A Rubric for Assessing Thinking Skills in Free-Response Exam Problems

Fatema Al Salmani and Beth Thacker

Texas Tech University

Abstract

We designed a rubric to assess free-response exam problems in order to compare thinking skills evidenced in exams in classes taught by different pedagogies. The rubric was designed based on Bloom's taxonomy (revised version) [1-3]. The rubric was then used to code exam problems. In particular, we analyzed exams from different sections of the algebra-based physics course taught by the same instructor with different pedagogies, often the same semester . We compared sections taught traditionally, with and without interactive-engagement, and those taught by inquiry. We discuss the instrument, present results and present plans for future research. The inquiry-based students demonstrated all of the thinking skills coded more often than the traditional students.

The Rubric

Bloom's taxonomy levels for physics free-response exam problem solutions:

1) None

i) Nothing
ii) Totally incorrect
iii) Statement only
iv) Partial Remember – includes some facts or concepts needed to solve a problem, but at least one necessary concept or fact is missing or incorrectly remembered.

2) Remember: Recall facts and basic concepts

Solution includes recollection of a fact or concept or set of facts or concepts needed to solve the problem in either mathematical or written form. (If one or more necessary concepts are missing, it counts as Partial Remember and not Remember.)

3) Understand: Explain ideas or concepts

Solution includes a correct application of the facts or concepts recalled. This could be verbal or mathematical; it includes recognition of any details or constraints given in the particular problem and how they would be applied. A picture must be supported by a verbal explanation.

4) Apply/Analyze: Use information in new situations/draw connections among ideas

Solution demonstrates a valid, logical and consistent process for determining an answer, applying (1) and (2), and demonstrates evidence of understanding an underlying model or fundamental principle.

5) Evaluate: Justify a stand or decision

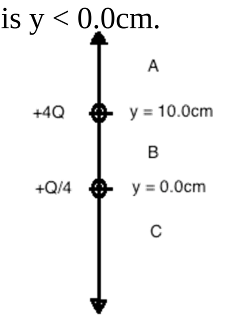
Solution includes a comparison of or discrimination between different possibilities or an explanation of why an answer is consistent or makes sense or why it is incorrect.

Introduction

In recent years there has been a move toward recognition of the need to assess not just conceptual understanding, but also students' thinking skills. At Texas Tech University (TTU) we have observed the need to promote and assess thinking skills. We have course sections taught by the same instructor using different pedagogies, one taught traditionally, and one taught non-traditional. We are interested in assessing both thinking skills and conceptual understanding in the different sections. The non-traditionally taught students were enrolled in an inquiry-based, laboratory-based physics course (INQ) taught without a lecture and without a text. Students worked through the directed-inquiry materials developed for the course [4], doing experiments to explore the world around them and developing qualitative and quantitative models based on their experimentation. We set out to design a rubric to assess thinking skills in the context of exam problem solving. Our goal was to design a simple, coarse rubric that was not hard to use but would give a general idea of the level of thinking skills applied in exam problem solving. Here we report on the creation of this rubric.

Example Problem 1

A positive charge +4Q is placed at y = 10.0cm and a charge of +Q/4 is placed at the origin of the y-axis, as in the picture below. Three regions of the y-axis are labeled. Region A is y > 10.0cm. Region B is 0.0cm < y < 10.0cm. Region C is y < 0.0cm.

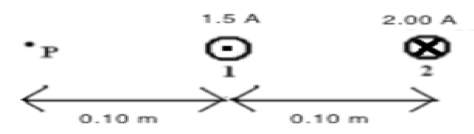


a) Is there any position along the y-axis that the net electric field could be zero? If so, in which region(s) could the net electric field be zero. Explain your reasoning.
b) If there is a position along the y-axis at which the net electric field is zero, determine that location. If there is not a position where the net electric field could be zero, explain why not. Show your work and explain your reasoning.

*Based on a problem from Arnold Arons. (1994). *Homework and Test Questions for Introductory Physics Teaching*. (John Wiley and Sons, NY).

