

Effect of Multiple Choice Testing on Student Performance in an Introductory Engineering Course

Dr. Jennifer M Peuker, University of Alaska, Anchorage

Jennifer Mott Peuker received her Ph.D. from the University of Illinois, Urbana-Champaign in 2012, where her research focus was on aluminum combustion in explosive fireballs. In addition, she has two teaching certificates from the University of Illinois Center for Teaching Excellence. In the Spring 2013 semester, she was a Term Assistant Professor at the University of Alaska Anchorage in the department of Mechanical Engineering, where she taught the freshman level engineering practices course, with an emphasis on computer programming using MATLAB and communication. Her teaching interests are in the area of thermo-fluids and freshmen engineering. Her current research is focused on the success of freshmen engineering students, and implementing a flipped classroom by using Team-Based Learning in engineering core courses. Jennifer can be reached at jmpeuker@gmail.com

Dr. Jennifer McFerran Brock, University of Alaska Anchorage

Dr. Jennifer McFerran Brock is an Assistant Professor of Mechanical Engineering at University of Alaska Anchorage. Her teaching focus is on fluid and thermal systems; she teaches classes in thermodynamics, heat and mass transfer, and fluid mechanics. She has done work with academic assessment, particularly relating to ABET. She can be reached at jmcferran@uaa.alaska.edu.

Dr. Steffen Peuker, University of Alaska Anchorage

Dr. Steffen Peuker is an Assistant Professor of Mechanical Engineering and the Director of the Thermal System Design Laboratory at the University of Alaska Anchorage. He is teaching the Thermal System Design, Thermal System Design Laboratory, HVAC Systems Optimization and Introduction to Engineering courses. His work in engineering education focuses on hands-on undergraduate engineering education in the HVAC&R area, student-industry cooperation, and developing innovative ways of merging engineering fundamentals and engineering in practice and research. Dr. Peuker's educational research also focuses on increasing student retention and success in engineering through a student success focused introduction to engineering course. He is an active member of the American Society of Heating, Refrigerating and Air Conditioning Engineers and can be reached at steffen.peuker@gmail.com.

Effect of Multiple Choice Testing on Student Performance in an Introductory Engineering Course

Abstract

This study aims to compare student performance on introductory engineering statics material by comparing the exam scores of students who are given both multiple choice (MC) questions and constructed response (CR) questions to see whether the type of exam question makes a difference in student performance and understanding. Seventy-five students in an introductory engineering course did either a MC version or a CR version of each statics problem, resulting in MC answers and a control group of CR answers to each statics problem. The students were also polled for feedback regarding their preferences of test question format at the end of the semester. All the exams were graded by one professor, and the results showed little difference between the scores on the MC versus the CR versions of a question. The average score for the MC version was 80%, while the average score for the CR version was 76%. While MC questions may not be appropriate in all circumstances, the high performance on the MC questions, and similar performance on CR questions indicates that not only do students not guess at the answer, but also are able to show understanding of basic statics problems.

Introduction

This study is intended to investigate the effect of multiple choice (MC) as opposed to constructed-response (CR)—‘traditional’ open ended problems—testing on student performance in an introductory engineering course. Most of the engineering educational literature is focused on the development of quizzes and web based questions¹⁻³. The main question this study intends to answer is: Does the use of multiple choice questions on an exam adversely affect the students’ performance? MC questions allow instructors to test a broader range of material on the exams than the traditional open-ended problem approach, and they also offer more efficiency and reliability in scoring because they are objectively rather than subjectively scored⁴. However, CR questions are often regarded as being a better teaching tool that emphasizes originality and depth of understanding. A possible drawback of using MC questions could be that this format will tempt the students to guess instead of solving a problem.

MC and CR questions are often seen as very different teaching and assessment tools, with MC questions emphasizing simple recall of facts (recognition) and CR questions giving students the opportunity to show originality and depth of understanding (generation), but in fact there is little empirical evidence to support this⁴⁻⁵, nor does empirical evidence support the notion that MC tests support poor study habits among students⁶. The College Board’s Advanced Placement (AP) tests are particularly well-suited to compare student performance on MC vs. CR questions, because these tests contain combinations of the two formats that cover the same material. Analysis of AP test scores suggests little difference in knowledge, skills or abilities measured using MC as opposed to CR questions, with correlations between MC and CR performance being especially high on AP tests for qualitative subjects, such as mathematics, physics and chemistry, and foreign languages⁴⁻⁵.

In particular, analysis of AP results for the Computer Science test (APCS) were found to have a very high correlation despite the CR questions having been developed to measure content more deeply than MC questions⁷. Explanations and caveats offered in APCS case seem to apply to college freshmen taking an introductory-level engineering course as well as to high school juniors and seniors taking an AP Computer Science course. For example, the population taking the exam would be expected to have a similar skill profile, with greater skill differentiation expected among individuals with more experience⁷. Also, CR questions at the introductory level do not represent the true length or complexity or real-world applications⁷. It is also pointed out that the scoring scheme for the APCS exam does not take into account efficiency, user-friendliness or originality⁷, which may also explain some amount of the close correlation between MC and CR results. However, this does not indicate that MC questions are inappropriate as a teaching tool at the *introductory* level, which is the issue under examination in this study.

