

# Tapping into Juniors' Understanding of E&M: The Colorado Upper-Division Electrostatics (CUE) Diagnostic

Stephanie V. Chasteen<sup>\*†</sup> and Steven J. Pollock<sup>†</sup>

<sup>\*</sup>*Science Education Initiative, University of Colorado, Boulder, CO 80309, USA*

<sup>†</sup>*Department of Physics, University of Colorado, Boulder, CO 80309, USA*

**Abstract.** As part of an effort to systematically improve our junior-level E&M I course, we are developing a tool to assess student learning of E&M concepts at the upper-division. Along with a faculty working group, we established a list of learning goals for the course that, with student observations and interviews, served as a guide in creating the Colorado Upper-Division Electrostatics (CUE) assessment. The result is a 17-question open-ended post-test (with an optional 7-question pre-test) diagnostic, and accompanying grading rubric. We present the preliminary validation of the instrument and rubric, plus results from 226 students in 4 semesters at the University of Colorado, and 4 additional universities.

**Keywords:** physics education research, course reform, electricity and magnetism, assessment

**PACS:** 01.30.Ib, 01.40.Di, 01.40.Fk, 01.40.G-, 01.40.gb

## INTRODUCTION

The PER community has investigated student understanding of introductory physics topics, including electromagnetism (E&M), in some detail [1,2,3]. Efforts to improve student learning in lower-division courses has been driven in large part by data on student performance from research-based conceptual tests, such as the Force Concept Inventory (FCI)<sup>4</sup>, Conceptual Survey of E&M (CSEM) [5] or the Basic Electricity and Magnetism Survey (BEMA) [6]. Not only have these instruments served to identify common and persistent student difficulties, but they have also been a powerful tool for supporting curricular reform. Faculty may be convinced about the persistence of certain learning difficulties when students perform poorly on areas of the diagnostic that they explicitly covered in lecture.

We lack similar tools for improving instruction at the junior level. Research on how students understand the more mathematically and conceptually sophisticated treatment of the material at the upper-division is just beginning [7,8]. We have undertaken a multi-year transformation of our upper-division E&M course, including identification of student difficulties, learning goals for the course, and student-centered instruction [8]. To assess the relative success of these transformations, as well as to document student difficulties, we developed a post-test for the course.

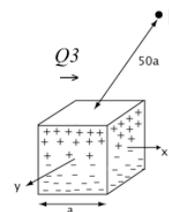
## THE CUE

The CUE is a 17-question conceptual test (15 electrostatics and 2 magnetostatics questions) testing students' ability to choose a problem-solving method and defend that choice, sketch electric field patterns, graph electric field strength and potentials, and explain the physics and mathematics underlying steps in common problems. A pre-test was developed from a subset of the questions on the CUE. With the exception of one multiple-choice question the exam is open-ended; 3 additional questions give students a multiple-choice alternative and require students to explain their answer to receive credit. The CUE is designed to be given in a single 50-minute lecture. Fig 1 shows two sample questions from the CUE [9].

## CUE Development

We began the process of course transformation by consulting PER and non-PER faculty with an interest in the course. In these individual interviews and informal biweekly brownbag meetings, we began to tackle the question of what students need to be able to do by the end of the course. The learning goals developed in these meetings -- such as "choose and apply the appropriate problem-solving technique" and "sketch the physical parameters of a problem." [10] -- are intended to describe the skills expected of a junior-

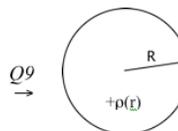
**Q3.** Give a brief outline of the EASIEST method that you would use to solve the problem. Do not solve the problem, we just want to know the general strategy and why you chose that method. A solid, neutral non-conducting cube, centered on the origin, with side length “a” and charge density  $\rho(z) = kz$ , Find E (or V) outside, at point P, off-axis, at a distance  $r=50a$ .



**Q9.** You are given a non-conducting sphere, centered at the origin. The sphere has a non-uniform, positive and finite volume charge density  $\rho(r)$ . You notice that another student has set the reference point for V such that  $V=0$  at the center of the sphere:  $V(r=0)=0$ .

What would  $V=0$  at  $r=0$  imply about the sign of the potential at  $r \rightarrow \infty$ ?

- (a)  $V(r \rightarrow \infty)$  is positive (+)
- (b)  $V(r \rightarrow \infty)$  is negative (-)
- (c)  $V(r \rightarrow \infty)$  is zero
- (d) It depends



**Briefly explain your reasoning:**

**FIGURE 1.** Two problems from the CUE, Q3 (5 points) and Q9 (10 points). Total exam is worth 118 points.

level physics student. While some of these skills are tested in traditional exams (such as setting up and executing line and surface integrals), many of these techniques and their conceptual underpinnings are not. The CUE was developed to address this gap.

