AC 2012-3932: USING SELF-ASSESSMENT IN AN INTRODUCTORY STRUCTURES COURSE FOR CONSTRUCTION MANAGERS

John Tingerthal P.E., Northern Arizona University

John Tingerthal joined the construction management faculty at Northern Arizona University in 2007. His engineering career spans a wide variety of design and forensic engineering experiences. He spent the first eight years of his career performing structural consulting engineering in Chicago. This work culminated with design work on the Minneapolis Public Library and the Overture Center for the Arts in Madison, Wis. He was also involved with forensic investigations in Iowa and Wisconsin and participated in structural coordination efforts at Ground Zero in September of 2001. He holds professional engineering licenses in the states of Arizona and Illinois. He is currently working on a doctorate of education in curriculum and instruction with an emphasis in higher education. His academic interests lie in the field of student-centered learning and teaching. He has been a primary instructor in a transdisciplinary course that incorporated engineering, construction, cultural anthropology, and emergency medicine in an immersive experiential setting that was aimed to prepare students for international development projects. Tingerthal is a member of Engineers without Borders, ASCE, AISC, and the Building Smart Alliance and advises the construction management student organization (CMO). He coordinates NAU's teams for the Associated Schools of Construction Regional competition in Reno, Nev., and coaches the BIM team.

©American Society for Engineering Education, 2012

Using Self-Assessment in an Introductory Structures Course for Construction Managers

Abstract

Assessment of student performance is a necessary component of every academic program, but all too often this is a one-way street with only the instructor performing the evaluation. Using assessment of learning by the students themselves is an approach that encourages students to actively engage in their own learning. Studies have shown that self-assessment can have a positive effect on achievement and that such assessments can be a reliable source of data. It is the formative use of their own assessments by the students that will allow them to focus and close the gap between current and desired performance. The objective of this paper is to demonstrate an approach to using a self-assessment of conceptual understanding in an introductory construction management structures course to help students improve learning of concepts that are personally challenging. Findings suggest that identification and focus on challenging concepts can have a positive effect on learning. In addition, the cyclic nature of action research used in this study provides stimulus for pedagogical improvement.

Introduction

Student assessment is both a necessary and a required part of any college curriculum.[1,2] Typically the student has little input or control over the format or content of assessments such as homework assignments, quizzes and tests: Self-assessments, on the other hand, give a student the opportunity and power to evaluate his or her own performance. These evaluations can take the form of reflective essays, confidence ratings of conceptual understanding, and responses to open-ended questions. Self-assessment has been defined as "the evaluation or judgment of 'the worth' of one's performance and the identification of one's strengths and weaknesses with a view to improving one's learning outcomes."[3] Empirical evidence shows that self-evaluation can contribute to student achievement and can improve student behavior.[4] Students like self-assessment evaluations when the outcome has an effect on their grade because it puts some control in their own hands. However, certain weaknesses of self-assessments have been identified. For instance, when the results of the self-evaluations are used as part or all of the student's grade, the loss of teacher control may lead to lower standards and inflated grades. In addition, students may balk at the additional work required to carry out these assessments. Ross recommends practices that can make self-assessments more effective: The instructor should explicitly define criteria, train students in applying the criteria, provide good feedback, and provide students help with using the data.[4] This paper aims to illustrate a method of fulfilling this last recommendation through the use of a self-assessment to focus preparation for a final examination.

Assessment can be used for both formative and summative evaluation. Formative assessment is intended to provide "feedback" on performance in order to improve and accelerate learning,[5] while summative assessment is intended to provide “feedout” that can be treated as an indicator of not only the student’s performance, but also of the performance of the faculty and the curriculum itself.[6] In college courses, homework assignments and quizzes are typical examples of formative assessments, while final examinations are typically used in a summative capacity. Although results of a final examination could be utilized by the student to improve and continue learning, once the final grade is assigned for a course, there is usually little motivation to continue to improve on the material. Final exams can provide the motivation for last minute learning, but the typical high stakes nature of them can also cause stress. Poor performance is assumed to indicate lack of comprehension of the concepts, but could also be attributed to other factors such as test anxiety.[7] Students can be overwhelmed by the sheer quantity of concepts they need to know for a final exam and may have difficulty even identifying where to focus their studying energy. Students are less motivated when they have little control over classroom activities [7], which is usually the case in a traditional final examination.