A revised version of Bloom's Taxonomy of knowledge types divides knowledge into four categories—factual, conceptual, procedural, and metacognitive⁸. Typically, MC questions can easily test factual and conceptual knowledge, such as testing vocabulary or fundamental theories. Instructors usually use CR or traditional open ended problems to test procedural knowledge—such as setting up and solving engineering problems. Procedural knowledge can be difficult to test in a MC format; however, the exam questions given to the students in this study both in MC and CR format were designed to test student knowledge of static problem solving methods and correct application of those methods (procedural knowledge). The Statics Concept Inventory has been used to measure student comprehension of statics material using multiple choice questions⁹; however, this study aims to compare student performance on introductory statics material by comparing the exam scores of students who are given both MC questions and CR questions to see whether the type of exam question makes a difference in student performance.

Experimental Method

The sample population was taken from students enrolled in an introductory engineering course at the University of Alaska Anchorage. Students from four majors are required to take this course, Computer Systems Engineering, Civil Engineering, Electrical Engineering and Mechanical Engineering, as well as undeclared engineering majors. One of the main topics covered is an introduction to engineering statics, including free body diagrams and calculation of resultant forces.

The introductory engineering course used for this study is intended as a broad survey of the engineering profession, with introductory units on the engineering method, problem-solving, reporting and displaying project results, simple engineering mechanics and materials science, and simple circuit analysis. Students will ideally take this course during their freshman year in college. The prerequisite for this course is pre-calculus, which is the mathematics requirement for the engineering program in general, so some students will have had one or more semesters of remedial mathematics and/or science courses upon enrolling in the introductory engineering course. The class meets for two 75 minutes lecture periods per week, with class periods devoted to either traditional lecture or in-class group activities.

For the final exam, four statics problems were given, each in two formats: MC and CR. Two versions of the exam were made—Exam A and Exam B—and each exam had a different combination of the four statics problems, two that had a series of MC questions, and two that were CR, (see the Appendix for all problems). Seventy-five students did either a MC version or a CR version of each problem, resulting in MC answers and a control group of CR answers to each statics problem. The students were also polled for feedback regarding their preferences of test question/problem format at the end of the semester.

The four statics problems used were as follows. The first problem, Figures A.1 (CR version) and A.5 (MC version) presented a concurrent force system and asked the student to calculate the x and y components and the resultant force in newtons, and determine the quadrant of the resultant force. The second problem, Figures A.2 (CR) and A.6 (MC), asked the students to analyze the forces on a kite. The third problem, Figures A.3 and A.7, presented a beam and asked the students to calculate the moments about two particular points. The final problem, Figures A.4 and A.8, presented a truss having a weight hanging from the middle. The students were asked to find the force in a particular part of the truss, state whether the member is in tension or compression, and determine the minimum diameter of the cable suspending the weight.

Multiple Choice Question Design

How to properly write multiple choice questions has been well documented¹⁰ and the design of the MC version of the problems included choosing distractors or wrong answers. The correct choice of alternatives in MC questions is important and can be the more time consuming part of developing MC questions and items. There is no reason to use random alternatives which students can immediately discard because they are obviously wrong, e.g. there is no need to increase the numbers of choices in order to always have four items per MC question. For example, the second MC question for the truss problem—see Figure A.8—asks if truss member AB is in a) Tension, b) Compression or c) Zero force member. Since there are, physically, only three possible senses for a truss member any additional item would be unnecessary.

Alternatives or distractors should be chosen to give the instructor and the student feedback about possible misunderstandings. This can be accomplished by creating alternatives by applying typical mistakes student might make when doing the problem. For example, the third MC question for the beam problem shown in Figure A.7 tests the understanding of a moment. There are four possible answers provided. Answer (b) is the correct answer. Alternative (c) is wrong because of the sign convention that states that positive moments act counterclockwise (right hand rule). Therefore, alternative (c), if chosen, indicates a misunderstanding of the directional sense of a moment or right hand rule. Alternatives (a) and (d) are derived by using wrong moment arms. A student choosing these alternatives demonstrates a misunderstanding of the perpendicular or shortest distance from a point of rotation to the line of action of the force.

An analysis of the student answers reveals that 19 out of 29 students answered the question correctly. Seven students chose answer (a) and two students answer (d), indicating that nine students or 31% percent demonstrate a misunderstanding of the perpendicular or shortest distance from a point of rotation to the line of action of the force. Only one student chose (c) as

an answer. Therefore, the instructor can assess that more emphasize needs to be spent on teaching the concept of the determining the arm of a moment.

Another example of distractor choice and student answers is from question 25 of the truss problem—see Figure A.8—which tests the understanding of determining truss member forces. Table 1 shows the result of the distractor analysis.