Example Problem 2

Consider two parallel, current-carrying wire, as in the figure. The current in wire 1 is 1.5A out of the page. The current in wire 2 is 2.00 into the page. The wires are 0



a) Determine the magnitude and direction of the net magnetic field at point P. Draw and label the net magnetic field vector in the diagram above. Show your work.

a) If a third wire, 1.0m long, were placed at point P and a net force of 2.0 x 10-6 N to the left was observed, what would be the magnitude and direction of the current in the third wire? Draw and label relevant vectors in the diagram. Show your work and explain your reasoning.

Results

The example problems were analyzed by two raters. We originally treated the rubric as non-hierarchical, assigning a one or a zero to each level. However, we observed that for our data, the results were hierarchical. No one received a one for a level without receiving ones for all the levels below it. We then assigned a score by adding all the ones to represent a students' thinking level. The larger the number, the higher the student's thinking level.

We used the weighted Cohen's Kappa to measure inter-rater reliability for two raters. The results for weighted Cohen's Kappa with two raters for the INQ class are 0.82 and 0.85 on the (a) and (b) parts of the example problem 1, respectively. For the traditional class, the weighted Cohen's Kappa is 0.86 and 0.97 for the same two parts, (a) and (b), of the same problem. For example problem 2 The results for the INQ class are 0.81 and 0.86 on the (a) and (b) parts, respectively. For hierarchical data, the weighted Cohen's Kappa is the correct statistic to use and a value in the range 0.81-1.00 is considered almost perfect agreement.

We show the results of the comparison of students taught by two different pedagogies, students in the INQ class and students in the traditional class, in Figs. 1, 2, 3 and 4. Fig. 1 shows the results of part (a) and Fig. 2 shows the results of part (b) of the example problem 1. Fig. 3 shows the results of part (a) and Fig. 4 shows the results of part (b) of the example problem 2.

A much higher percentage of the INQ students demonstrated each of the thinking skills assessed by our rubric. This is consistent with previous data using a problem specific rubric. We have evidence of much more frequent use of all the thinking skills by the INQ students.

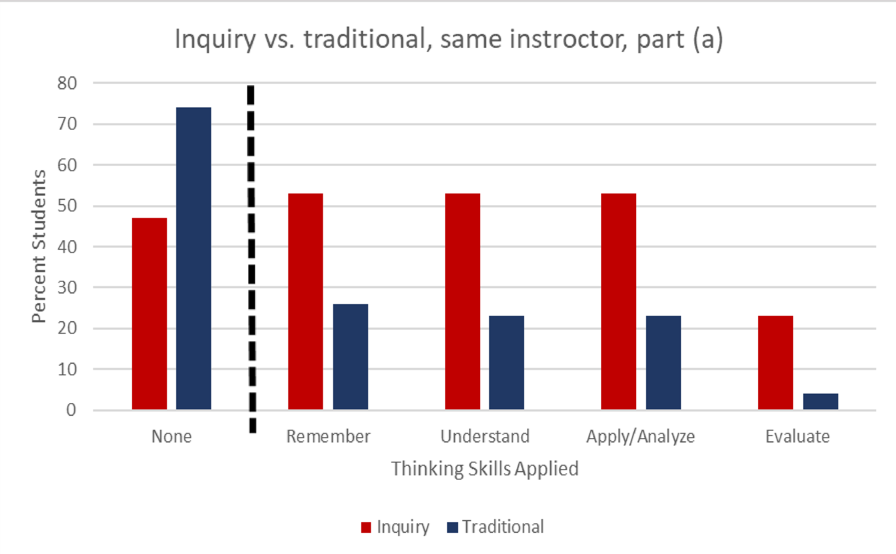


Fig.1. The results of the general rubric applied to two algebra-based physics sections, one INQ, one more traditional, part (a).Same instructor, same semester, different pedagogies. The dotted line separates the None category from the Thinking Skills categories.