Questions were developed based on common difficulties observed in student interviews and homework help sessions, and in collaboration with the faculty working group, considering the learning goals. Questions were then validated in think-aloud interviews as students worked through the diagnostic, as we reviewed their performance on the diagnostic, and through evaluating the quality of responses on the test (i.e., did they understand what we were asking?) over 3 semesters of administration. Since the first version, a total of 7 questions were dropped, 2 questions were added, and 5 questions were substantially modified, with most questions reworded for clarity. We are in the process of further evaluating the diagnostic through solicitation of expert responses.

## CUE Grading Rubric

A detailed grading rubric was developed in order to explicitly define what points should be assigned to each question. This was a challenging task; the rubric needed to clearly specify the points to be assigned for a variety of student responses. For example, Q3 (5 points: Figure 1): 3 points for the correct answer for Q3 with partial points specified for common incomplete answers (e.g., Coulomb’s Law, or “multipole” without mentioning dipole nature), and 2 points for a satisfactory *explanation* with partial points specified for common incomplete answers.

Not every question is that straightforward. Some sample responses for Q9 (10 points: Figure 1) are:

(answer: b) *V is a negative quantity, since it is 0 at the center, it must be negative far away.*

That student gained 3 points for correctly indicating that V decreases as you move away from the sphere, but lacking a correct definition of the potential

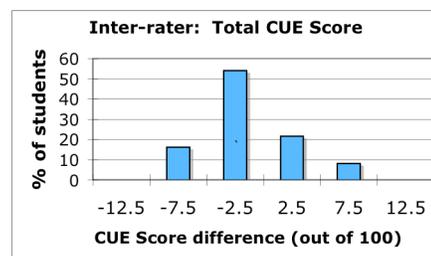
difference and logical incompleteness. For comparison, a 10-point answer read

(answer: b) *If  $V(r \rightarrow \infty) = 0$ , then  $V(r=0)$  is positive, so  $V(r \rightarrow \infty)$  should be negative if we set  $V(r=0) = 0$ .*

while a 0-point answer read

(answer: c) *Potential should always go to zero as the distance goes to infinity.*

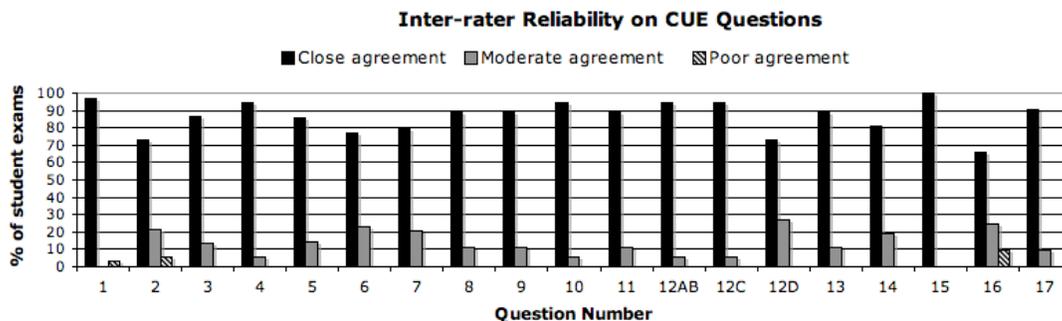
What is important is not that all parties agree with our point scheme, but rather that independent graders will achieve consistent results using the rubric. The rubric underwent significant revision over time as part of the validation process, resulting in an 11-page document (unpublished). To check inter-rater reliability on the final rubric, two independent graders graded a total of 36 exams from another institution. Graders were trained on the rubric through grading and subsequent detailed discussion of point assignments on a subset of exams. Grader differences on the “Total CUE Score” are shown in Figure 2.



**FIGURE 2.** Grader differences in “Total CUE Score”. Blank responses were included as part of this score.

Inter-rater reliability for the CUE exam as a whole is very high. Averaged across all exams, the difference in the total CUE score between graders was -1.4 points (out of 100) with the standard deviation (of the difference) of 3.7. As can be seen in Figure 2, graders agreed within 10% on CUE scores for all students and within 5% on the majority of students (76%).