Table 1 Distractor Analysis Question 25

Answer	Distractor Design	Number of Student Answers
a.	Switched sine and cosine, F_g positive	2
b.	Switched sine and cosine, used $F_g=150\text{ N}$	6
c.	Trivial solution	2
d.	Correct solution	18
e.	Used 50° instead of 25°	1

One of the advantages of the CR or traditional open-ended problem approach is the possibility for the instructor to give partial credit, e.g. to indicate to the student that he/she has understood one concept but not the other. MC questions might be thought of as excluding this partial credit option. Question 24 from the concurrent force system problem, see Figure A.5, is an example how partial credit can be given for MC questions.

A concurrent force system is given to the student and the first two questions (numbers 21 and 22 in Figure A.5) ask for the resultant force in the x-direction and y-direction, respectively. Question 24 asks which quadrant the resultant vector will be located. To answer this question, students need to consider their answers from question 21 and 22. If they incorrectly determined the x- and y- forces, but did understand how to determine the quadrant of the resultant force correctly, they would not receive credit for question 24. However, if they were to make the same mistake on a CR problem the instructor most likely would give partial credit for choosing the quadrant that was consistent with their previous work. Therefore, partial credit could be given for the MC problem, if the “correct” quadrant was chosen for incorrectly calculated forces.

Based on analyzing the student responses to the quadrant question of the concurrent force system problem, 36 students marked the correct answer, and 12 students did not. However, two of the 12 students actually chose the correct quadrant based on their marked answers for the force calculations, and therefore it can be argued that these students understood the concept of determining in which quadrant the resultant force lies based on x- and y-component of the resultant vector.

A more convincing argument of awarding partial credit for properly designed MC questions is question 26 from the truss problem—shown in Figure A.8—which asks to identify the sense of truss member AB. The correct answer is (a) based on the correct result for question 25 (d): determining the force in member AB. Question 25 tests the understanding of determining the force of a truss member, a different concept than determining the sense of a force member. Out of 29 students, 18 answered both question correctly and 11 students answered question 26 incorrectly. However, it turns out that 10 out of the 11 students did answer question 26

consistently with their answer for question 25, demonstrating that they understood how to interpret the sign of the force to determine the sense of the member. If each MC question is graded independently, in the presented case, 34.5% of the students would have received a 0 score for question 26 although they actually understand the concept tested.

Results and Discussion

All exams were graded by one professor in order to reduce bias and ensure the same grading scheme was used for all problems/students. Partial credit was given on the CR problems and for MC questions in which the question depended on the answer from the previous questions, e.g., determining the quadrant of the resultant vector. Figure 1 shows the average score (out of 10) for each problem. Most problems—both MC and CR type—had at least a 60%, and two of the problems had scores of at least 70%.

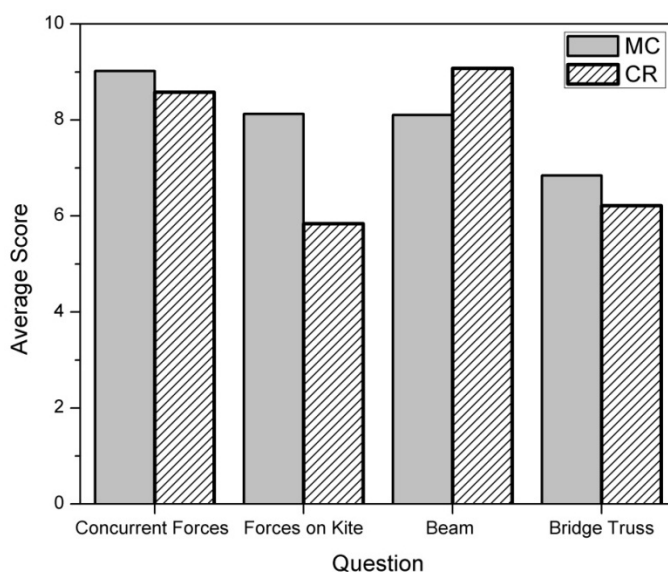


Figure 1 Average scores out of 10 for each exam question.

The high success rate on the MC versions shows that the students are not guessing, and are exhibiting an understanding of the material tested. Statistically speaking, guessing does not play a role in the success of the MC questions. Assuming an exam consists of four MC questions, and each question has four possible answers, the probability to answer all four questions correct by guessing is 0.0039 or just 0.39 percent. Therefore, students cannot be guessing their way to a passing grade. However, students might be tempted, psychologically, to guess if they encounter MC questions. Feedback from the students, though—shown in Figure 2—indicates that only 11% of the students report that they are tempted to guess, while 76% report they are not tempted to guess (13% neutral), which indicates that the students understand they have a better chance of getting the correct answer if they attempt the problem to at least narrow down the choices, enabling the students to make an educated ‘guess’, if not solving the problem correctly.