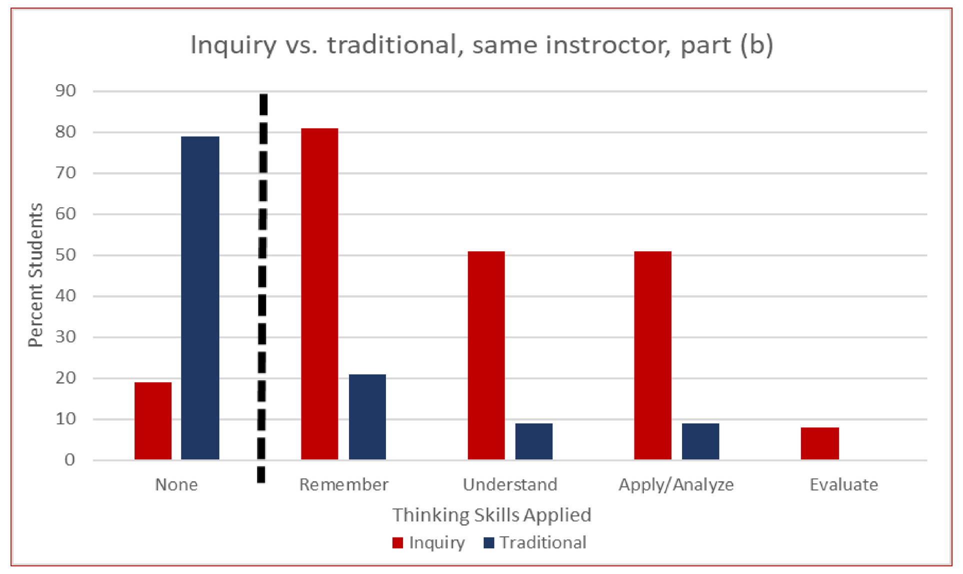


Fig.2. The results of the general rubric applied to two algebra-based physics sections, one INQ, one more traditional, part (b).Same instructor, same semester, different pedagogies. The dotted line separates the None category from the Thinking Skills categories.

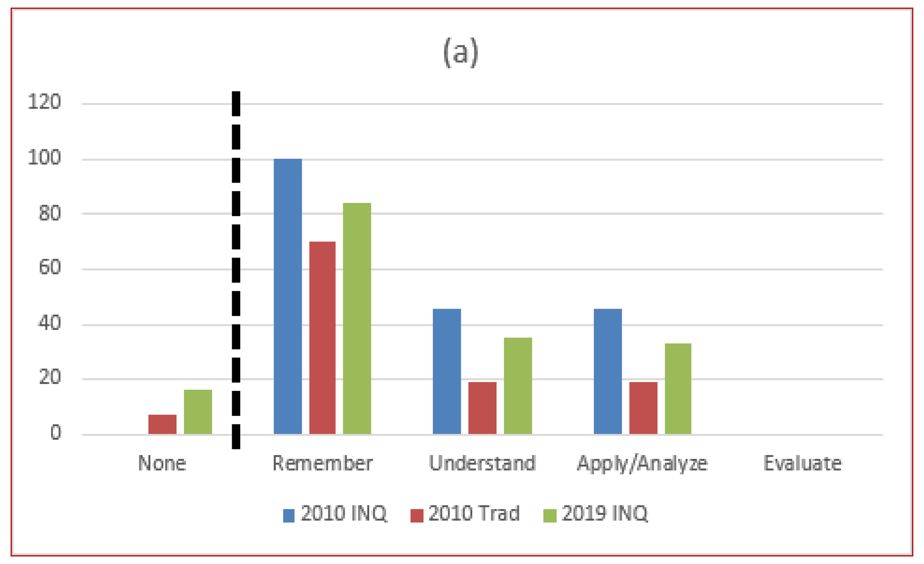


Fig.3. The results of the general rubric applied to three algebra-based physics sections, one small class INQ (24 students), one interactive-engagement traditional (64 students) , and one large class INQ (64 students) part (a). All were the same instructor. The 2010 sections were the same instructor, same semester, different pedagogies. The dotted line separates the None category from the Thinking Skills categories.