In an effort to address these issues surrounding a final examination and to provide an evaluation that is personalized to the individual student, an approach has been developed to use self-assessments in an introductory structural mechanics course. This paper describes an approach to creating individualized final examinations through a process that helps students identify and focus on concepts that are personally most difficult. The findings suggest that getting students to engage in metacognitive activity about the topics in a course, then having them focusing on problematic or troublesome concepts can improve performance on those same concepts. The self-assessment data also provides strong feedback to the instructor on which concepts are the most challenging to students.

Background

Statics is a branch of mechanics that is concerned with the equilibrium of forces on a rigid body. In structural, mechanical, aerospace and related fields of engineering, the university curricula contain a course in statics which is the foundation for all subsequent studies in mechanics. Goldfinch asserts that the difficulties that students have in fundamental mechanics courses is a widespread and persistent problem.[8] Steif and Dóllar posit that improving student learning in statics deserves significant attention and have subsequently devoted much effort to this cause.[9] They propose actively engaging students in the learning process, using iterative inquiry methods and scaffolding upon student’s existing knowledge. The foundational concepts in statics are what Meyer and Land call ‘threshold concepts.’ Threshold concepts are conceptual building blocks that, when mastered, progress the understanding of a particular subject matter. Mastering a threshold concept is likely to shift a person’s perception of the subject matter in a way that exposes previously hidden relationships in an irreversible way.[10]

Construction management students are introduced to statics and mechanics within the larger framework of general structural behavior in an introductory course in structures. While the concepts do not differ from those studied by engineers, construction management students typically have only completed one semester of classical non-calculus based physics and precalculus math, and thus have a weaker technical background than engineering students. In addition, construction management structures courses include topics that engineers cover in more depth over the course of many classes. This makes the topic of structural mechanics especially difficult for construction management students.

Method

The present study is a the result of action research carried out over the course of three semesters in an introductory course in structures as part of a construction management program at a public, southwestern university. The participants included students (*N=66*) enrolled in Introduction to Structural Design during the Fall 2010 (*N=20*), Spring 2011 (*N=24*) and Fall 2011 (*N=22*) semesters, constituting a convenience sample of sophomore level construction management students studying structural analysis.

Action research "is about the systematic study of attempts to improve educational practice by groups of participants by means of their own practical actions and by means of their own reflection upon the effects of those actions."[11] In the case of this study, the practical actions took the form of creating partially individualized final examinations based on self-assessments performed by the students. Action research, using the Kemmis Model,[12] moves through a cycle of planning, action, observing and reflecting, returning back to a revised plan, followed by a repeat of the cycle. Action research is categorized as non-experimental due to the lack of random assignment of treatment and a lack of control. The following process was thus developed over the course of the study, with slight modifications occurring in each cycle informed by observation and reflection.

During the last week of each semester, a three-step process was initiated that included having the students complete a personal assessment of their own confidence with course material, a reflective exercise that probes topics and a follow-up assignment to focus their study for a final examination. Since each step is based on the results of the previous step, students are not informed of the entire process at the outset in an effort to keep them from tailoring their responses to simplify later tasks. These three steps are crafted to assist the students with identifying the concepts that are personally problematic.

In the first step the students were presented with a self-assessment inventory of concepts from the current semester (see Appendix A), listed in the order that they were introduced in the classroom. Students were then asked to rate their level of understanding and confidence with each topic on a scale from 1 to 5. The descriptors for each level are listed in Table 1.

Table 1.

