning and Cognition

Whereas much attention has been paid to evaluating student knowledge and cognition and effectiveness of teaching strategies, comparatively little attention has been paid to textbook quality and their role in fostering student learning. I know of no study that has attempted to show a direct link between textbook quality and student learning. In such a study, it would be difficult to control for textbook effect in the presence of many other environmental factors (e.g. instructional methods, student use of text). I will nevertheless assert that at a minimum, textbooks should consistently model the appropriate methods and approaches that are expected of students, and therefore they should be carefully examined. However, as has been pointed out by other educators, good teachers are able to compensate for flaws in textbooks, and as appropriate, use textbooks as complementary instead of primary sources.

Of the research that does exist related to quality and evaluation of engineering textbooks, the bulk to be focused on K-12 math and science texts. In particular, the AAAS Project 2061 includes a very systematic review of middle and high school science texts.

I located some work regarding evaluation of undergraduate engineering textbooks. McClelland[17] commented that most undergraduate mechanics texts are replete with inconsistent and imprecise definitions, leading to student confusion. He did not, however, document or cite any specific examples from textbooks, and proposed complicated alternative definitions that he stated are universally applicable and accurate. Rosati[29] conducted a study of student attitudes of mechanics textbooks by surveying 110 intersession students over a three year period. His study did not identify any texts, but provided student comments indicating that they preferred the text that gave briefer rather than lengthy explanations. He also discovered that students tend to use the textbook as a secondary, and not a primary study tool. And when students did consult the text, they primarily studied prepared sample problems or worked other problems; actual reading of the text was the least engaged mode of use. Hughes et al.[15] and her colleagues did conduct a review of specific texts in biological engineering, documenting principally the topics covered by six different texts (which were identified by title and author). However, this study did not provide any data on the quality of the content.

My own interest in studying textbooks parallels my commitment to teaching students how to reason systematically and helping them to develop procedural knowledge. I am generally dismayed by corner-cutting that appears in so many standard textbooks, both in the text and in worked sample problems. Early in my teaching career I developed the attitude that I needed to “teach around the text” by providing additional explanations, insights, approaches, and probing questions. I imagine that many instructors do likewise.

In an attempt to quantify the reasons for these attitudes, Rahman, Bostwick, and I reviewed several standard textbooks, first against the same topics as we reviewed student work (VCS, FBD, UNITS)[22] and then against other techniques that we postulated were relevant to strengthening rigor in student reasoning[23]. In general, we discovered inconsistencies in presentations across the different texts, although two texts met the majority of our expectations. We also argued that it is imperative for textbooks to present material in a manner that does not undermine the teaching of rigorous systematic reasoning.

As technology evolves, more and more texts and other teaching resources are moving online. While the era of the traditional textbook may be slowly coming to an end, the expectations with online materials should remain the same. Therefore, the principles of textbook evaluation that we outline here can be applied to online materials. As a next step in this work, I would like to solicit interest from other colleagues to develop a more robust textbook evaluation rubric and process. There would be great value in a body of educators convening to establish general features against which texts and other teaching materials can be evaluated.

7. Conclusions and Discussion

I have argued that studying patterns of student understanding, learning, and thinking in mechanics will serve the study of engineering pedagogy as a whole. Another reason to focus attention on mechanics pedagogy is that a large body of data is available from mechanics educators, including those in physics and engineering. Most notably, mechanics educators have contributed a significant body of research regarding student misconceptions and cognition in mechanics, including the use of concept inventories to identify and address those misconceptions.

Concept inventories have and continue to be used to evaluate effectiveness of various teaching pedagogies. This research has yielded compelling evidence to demonstrate that engaged and interactive teaching methods foster greater learning, and the continued development and use of concept inventories, such as through the Foundation Coalition, will be fruitful for helping students learn concepts and for helping researchers better understand what students learn.

As I mentioned, however, there are limitations in the ability of the concept inventories to measure student ability to identify concepts in more complicated settings. For this reason, interviewing techniques are also helpful in studying student cognition. In addition, assessing student performance in sequel courses is another method (perhaps somewhat indirect) that can be used to measure the degree to which student retain knowledge of concepts and develop the ability to apply them.

However, concept inventories, which present idealized problems focused on single concepts, are limited in their ability to measure students' ability to apply and transfer knowledge. I suggest that development of procedural knowledge is also critical. Moreover, other, complementary assessment methods, such as interviews and detailed analysis of student written work, can provide direct insights into how students use procedure, and whether the use of procedure is, in fact, a strong indicator of problem-solving ability. In addition, longer-term studies are necessary to track the degree to which students retain knowledge and can apply it new situations. Research is beginning to emerge in which student performance in a given class is being compared to learning outcomes in prior courses. I believe that further efforts in this area should be undertaken in order to develop a better understanding of the learning that truly takes place in foundational courses such as in mechanics.

Finally, I argue that textbooks and other teaching materials should emphasize the procedural knowledge that is so vital to problem-solving. I advocate for continued studies of these materials to provide another window into the types of problem-solving habits that students might adopt, and for new materials to be developed that foster procedure-based problem-solving methods.