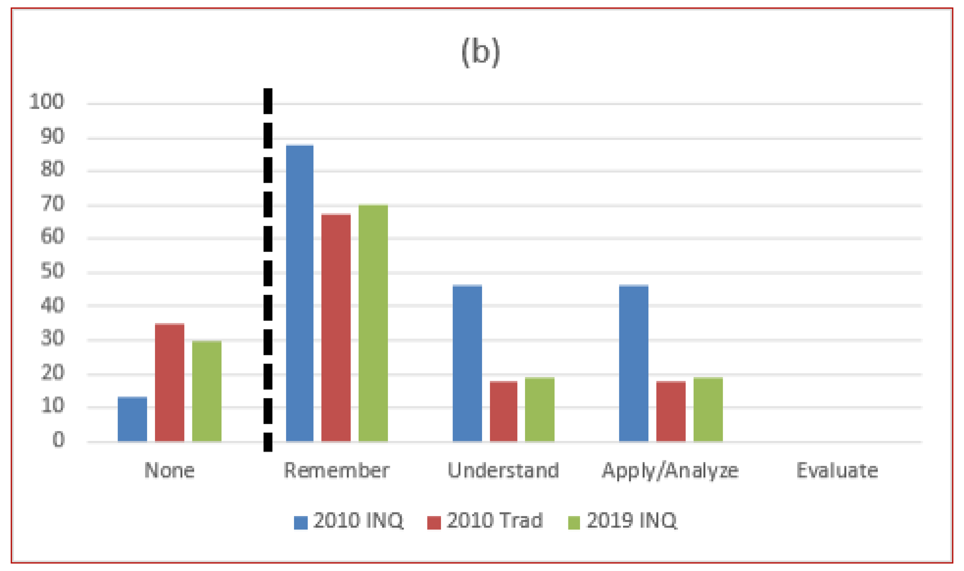


Fig.4. The results of the general rubric applied to three algebra-based physics sections, one small class INQ (24 students), one interactive-engagement traditional (64 students) , and one large class INQ (64 students) part (b). All were the same instructor. The 2010 sections were the same instructor, same semester, different pedagogies. The dotted line separates the None category from the Thinking Skills categories.

Methods

- **Identification of common exam problems :** We first identified exam problems that had been administered and were common across classes. We also identified, in particular, sections of the algebra-based course taught by the same instructor using different pedagogies.
- **The rubric and analysis :** Our goal was to design a general rubric to analyze free-response physics problems. In the development of the general rubric we started with a rubric with seven levels but soon moved to a five- level rubric. The rubric was designed so that each student answer would be coded as to whether or not the written answer supplied evidence of the thinking skill described at a particular level of the rubric. The student answer was given a 1 or a 0 for that level on each part of the problem (for multipart problems). After we were satisfied with the rubric wording, we used it to analyze different problems from traditional and INQ classes. We assessed inter-rater reliability for two raters by calculating percent agreement and Cohen's Kappa. We also interviewed students to establish construct validity. We report on our present results and present the results for two of the questions we coded for exam problems comparing students taught with different pedagogies by the same professor.

Conclusion

We have developed a general coarse rubric that is reliable to compare students thinking skills across pedagogies in the context of FR exam problems.

We have evidence, based in our rubric, that students in an INQ section of algebra-based physics demonstrate a more frequent use of all the thinking skills analyzed in our example problems compared to a traditional section. This is consistent with earlier results with a problem specific rubric and indicates that the combination of the instructional method (Socratic questioning) and the evidence-based instructional materials [4] are more effective at promoting the thinking skills assessed than traditional instructional methods.

We are presently continuing this research to tweak our rubric, apply it to other problems and to problems from other universities.

References

1. Bloom, B.S. (1956). Taxonomy of educational objectives: The classification of educational goals. New York, NY: Longmans, Green.
2. Anderson, L.W. & Krathwohl, D.R. (2001). A taxonomy for teaching, learning, and assessing: A revision of Bloom's taxonomy of educational objectives. New York, NY: Longman.
3. L.W. Anderson, D.R. Krathwohl, P.W. Airasian, K.A. Cruikshank, R.E. Mayer, P.R. Pintrich, J. Raths, and M.C. Wittrock, (Longman, New York, 2001).
4. B. Thacker and K. West, Inquiry-based Experimental Physics Volumes I and II, (Stipes Publishing L.L.C., Illinois, 2019).