Levels of Understanding

Level	Description
1	Fully understand and confident that I could answer questions / solve problems on this topic now
2	Understand but would have to review to answer questions / solve problems on this topic now
3	Have an idea of what this is in concept, but need Considerable study to understand
4	Recognize the topic but would need extensive studying to understand
5	I don't even know what this means

After completing this matrix, students were given a worksheet upon which they must complete two tasks (see Appendix B). In the first task, they selected three items that they rated '1' (or their top three if they did not have three that were all rated a '1') and were instructed to compose exam-style question and answer for each of the three topics. The second task involved selecting four troublesome items rated either '4' or '5' and posing two questions about each for which the student does not have an answer (for a total of eight questions). The results of this second task were then used as the basis for the third task.

In preparation for the third task, the instructor collected and recorded responses from the first two steps. Each concept was then ranked on a basis of average score across all student responses. In the class immediately following the completion of the first two steps, the ranked list of average scores was presented to the class as a whole and was used as a guide for an in-class review, starting with topics that were ranked highest (lowest confidence) by the students. The students were then given back their worksheets and instructed to use them to review for the final examination and to select and answer four of the eight troublesome questions. Their one-page response to each of these four questions would constitute a take-home final exam, making up 20 percent of the total final exam grade (the other 80 percent would be a two-hour sit-down exam).

The data for this study consists of results of the self-assessment inventories, topics chosen for the take-home exam, and performance on the sit-down exam. Questions on the sit-down exam were cross referenced to the topics from the self-assessment. In order to compare student performance on topics that were chosen for the take-home exam with those not selected, three scores were generated from the data: Score on topics chosen for take-home, X_{TH}; score on topics rated with the lowest confidence (not including those chosen for take-home), X_{LC}; and score on topics rated with highest confidence, X_{HC}. Each score, X_i, was calculated as the ratio of points earned to possible points in each category. The score for each individual student was based on the self-identified topics. Since the topics varied slightly each semester, there was an overlap of only 23 of 32 topics across the three semesters, and of 66 participants, only 60 provided full set of self-assessment data. Course policy allowed for students at 95% to opt out of the sit-down portion of the final exam. In addition, some exams were not available for evaluation, so analysis of performance is based on N=46.

Results

The first result worth noting is the distribution of rankings across the different levels of understanding (recall that low score represents high confidence, see Table 1). The distribution shown in Figure 1 shows that the most common response was '2' (41%), corresponding to "understand but would have to review to answer questions / solve problems on this topic now" while response '5' ("I don't even know what this means") garnered only 1% of the responses. Response '1' ("Fully understand and confident that I could answer questions / solve problems on this topic now") had the most variability ($s^2=1.68\%$) across the topics ranging from 7% (section modulus) to 54% (free body diagrams), while response '2' was much more regular ($s^2=0.5\%$) ranging from 24% (reactions) to 55% (stability).