Appendix A. Statics Problem Solving Strategy[24]

University of Wisconsin-Milwaukee **Department of Civil Engineering and Mechanics**

Statics Problem Solving Strategy

In general, include each of the following elements in solving a statics problem. Note that in some cases, these elements may not occur in sequential order (this is because solving part of a problem may lead to a clearer understanding of a later part of a problem, but you might not be able to see that ahead of time).

1. **Determine the kinematics.** Do you know anything about the geometry of the system before doing any calculations? **Clearly state or define** all variables, coordinates, base vectors, and their sign conventions. **State any assumptions** (*e.g.* type of constraint) and consider how these assumptions affect your problem.

2. **Draw clear Free Body Diagrams** (FBD) that illustrate all possible forces and moments that may be present. Care should be taken to draw forces at their appropriate points of application, and in the correct direction if that information is given. **State any assumptions** (*e.g.* frictionless contact) and consider how these assumptions affect your problem.

3. **Write Equilibrium Equations**. Depending on the application, some or all of the following equations will be required:

   a. Force Equilibrium

   $$\sum \mathbf{F} = \mathbf{0}$$

   b. Moment Equilibrium

   $$\sum \mathbf{M}_{/o} = \mathbf{0}$$

   These equations can be applied to any given system or sub-system. Part of your job is to consider how to apply these equations, and to which system – the FBD will help you decide. **No equations should be written without a corresponding FBD!** For example, in a structure with several components, you should be clear to state if your equations apply to the entire structure or a portion of it. **Also, note how few general principles you need to use in this course!**

4. **Solve Equations.** Check that your equations are *well-posed* (*e.g.* #unkowns = #equations, physical units balance, vectors balance). After solving the equations, double check that the units balance and consider if the order of magnitude is sensible.

NOTE: This strategy will not be completely applicable until Part II of the course, when equilibrium is studied. Part I deals learning the 'language' of vectors, and is associated with Kinematics (Item 1).

Bibliography

1. Buck, John R., Kathleen Wage, Margaret A. Hjalmarson, and Jill K. Nelson. "Comparing Student Understanding of Signals and Systems Using a Concept Inventory, a Traditional Exam, and Interviews", *Proceedings of the 37th ASEE/IEEE Frontiers in Education Conference,* 2007.

2. Brown, Shane, Devlin Montfort, and Kip Findley. "Development, Implementation, and Assessment of a Bending Stress Tutorial", 37th ASEE/IEEE Frontiers in Education Conference, 2007.

3. Clement, John. "Students' Preconceptions in Introductory Mechanics", American Journal of Physics 50(1), pp. 66-71, 1982.

4. Clement, John. "Students' Alternative Conceptions of Mechanics", in *Proceedings of the First International Seminar on Misconceptions in Science and Mathematics*, 1990.

5. Cloete, Adrian. "Solving Problems or Problem Solving: What are we Teaching our Students?", *Proceedings of the ASEE Annual Conference & Exposition*, 2001.

6. Evans, D. L., G. L. Gray, F. Costanzo, P. Cornwell, and B. P. Self. "Rigid Body Dynamics: Student Misconceptions and Their Diagnosis", in *XXI International Congress of Theoretical and Applied Mechanics*, Warsaw, Poland, August 15–21, 2004.

7. Froyd, Jefferey, Xiafeng Li, Arun Srinivasa, William Bassichis, Jacque Hodge, and Donald Maxwell. "How do Students in a Project-Based First-Year Engineering Curriculum Perform in a Sophomore Mechanics Course?", *Proceedings of the ASEE Annual Conference & Exposition*, 2007.

8. Gray, Gary L., Francesco Costanzo, and Michael E. Plesha. "Problem Solving in Statics and Dynamics: A Proposal for a Structured Approach", *Proceedings of the ASEE Annual Conference & Exposition*, 2005.

9. Gray, Gary L., Don Evans, Phillip Cornwell, Francesco Costanza, and Brian Self. "Toward a Nationwide Dynamics Concept Inventory Assessment Test", *Proceedings of the ASEE Annual Conference & Exposition*, 2003.

10. Gray, Gary L., Francesco Costanza, Don Evans, Phillip Cornwell, and Brian Self. "The Dynamics Concept Inventory Assessment Test: A Progress Report and Some Results", *Proceedings of the ASEE Annual Conference & Exposition*, 2005.

11. Hake, Richard R. "Interactive Engagement vs. Traditional Methods: A Six Thousand Student Survey of Mechanics Test Data for Introductory Physics Courses", *American Journal of Physics* 66(1), pp. 64-74, January 1998.

12. Halloun, Ibrahim Abou, and David Hestenes. "The Initial Knowledge State of College Physics Students", *American Journal of Physics*, 53(11), pp. 1043-1048, November 1985.

13. Hamilton, Scott R. "Peer Review: Modeling Civil Engineering Practice, Another Way to Improve Learning", ASEE Annual Conference & Exposition, 2005.