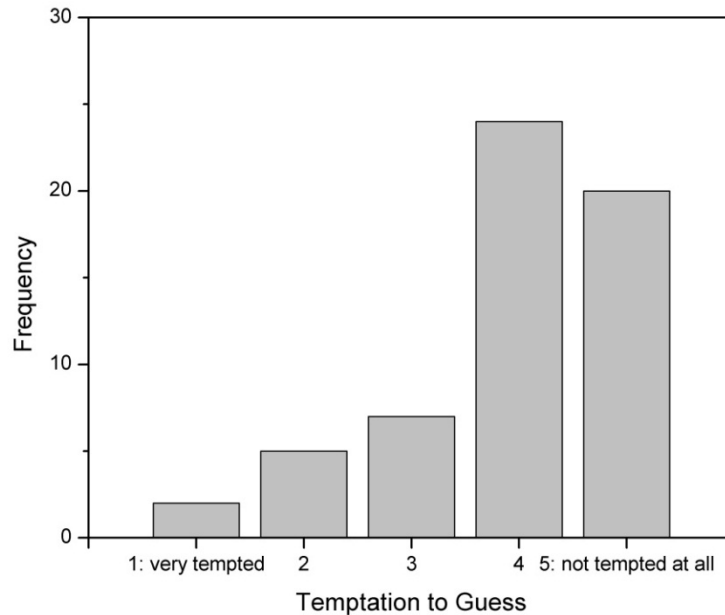


Figure 2 Students' self-reported temptation to guess on MC style questions.

Table 2 lists the average scores for each question and t-test values (two-tailed). Three of the four questions scores were not statistically significant for an error of less than 1% ($p < 0.01$), which means that there was no difference between the MC and CR scores, and therefore the style of question (MC or CR) did not make a difference in students' performance. The scores from the forces on a kite question are statistically significant at an error of less than 0.1 percent ($p > 0.001$), indicating that the type of question *did* make a difference in the student's performance.

Table 2 Average question score and t-test value

Question	MC	CR	value of t
Concurrent Forces	9.1 (N=75)	8.6 (N=59)	1.71***
Forces on Kite	8.1 (N=75)	5.8 (N=59)	5.67*
Beam	8.1 (N=59)	9.1 (N=75)	2.44**
Bridge Truss	6.8 (N=59)	6.2 (N=75)	1.25

*significant at $p > 0.001$

** significant at $p > 0.02$

*** significant at $p > 0.1$

The difference between the scores on the kite question indicates that the way a question and MC options are formulated can make a difference on student performance. In the accompanying survey on test format preference, 58% of students indicated that they feel that MC questions

allow them the opportunity to check their answers and/or get new ideas from the choices presented. Most likely, this occurred for the kite problem, from the format of the MC questions and their possible choices providing hints that were not available to the CR question group. The first part of both types of problems asked to draw the free body diagram (CR) or circle the correct free body diagram (MC). If students are able to correctly identify or draw the correct free body diagram, they are more likely able to correctly calculate the x- and y- components of the forces, determine the magnitude, and direction of the total force. The MC version of the question could help students who are unsure or who may forget a force (such as the weight) determine the correct free body diagram, boosting their chance of correctly solving the remaining questions.

In addition, 68% of the students indicated that they prefer MC to CR exam questions. Table 3 summarizes by theme the students' responses to the survey question of why they like or dislike MC questions. Fifty-eight percent of the free answer comments from students indicate that the students like to be able to check their answer, get new ideas from the answers, and feel that the MC questions boost confidence during the exam. Only four percent of the students indicated strong feelings of doubt when it came to the distractors, and only 13% of the students dislike the idea of no partial credit. A low percentage of students—eleven percent—reported they are tempted to guess. Even fewer students—four percent—perceive that they have a better chance on receiving credit on a MC questions by guessing. Finally, an interesting note, that seven percent of the students commented that MC questions are easier for the instructor to grade and note that for the students this means faster returns. In a busy and quick term, the students recognize that a faster turnaround for feedback on midterm exams and quizzes is important for applying lessons learned to future assignments.

Table 3 Comments from students about why they like or dislike MC questions.

Theme	Percent
No partial credit, or feedback	13%
Can guess	11%
New ideas from answers, Check answers, Confidence building	58%
Distractors bad, Doubt	4%
Better chance	4%
Easier grading, Quicker returns	7%

Students were asked whether and how they would study differently if they knew an exam would consist of only MC questions. Seventy-seven percent of students self-reported that they would not study differently, while 23%, or 13 students, reported that they would study differently. Of these, two students reported that they would study more because there is no partial credit available, or because they feel they would need to know how to do problems in multiple ways. The largest number, four students, reported that they would expect to “reverse engineer” their responses from the available choices and might use this opportunity to replace some studying. Several students commented that they would expect MC tests to involve less memorization, and one student in particular identified that whether a test were to be open or closed notes would affect studying more than whether it were to be MC or CR. Only three students reported that they would study less because they feel that MC questions test at a lower level than CR questions. These results are consistent with a 1968 empirical study which found that students

studying for an MC exam were motivated to perform at least as well as students studying for a CR exam⁶.

The students were also surveyed for their perception of how fair the final exam questions were compared to the course material—homework and in-class activities. The results are presented in Figure 3. Overall, the majority of students perceived the final exam questions as very fair, and no one reported thinking the exam questions were ‘not fair at all’.