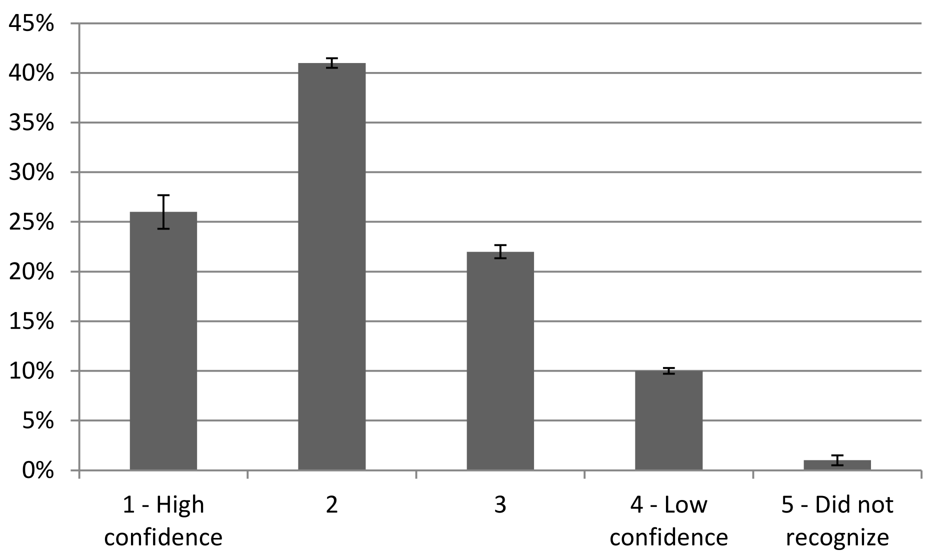

Figure 1. Distribution of levels of understanding scores assigned in self-assessment. Sixty-six participants in 32 topical categories

The average confidence levels across all topics for individual participants ranged from 1.39 to 3.46 (lower numbers indicating higher confidence). These average values are plotted in Figure 2 against student final course grade. There is generally a positive correlation between confidence and performance (*r=-.42, N=66*, reflected statistically as a negative correlation between level of understanding score and course grade). This is a low-moderate correlation.[13] It is notable that nine above average performing students rated themselves below average level of understanding (upper right quadrant) while 12 below average performing students rated themselves above average level of understanding (lower left quadrant).

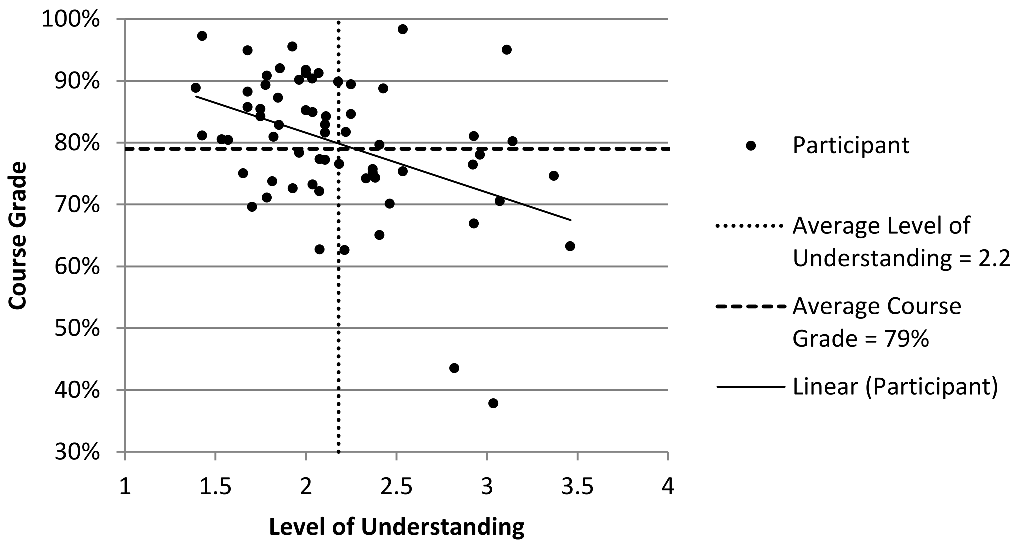

Figure 2. Participant final course grade vs. participant average level of understanding. (1 = high confidence in understanding, 4 = low confidence in understanding)

Turning now to the topics themselves (see Table 2), the average ranking of each topic ranged from 1.58 for "free body diagrams" to 2.65 for "section modulus" ($N=60$). Most confidence was shown in topics taught early in the semester and dealing with force and load. Least confidence was shown with concepts taught later in the semester, and dealing with section properties (S), bending (stress, beam moment) and pure stress and strain. These lower ranked concepts deal with more complex concepts. The notable exception to the relationship between order of presentation and ranking are the topics dealing with wood properties and design.