14. Hestenes, David, Malcom Wells, and Gregg Swackhamer. "Force Concept Inventory", *The Physics Teacher* Vol. 30, pp. 141-158, March 1992.

15. Hughes, Kerry, Dawn Farver, Ann Christy, and Marybeth Lima. "A Review of Texts for Biological Engineering Courses", *Proceedings of the ASEE Annual Conference & Exposition*, 2001.

16. Kowalski, Eileen and Joe Manous. "Cognitive Learning in Introductory College Science Education", *Proceedings of the ASEE Annual Conference & Exposition*, 2006.

17. McClelland, J.A.G. "Persistent Inconsistencies in Textbook Mechanics", in *Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*, 1993.

18. McDermott, Lillian C. "Research on Conceptual Understanding in Mechanics", *Physics Today* 37(7), pp. 24-34, 1984.

19. Montfort, Devlin, Shane Brown, and Kip Findley. "Using Interviews to Identify Student Misconceptions in Dynamics", 37th ASEE/IEEE Frontiers in Education Conference, 2007.

20. Papadopoulos, Chris, Adeeb Rahman, and Josh Bostwick. "Assessing Critical Thinking in Mechanics in Engineering Education", *Proceedings of the American Society for Engineering Education Annual Conference & Exposition*, 2006.

21. Papadopoulos, Chris, Adeeb Rahman, and Josh Bostwick. Interviews to document student reasoning in Statics and Dynamics, unpublished, 2003.

22. Papadopoulos, Chris and Josh Bostwick. Web-based query of use of feedback and revision of student work in engineering, unpublished, 2004.

23. Papadopoulos, Chris, Josh Bostwick, and Andrew Dressel. "Promoting Holistic Problem-Solving in Mechanics", *Proceedings of the American Society for Engineering Education Annual Conference & Exposition*, 2007.

24. Papadopoulos, Chris. Statics Problem-Solving Strategy. University of Wisconsin-Milwaukee, 2003.

25. Prescott, Anne, and Michael Mitchelmore. "Student Misconceptions about Projectile Motion", Proceedings of the 28th Annual Conference of the Mathematics Education Research Group of Australia, 2005.

26. Reed-Rhoads, Teri, P.K. Imbrie, Kirk Allen, Jeff Froyd, Jay Martin, Ron L. Miller, Paul Steif, Andrea Stone, and Robert Terry. "Panel – Tools to Facilitate Better Teaching and Learning: Concept Inventories", *37th ASEE/IEEE Frontiers in Education Conference*, 2007.

27. Richardson, Jim, Paul Steif, Jim Morgan, and John Dantzler. "Development of a Concept Inventory for Strength of Materials", 33rd ASEE/IEEE Frontiers in Education Conference, 2003.

28. Richardson, Jim. "Concept Inventories: Tools for Uncovering STEM Students' Misconceptions", AAAS, 2005.

29. Rosati, Peter. "Mechanics Texts are for Students", *Proceedings of the ASEE Annual Conference & Exposition,* 1996.

30. Sjøberg, S. Lie. Technical Report 81-11, University of Oslo, 1981.

31. Steif, Paul. "An Articulation of the Concepts and Skills Which Underlie Engineering Statics", *Proceedings of the 34th ASEE/IEEE Frontiers in Education Conference*, 2004.

32. Steif, Paul S. and John A. Dantzler. "A Statics Concept Inventory: Development and Psychometric Analysis", *Journal of Engineering Education*, October 2005.

33. Steif, Paul S. and Mary A. Hansen. "New Practices for Administering and Analyzing the Results of Concept Inventories", *Journal of Engineering Education*, July 2007.

34. Sweeney, Shannon, Richard Englund, and Robert Edwards. "Direct Assessment of Mechanics of Materials Learning with Concept Inventory", *Proceedings of the ASEE Annual Conference & Expostion*, 2007.

35. Taraban, Roman, Edward E. Anderson, Alli DeFinis, Ashlee G. Brown, Arne Weingold, and M.P. Sharma. "First Steps in Understanding Engineering Students' Growth of Conceptual and Procedural Knowledge in an Interactive Learning Context", *Journal of Engineering Education*, January 2007.

36. Taraban, Roman, Alli DeFinis, Ashlee G. Brown, Edward E. Anderson, and M.P. Sharma. "A Paradigm for Assessing Conceptual and Procedural Knowledge in Engineering Students", *Journal of Engineering Education*, October 2007.

37. Taraban, Roman, Edward E. Anderson, Curtis Craig, Jacob Fleming, Alli DeFinis, and Ashlee G. Brown. "An Assessment of Problem Solving Strategies in Undergraduate Statics", Conference *On Being an Engineer: Cognitive Underpinnings of Engineering Education*, Texas Tech University, February 2008.

38. Turns, Jennifer, Cynthia J. Atman, Robin S. Adams, and Theresa Barker. "Research on Engineering Student Knowing: Trends and Opportunities", *Journal of Engineering Education*, January, 2005.