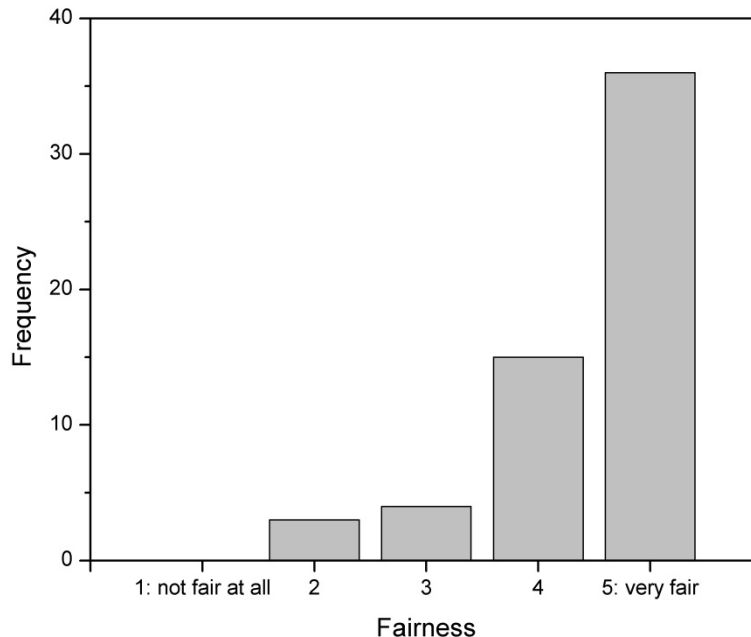


Figure 3 Student perception of the fairness of exam questions.

Based on the responses to the student surveys, the authors speculate that MC problems on an exam may be able to reduce student fear and test anxiety. Students self-reported that MC problems boost their confidence during the exam. Building confidence in freshman engineering students can also encourage more students to continue in engineering, and therefore could improve engineering student retention, and would be an interesting study to continue.

Conclusions

While MC questions may not be appropriate in all circumstances, the high performance on the MC questions, and similar performance on CR questions indicates that not only do students not guess at the answer, but also are able to show understanding of basic statics problems. MC questions are good at testing calculations, but can be poor at testing problem set up, depending on how the problem is formulated and what the MC options are. Although more difficult to implement because of the more sophisticated design and analysis of MC questions, it is possible to give "partial credit" even for MC questions if, e.g., a MC question tests a different concept but the answer is dependent on previous answer(s) of related MC question(s).

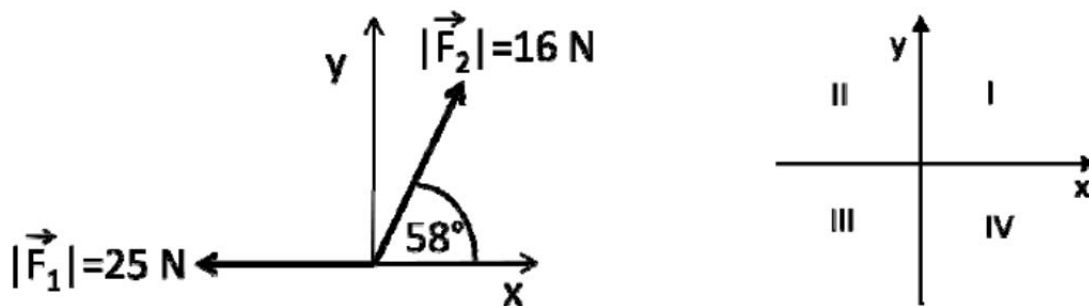
As discussed above, MC distractors can give insight to student misconceptions in the same way that the traditional CR problems can provide the instructor information about common misunderstandings. The MC question approach also allows for very fast turn-around in terms of

grading, and can be used to give feedback to the students about potential pitfalls. If an MC exam were to be given using electronic means, e.g., clickers, the instructor could immediately address common misconceptions directly after the exam—including in-class quizzes or online quizzes/exams—has been taken, whereas CR problems typically have a longer turnaround time for student feedback.

As long as students in an introductory engineering course can practice critical thinking and the engineering solving method in homework assignments and in-class activities, ideally combined with hands-on experience, the use of MC questions on the exam will not have an adverse effect on the students' overall performance.

Appendix

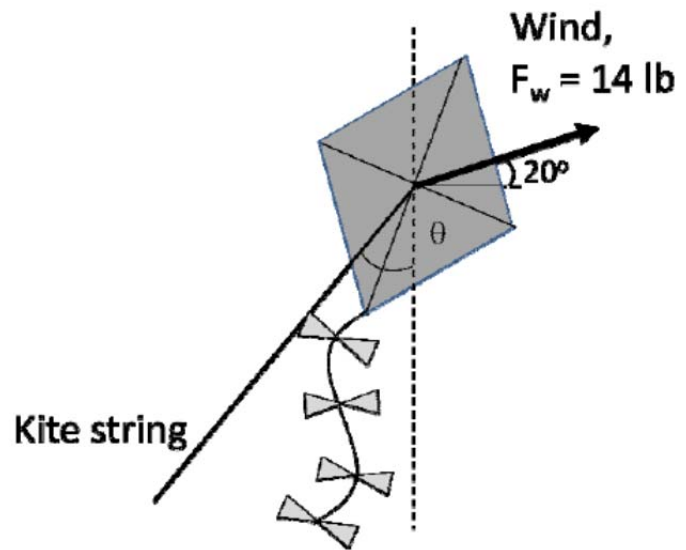
Consider the concurrent force system shown below.