Looking at performance on exam, a paired-sample t-test was used to compare the scores on topics chosen for take-home, X_{TH}; scores on topics rated with the lowest confidence (not including those chosen for take-home), X_{LC}; and score on topics rated with highest confidence, X_{HC}. No significant difference was found between X_{HC} and X_{TH} ($\bar{X}_{HC} = 81.4\%, \bar{X}_{TH} = 79.5\%, t = .678, p = .501, N=46$), however there was a significant difference between X_{HC} and X_{LC} ($\bar{X}_{HC} = 81.4\%, \bar{X}_{LC} = 71.6\%, t = -2.254, p = .029$).

The quality of work submitted on the take-home exams was qualitatively better than the work produced for homework assignments. Some students engaged in meaningful reflection on the earlier difficulties with the particular topics, although this reflection was not an explicit requirement of the assignment. On the topic of solving shear and moment diagrams, one student commented "when looking at my original attempt at this problem I had the process right, I just made an error when solving for reactions" indicating the importance of mastering fundamental concepts (reactions) before moving on to more complex topics (shear and moment diagrams).

The difficulty with mastering the threshold concept of calculating beam reactions was a recurring theme.

Table 2.

Topical Rankings

Topic	Rank[1]	Average level of understanding[2]	Order of presentation during Semester
Free Body Diagrams	1	1.58	9
Load Tracing	2	1.73	4
Force (concept)	3	1.75	5
Load determination	4	1.83	3
Stability	5	1.85	1
Reading Structural Plans	6	1.95	2
Equilibrium	7	1.95	10
Reactions	8	1.98	11
Properties of Wood	9	2.03	28
Vectors	10	2.05	6
Designing with Wood	11	2.08	29
Moment (concept)	12	2.13	7
General Material Properties	13	2.15	14
Moment (calculations)	14	2.22	8
V+M diag.: equilibrium method	15	2.33	23
Temperature Effects	16	2.35	17
Centroid	17	2.37	18
Axial Stress (f=P/A)	18	2.38	12
Moment of inertia, I	19	2.50	19
Axial Strain (e = DL/L)	20	2.55	13
Beam Shear and Moment (concept)	21	2.57	21
Bending Stress (f=M/S)	22	2.60	27
Section Modulus, S	23	2.65	20

[1] Only topics that were assessed during all three semesters are reported

[2] Level of understanding value of 1 = high confidence

Another recurring theme had to do with the difficulty of getting started on a problem: “My main problem with this question [equilibrium of forces] was once again figuring out what it was asking and knowing where to start.”

Some comments indicated that the student did indeed increase their topical comprehension while completing the take-home exam, for example: “I was unsure what the difference between elongation, stress and strain was. With these four problems, I now remember and understand what the difference between elongation, stress and strain [is] and how to do the math to compute these types of problems." Regarding beam shear, one high-performing student remarked, "I never actually understood what I was solving for when I found the maximum shear in beams." He went on to describe other concepts of shear that he was familiar with and then adequately described beam shear.

Discussion:

The distribution of levels of understanding scores assigned in self-assessments shown in Figure 1 indicates that students in the sample generally believe that they understanding the topics, but do not have confidence in applying the principles to problem solving situations. This reflects a low level of cognition on Bloom's Taxonomy.[14] That very few students responded that they did not recognize the topic ("5") indicates that the course was either effective in introducing the concepts, that the students were already familiar with the concepts, or that the students chose not to demonstrate their ignorance. This last explanation is viable since the inventories were not anonymous.

It is not surprising that the self-assessment of level of understanding correlates to performance, or that this correlation is not a strong one. Performance on previous formative assessments in the class and resulting feedback gave the students an idea of how well they could perform on the various tasks. The spread may be related to the student's attribution of prior success or failure in a given area, e.g. students attributing failure to external influences may not believe that their understanding is related to assessment measures. Further, a student who has high self-efficacy and does not put fort the effort to succeed may continue to have high confidence in their understanding. One may also question the validity of using course grades to reflect understanding. These factors may explain data points in the lower left quadrant of Figure 2. Explanation of highly performing students who self-assessed at low levels of understanding (upper right quadrant of Figure 2) may be related to the notion that 'the more you know, the less you know.' High performing students may also have set high standards for themselves.