We also examined inter-rater reliability on the individual questions on the exam. All questions are normalized to 100 points for this analysis. One long



**FIGURE 3.** Inter-rater differences on CUE questions. “Close” agreement is agreement within  $\pm 20\%$  ( $\pm 1$  point on a 5 point question), “moderate” is within  $\pm 20\text{-}50\%$  ( $\pm 1\text{-}2.5$  points on a 5 point question) and “poor” is off by 50-100%. Blanks omitted.

question, worth 26 points, is split into 3 parts, for a total of 19 questions. The average inter-rater difference is low, but variable, for individual questions, as indicated by a standard deviation ranging from 0 to 28 (on a scale of 100), with an average of 12.3 over all questions. Average grader differences are on the order of 0.1 to 0.5 points (on a scale of 100) per problem. As seen in Figure 3, graders were in “close” agreement for at least 75% of students on all questions but two. The two graders were in *exact* agreement for 45% or more of the students on all questions but one.

Thus, on average, we are able to discern the difference between total exam scores within 5%, and on individual questions within 20%. These results indicate that it is meaningful to compare student performance on both individual questions and the exam as a whole, though some caution is required on a few questions with lower reliability (e.g., Q16). In conclusion, the agreement level is high, particularly considering the variability of student responses, and difficulty in interpreting many answers.

## RESULTS

The CUE was administered to upper-division E&M students for 4 semesters at the University of Colorado, and in 5 courses at 4 outside institutions, for a total N=226 students. The exam was administered at the end of the semester, in-class as part of the course, resulting in high response rates (75-100%), except “Trad” (63% response due to poor attendance). One course (C1) gave the CUE as a take-home test, which may explain the high variance for that course. Some instructors announced the CUE in advance and some students received participation credit. Because the CUE was modified from semester to semester, a total of three different versions were given across institutions. Thus, there are two scores of interest, the “Comparison CUE Score” (88 points), made of questions given in common across all exams and a “Total CUE Score” (118 points) which represents the

exam in its final version. Institutions are described in Table 1. Courses IE1-IE3 represent 3 semesters of a transformed course at CU with similar research-based IE techniques and materials, also adopted by C-IE [8].

**TABLE 1.** Institutions administering CUE.

*An \* indicates transformed courses using research-based interactive engagement (IE) materials as described in [8]*

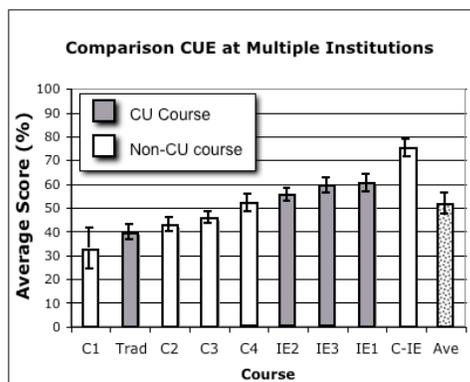
Code	Institution	Pedagogy <sup>†</sup>	N	Total CUE
Trad	CU	L	26	N/A
IE1*	CU	CQ-PI, Tut,	21, 48,	N/A,
IE3*	(3 consecutive semesters)	GPS, L	27	53%, 60%
C1	Public Univ.	GPS, Q, L	6	N/A
C2	Private Eng. College	GPS, L	18	40%
C3	Public Univ.	GPS, Q, L	52	45%
C4	Public Univ.	GPS, Q, L	39	50%
C-IE*	Private Liberal Arts	CQ-PI, Tut, GPS, L	12	73%

<sup>†</sup> Clickers with Peer Instruction (**CQ-PI**), Conceptual Tutorials (**Tut**), Group Problem Solving sessions, in or out of class (**GPS**), Frequent Quizzes (**Q**), Lecture (**L**).

The “Comparison CUE Score” was moderately correlated with student course grades ( $r=0.49$ ,  $p \ll 0.01$ ) at CU, and individual questions were well correlated with the exam score as a whole (Cronbach's  $\alpha$  for numbered items is 0.82, indicating strong internal statistical reliability). Comparison CUE scores are given in Fig. 4, and CUE-V16 scores in Table 1. Institutional scores on the Comparison CUE range from 33%-74%, with an average of 51% across all courses (N=9) and across all students (N=226). On 11 of the 17 questions, average student performance was under 50%. This is a challenging test.

It is difficult to make direct comparisons between different institutions due to variations in curricula, administration of the CUE, and student body demographics. However, we note that all courses above the mean used research-based interactive engagement (IE) techniques, as described in [8].