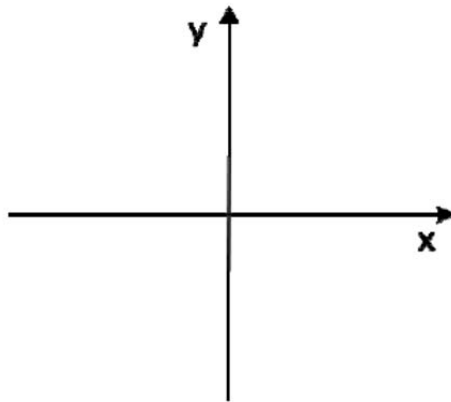
- Calculate R_x , R_y and resultant force, in N.
- Determine the quadrant of resultant force.

Figure A.1 CR version of concurrent force system question.

A kite experiences a wind force of 14 lb at an angle 20° from the horizontal direction. The kite weighs 2 lb and is held in place by a tensile force in the kite string F_s measured at an angle θ from the vertical direction. Assume a condition of static equilibrium. Assume that this is a 2D problem, and that the weight of the kite is concentrated at its center.



a. Draw the appropriate free body diagram for this problem on the axes provided

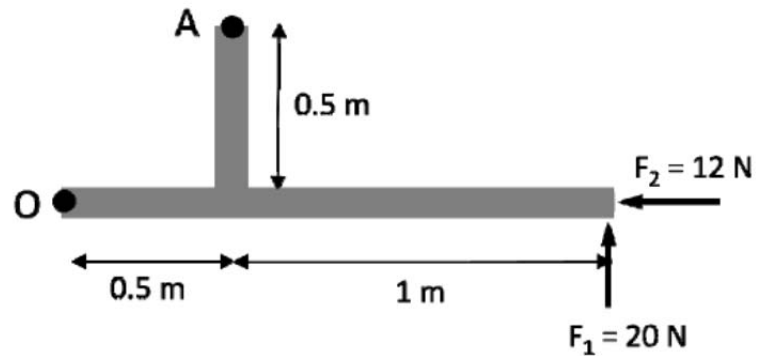


b. Find the magnitude of the force in the string F_s

c. Find the direction of the kite string θ

Figure A.2 CR version of forces on a kite question.

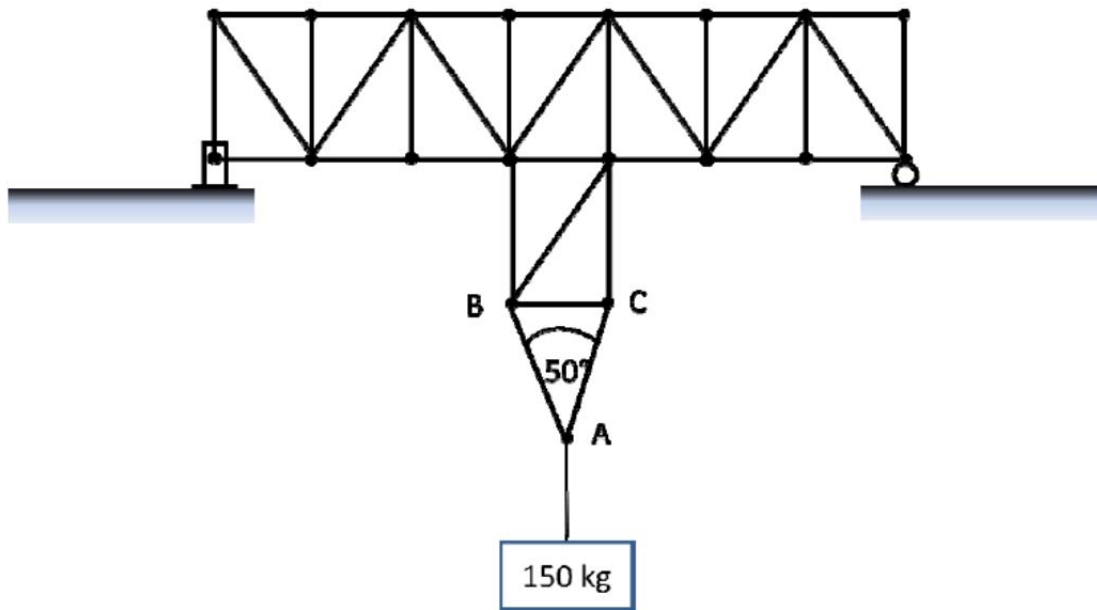
Consider a body subjected to forces F_1 and F_2 as shown.