The topics ranked with the most confidence (Table 2) tended to be simple conceptual building blocks that were introduced early in the semester. This means that these topics were seen and practiced more often than the more complex concepts that relied upon them. This is consistent with the notion of improving declarative and procedural knowledge through practice.[15]

Topics ranked with the least confidence in level of understanding, on the other hand, tended to be more complex (beam bending) and abstract (section modulus, moment of inertia, shear, moment). Although these topics were taught later in the semester, any recency effect[16] was overcome by other factors, perhaps including lack of practice and degree of complexity.

When looking at the extremes of scores on the exam (X_{HC} v. X_{LC}) on an individual basis, there was an average of 9.8% difference in performance between the topics that students selected for the take-home exam and the topics in which they indicated low confidence. This does provide support for the validity of using self-assessment of level of understanding to predict performance. The insignificant difference between performance on topics on the take-home exam (X_{TH}) and those identified with high confidence (X_{HC}) suggests that the take-home exam did have a positive effect on performance in the topics studied. This result does need to be tempered by the fact that students were allowed to consult their take-home exam while completing the sit-down exam, possibly inflating their score on those topics. Nevertheless, the quality of the take-home exams and the reflective comments do support the idea that learning did occur in the process.

Limitations

This study is not without limitations. The dynamic nature of action research means that the research questions and procedures emerged and evolved over time. The slight changes between the semesters are not reflected in the aggregate data presented here, however, analysis of such may provide insight into the effectiveness of the modifications made. The sit-down final exam was not developed using rigorous psychometric methods and its reliability and validity were not estimated. Since mastery of some foundational topics is required to understand more complex ones, the topics in this study are not independent. Some topics were not explicitly assessed in the final examination, while others had different weights, thus potentially resulting in a sensitivity effect.

The procedure described relies on students not knowing that they are writing a portion of their own final exam. If word gets out, students may change their responses in order to make the take-home exam easier. The fact that students have access to the take-home exam while completing the sit-down may have an effect on the results. Finally, the best students (scoring greater than 95% coming into the final) did not take the final sit-down exam.

Recommendations and Future Work

In the next iteration of this study, a few modifications and improvements are warranted. First of all, the take-home exam should explicitly require a reflective component. This extends the metacognitive process started with the inventory itself. The take-home exam should either include all questions written so that the students do not select the ‘best of the worst’. This will keep them from avoiding the few questions that were the most problematic. Analysis of the sit-down exam indicated that the different topics were not consistently assessed (some were not assessed at all.) There is therefore room for improving the psychometric properties of the sit-down exam in order to improve reliability of the data and validity of conclusions.

As noted in the introductory paragraphs, this study addresses only one of the practices recommended by Ross (providing students with help using the data). The other three recommendations, (explicitly defining criteria, training students in applying criteria and providing good feedback) are all areas of potential improvement.

Conclusion

Based on the results of this study, the process of having students perform self-assessment of their level of understanding of particular topics does force them to identify problematic areas and motivates them to direct effort toward improving their understanding. The results suggest that focusing on these problematic areas may improve performance. The cyclic nature of action research is an effective way to stimulate pedagogical improvement for courses, as the recommendations provided in the preceding paragraph would not have emerged without careful analysis of data and reflection on the findings of this study.

Bibliography

1. American Council for Construction Education. Document 103: Standards and criteria for accreditation of postsecondary construction education degree programs. 2010:1-32.

2. ABET Engineering Accreditation Commission. Criteria for accrediting engineering programs. *Engineering*. 2010.

3. Klenowski V. Student self-evaluation processes in student-centered teaching and learning contexts of Australia and England. *Assessment in Education: Principles, Policy & Practice*. 1995;2(2). (p. 146)

4. Ross JA. The reliability, validity , and utility of self-assessment. *Practical Assessment Research & Evaluation*. 2006;11(10). Available at: http://pareonline.net/pdf/v11n10.pdf.