Additionally, students in courses using IE techniques such as clickers (all CU-IE courses and C-IE) scored higher, on average, on the Comparison CUE, and those students outperformed students in other courses on all questions. This suggests that IE improves conceptual understanding at the upper division. This conclusion is strengthened by the fact that 3 of the instructors (IE1, IE2 and C-IE) had never taught E&M previously, and C-IE5 was a new faculty member (but C1, C2 and Trad had taught the course several times).



**FIGURE 4.** Courses C1-4 and C-IE are non-CU. “Trad” is a traditional lecture-taught course, and IE1-3 used similar research-based course transformations.

We were particularly interested in the CUE performance of 4 semesters at CU – one traditionally taught lecture and three taught using our research-based transformations. The CUE serves as one of our main measures of the success of these transformations. The Comparison CUE scores of the transformed courses (cross-course average 62%) are statistically significantly higher than the traditionally taught course (42%), and all are higher than the average across all courses (statistically significant except for IE2). As to computational ability, we previously observed that students in IE1 outperformed students in Trad on standard exam problems [8]. Overall, these results suggest that the interactive techniques were consistently successful, over 3 semesters, in improving students’ conceptual and computational skills.

A pre-test was developed using 7 of the 17 questions which an incoming junior might be expected to know, and given in some courses (Table 2). Pre-test scores are relatively consistent at about 30%, either before Junior E&M or just after Freshman E&M, suggesting a remarkably low “fade” in the intervening 2 years [11]. Pre-test scores were compared to the same 7-questions on the post-test (“7-Q Post test”). Students in C-IE begin at a higher level (as would be expected of students in an elite private college), but experience similar gains (28%) to other courses using the transformed curriculum.

**TABLE 2.** Pre/post test scores (matched by student). “CU Freshmen” is N=25 after freshman E&M. Both tests are a 7 question subset of the V16 CUE (60 points out of 118).

Code	Pre-test (%)	7-Q Post-Test (%)	Gain (Post-Pre)
CU Freshmen	30 ± 3.0	N/A	N/A
IE2	30 ± 2.3	51 ± 2.9	21 ± 2.8
IE3	33 ± 3.2	61 ± 3.4	28 ± 3.0
C-IE	43 ± 6.3	71 ± 5.9	29 ± 7.6
C1	33 ± 5.3	47 ± 12.5	15 ± 9.3

## CONCLUSIONS

We have developed a validated instrument for use in studying and assessing student conceptual understanding of upper-division E&M. The grading rubric shows good interrater reliability, and the diagnostic is able to differentiate between different student populations and instruction methods. This assessment can serve as a tool for the physics community to support ongoing curriculum development at the upper-division.

## ACKNOWLEDGMENTS

We acknowledge the generous contributions of many faculty members at CU, for assistance with creating and/or administering the CUE. Many thanks to A. Bloom for grading the CUE and assisting with the rubric. We are grateful to the 4 instructors at external institutions who provided us CUE scores from their courses. This work is funded by the Science Education Initiative and NSF-CCLI grant #0737118.

## REFERENCES

1. C. Singh, *Am. J. Phys.* **74**, 923 (2006).
2. L. McDermott and E. Redish, *Am. J. Phys.* **67**, 755, (1999) and references within
3. R. Chabay and B. Sherwood, *Am. J. Phys.* **74**, 329 (2006).
4. R.R. Hake, *Am. J. Phys.* **66**, 750-74 (1998)
5. D. Maloney et al, *Am. J. Phys.* **69** (2001) S12.
6. L. Ding et al., *Phys Rev ST:PER* **2**, 010105 (2006)
7. For example: C. A. Manogue, et. al, *Am. J. Phys.* **74**, 344 (2006); J. Bilak and C. Singh, *PERC Proc.* **951**, AIP, Syracuse, NY, 2007, p. 49;; T. J. Bing and E. F. Redish, *PER Proceedings* **883**, AIP, Syracuse, NY, 2006, p. 26.; E. F. Redish, *World View on Physics Education in 2005: Focusing on Change Conf. Proc.*, Delhi, 2005.
8. S.V. Chasteen and S.J. Pollock *PERC Proc.* **1064**, AIP, Syracuse, NY, 2008, p 91-94.
9. To see the CUE and provide feedback, visit [www.colorado.edu/sei/surveys/Faculty/CUE/Sp09\\_CUE.html](http://www.colorado.edu/sei/surveys/Faculty/CUE/Sp09_CUE.html)
10. All course material (e.g., learning goals, clicker questions, CUE) are at [www.colorado.edu/sei/departments/](http://www.colorado.edu/sei/departments/)
11. See also S. Pollock and S.V. Chasteen, *PERC Proc.*, AIP, Syracuse, NY, 2009 (in press).