- Determine the sum of the moments about point O.
- Determine the sum of the moments about point A.

Figure A.3 CR version of the beam question.

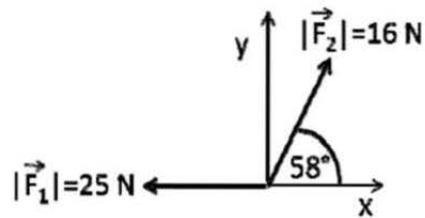
Consider the truss shown below. A mass of 150 kg is suspended by a cable at joint A.



- Determine the force in member AB, in N. Clearly state whether the member is in tension, compression or a zero force member.
- Determine the necessary minimum diameter of the cable to support the weight assuming the cable material is stainless steel having a yield strength of 210 MPa.

Figure A.4 CR version of the bridge truss question.

The following questions, 21.-24., refer to the concurrent force system shown below.



21. What is the x-component of the resultant force, R_x ?

- a. -38.6 N
- b. -33.5 N
- c. -16.5 N
- d. -11.4 N

22. What is the y-component of the resultant force, R_y ?

- a. -13.6 N
- b. -11.4 N
- c. 8.5 N
- d. 13.6 N

23. What is the resultant, $|R|$?

- a. 8.7 N
- b. 14.2 N
- c. 21.4 N
- d. 202.6 N

24. The resultant vector is in which quadrant?

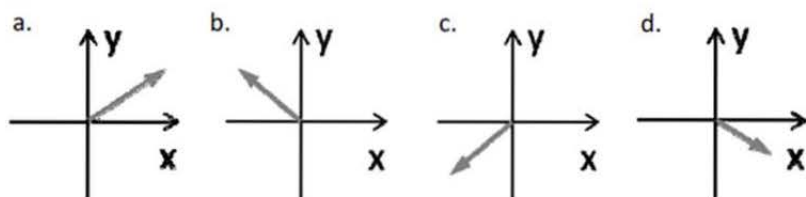
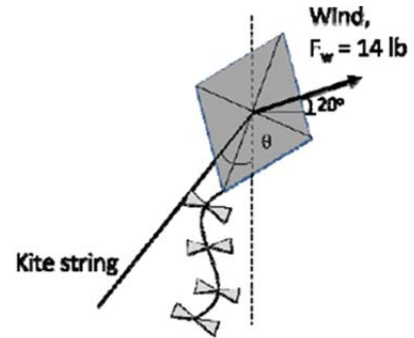
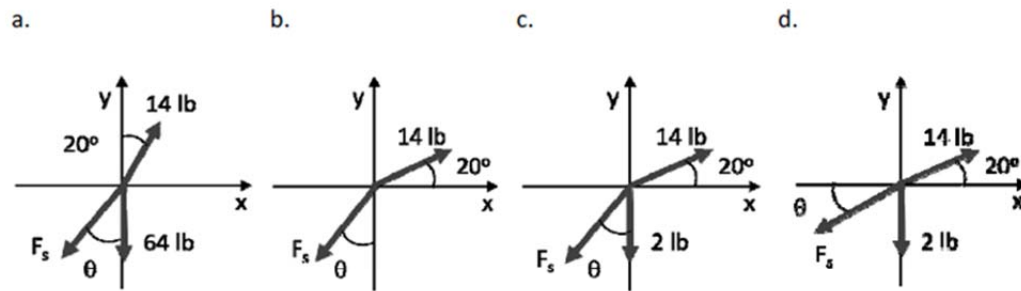


Figure A.5 MC version of concurrent force system.

The following questions, 25.-29., refer to the diagram at the right. A kite experiences a wind force of 14 lb at an angle 20° from the horizontal direction. The kite weighs 2 lb and is held in place by a tensile force in the kite string F_s measured at an angle θ from the vertical direction. Assume a condition of static equilibrium.



25. Circle the correct free-body diagram for this problem.



26. What is the magnitude of the x-component of the force F_s ?

- a. 13.2 lbf
- b. 4.8 lbf
- c. 14.0 lbf
- d. 7.4 lbf

28. What is the magnitude of the force F_s ?

- a. 14.0 lbf
- b. 12.2 lbf
- c. 5.2 lbf
- d. 13.5 lbf

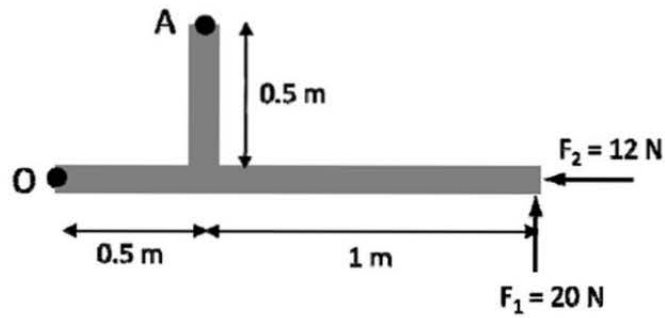
27. What is the magnitude of the y-component of the force F_s ?