5. Sadler DR. Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*. 1998;5(1):77-84. Available at: http://www.tandfonline.com/doi/abs/10.1080/0969595980050104 [Accessed August 3, 2011].

6. Knight PT. Summative assessment in higher education: Practices in disarray. *Studies in Higher Education*. 2002;27(3).

7. Ormrod JE. *Human learning*. 5th ed. Upper Saddle River: Pearson / Merrill Prentice Hall; 2008.

8. Goldfinch T. Improving learning in engineering mechanics: The significance of understanding causes of poor performance. In: *AaeE Conference*. Yeppoon; 2008:1-6.

9. Steif PS, Dóllar A. Reinventing the teaching of statics. *International Journal of Engineering Education*. 2005;21(4).

10. Meyer JHF, Land R. Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practicing within the disciplines. In: Rust C, ed. *Improving Student Learning – Ten Years On*. Oxford: Oxford Centre for Staff and Learning Development; 2003:1-16.

11. Hopkins D, Ahtaridou E. Applying research methods to professional practice. In: Lapan SD, Quartaroli MT, eds. *Research Essentials*. San Francisco: Jossey-Bass; 2009. (p. 276)

12. Kemmis S, McTaggart R. *The action research planner*. 3rd ed. Victoria, Australia: Deakin University PRess; 1988.

13. Grimm LG. *Statistical applications for the behavioral sciences*. New York: John Wiley & Sons; 1993.

14. Bloom BS, Englehart MB, Furst EJ, Hill WH, Krathwhol DR. *Taxonomy of educational objectives, the classification of educational goals, Handbook I: Cognitive domain*. New York: Longmans, Green and Co. 1956.

15. Ericsson KA, Polson PG. A cognitive analysis of exceptional memory for restaurant orders. In: Chi MTH, Glaser R, Farr MJ, eds. *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

16. Baddeley AD. How many kinds of memory? The evidence for STM. In: *Human memory: theory and Practice*. Boston: Allyn & Bacon; 1990.

Appendix A: Self-assessment Inventory

	Topic	Level of Understanding and Confidence				
		1 **Fully understand and confident** that I could answer questions / solve problems on this topic now	2 **Understand** but would have to review to answer questions / solve problems on this topic now	3 **Have an idea** of what this is in concept, but need Considerable study to understand	4 **Recognize the topic** but would need extensive studying to understand	5 **I don't even know what this means**
Before Midterm	Stability					
	Reading Structural Plans					
	Load determination					
	Load Tracing					
	Force – concept					
	Vectors					
	Moment – concept of moment					
	Moment – calculations					
	Free Body Diagrams					
	Equilibrium					
	Reactions					
Since Midterm	Axial Stress (f=P/A)					
	Axial Strain (e = □L/L)					
	General Material Properties					
	Elastic-plastic Behavior					
	Modulus of Elasticity					
	Temperature Effects					
	Centroid					
	Moment of inertia, I					
	Section Modulus, S					
	Shear and Moment in Beams (V+M) - concepts					
	V+M diag.: equilibrium method					
	V+M diag: semi-graphical method					
	Bending Stress (f=M/S)					
	Properties of Wood					
	Designing with Wood					
	Trusses					
Other						

Appendix B: Follow-up questions

Take 3 items from column 1 (or cols 1&2 if needed) (Fully understand and Confident) and write a good exam-style question (**not a True/False or Multiple choice question**)

1. Topic:__________________   Answer:
2. Topic:__________________   Answer:
3. Topic:__________________   Answer:

Take 4 items in the 4th and/or 5th columns and **write <u>two</u> questions** about unclear aspects of these topics

1. Topic:__________________
2. Topic:__________________
3. Topic:__________________
4. Topic:__________________