- a. 4.8 lbf
- b. 11.2 lbf
- c. 2.8 lbf
- d. 2.0 lbf

29. What is the direction θ of the kite string?

- a. 12.0°
- b. 78°
- c. 67°
- d. 81°

Figure A.6 MC version of forces on a kite problem.



21. What is the moment exerted about point O by F_1 ?

- a. -30 Nm
- b. 30 Nm
- c. -10 Nm
- d. 0 Nm

22. What is the moment exerted about point O by F_2 ?

- a. -18 Nm
- b. 18 Nm
- c. 6 Nm
- d. 0 Nm

23. What is the moment exerted about point A by F_1 ?

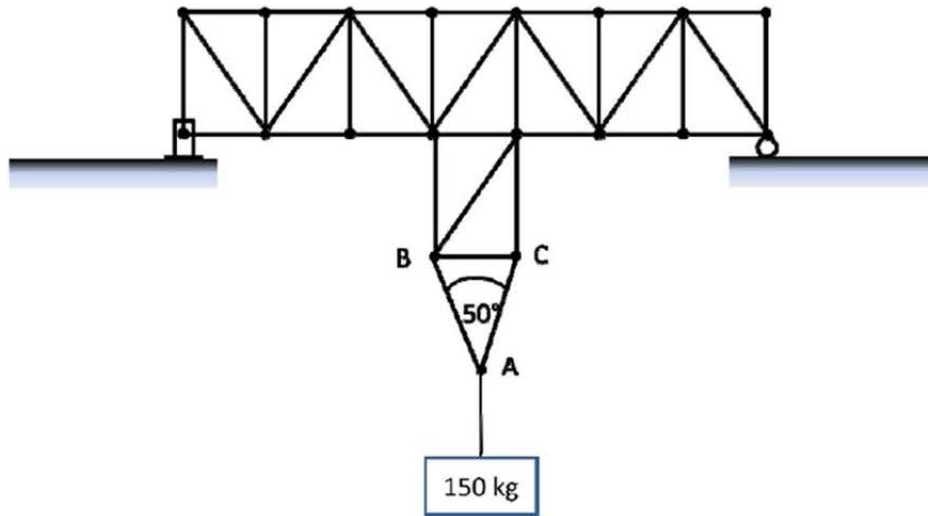
- a. 10 N-m
- b. 20 N-m
- c. -20 N-m
- d. 0 N-m

24. What is the moment exerted about point A by F_2 ?

- a. 12 Nm
- b. -12 Nm
- c. -6 Nm
- d. 0 Nm

Figure A.7 MC version of the beam problem.

The following questions, 25.-27., refer to the truss shown below. A mass of 150 kg is suspended by a cable at joint A.



25. The force in member AB, F_{AB} , is?

- a. -1741 N
- b. -178 N
- c. 0 N
- d. 812 N
- e. 1145 N

26. The member AB is in?

- a. Tension
- b. Compression
- c. Zero force member

27. What is the necessary minimum diameter of the cable to support the weight assuming the cable material is stainless steel having a yield strength of 210 MPa?

- a. 0.0010 m
- b. 0.0015 m
- c. 0.003 m
- d. 3.0 m

Figure A.8 MC version of the truss question.

Bibliography

1. Mehta, S.I., and Schlecht, N.W., "Computerized Assessment Technique for Large Classes," *Journal of Engineering Education*, 167-172, April, 1998.
2. Book, N.L., and Sitton, O.C., "Evaluation of Computer-Based Methods for Engineering Courses," *Proceedings of the 2005 American Society of Engineering Education (ASEE) Annual Conference & Exposition*, Portland, OR, June 12-15, 2005.
3. Marks, B.P., "Web-Based Readiness Assessment Quizzes," *Journal of Engineering Education*, 97-102, January 2002.
4. Walsted W. and Becker W., "Achievement Differences on Multiple-Choice and Essay Tests in Economics," *American Economic Review*, v. 84, n. 2, 193-196, 1994.
5. Wainer H. and Thissen D., "Combining Multiple-Choice and Constructed-Response Test Scores: Toward a Marxist Theory of Test Construction," *Applied Measurement in Education*, v. 6, n. 2, 103-118, 1993.
6. Sax G. and Collett L.S., "An Empirical Comparison of the Effects of Recall and Multiple-Choice Tests on Student Achievement," *Journal of Educational Measurement*, v. 5, n. 2, 169-173, 1968.
7. Bennet R.E., Rock D.A. and Wang M., "Equivalence of Free-Response and Multiple-Choice Items," *National Council on Measurement in Education*, v. 28, n. 1, 77-92, 1991.
8. Kranthwohl, D. R., "A Revision of Bloom's Taxonomy: An Overview," *Theory into Practice*, Vol. 41 Num. 4, 2002
9. Steif, P. S., Dantzler, J. A., "A Static Concept Inventory: Development and Psychometric Analysis", *Journal of Engineering Education*, 33: 363-371, 2005
10. Collins, J., "Writing Multiple-Choice Questions for Continuing Medical Education Activities and Self-Assessment Modules. *Radiographics*, 26:543-551, 2006