Reworking exam problems to incentivize improved performance in upper-division electrodynamics

Andrew J. Mason[1], John S. Colton[2]

[1]Department of Physics and Astronomy, University of Central Arkansas, Conway, AR 72035

[2]Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602

Goal of the Project

A previous study shows that incentivizing students to correct mistakes on unit exam problems within an upper-division quantum mechanics course improved students' problem-solving efforts on those same problems in a final exam environment, relative to a comparison group of students who were not incentivized. [3]

We replicate the quantitative portion of this study within a first-semester upper-division electromagnetism course, specifically examining students' invoking correct concepts and applying those concepts correctly.

Background

- Upper-division undergraduate physics students may still exhibit novice-like tendencies with problem solving attempts, do not reflect on mistakes [1-3]
- Cognitive apprenticeship [4-5]: reinforcement via scaffolding
- QM study [3]: Explicitly incentivizing students to rework unit exam problems assists students' final exam performance in quantum mechanics, indicated need for future research on specific QM topics, e.g. [6-7]
- Current study: Replicate QM study for upper-division electromagnetism
- Potential difference between QM and EM: QM introduces concepts at same time as new mathematical rigor [6-8], EM has more conceptual foundation? [9]

Questions:

1) **Does incentivized reworking of unit exam problems help students perform better on the same problems featured verbatim on the final exam semester in a first-semester upper-division electromagnetism course?**
2) **How does the intervention appear to help for different kinds of problem structures (invoking vs. applying)?**

Methodology

- ***Sample:*** 3 in-person sections (Fall 2019, Spring 2021*, Fall 2021) of EM1 for physics majors at BYU, same instructor (see Table I)
  - *Note: "Spring 2021" is summer term, 7 weeks long
- ***Instructional mode:*** Traditional lecture with specific active learning components (formative conceptual quizzes to begin each class, Socratic dialogue, individual or paired attempts at problem solving)
- ***Exams:*** Three unit exams and a cumulative final exam
  - Exam mode: Dedicated testing center outside class time; single page of handwritten notes permitted; students take exam individually within a period of a set number of days (3-5)
- ***Intervention:*** After unit exams graded and returned:
  - Opportunity to rework missed points on all problems as desired (none, some, all)
    - ***Groups:*** reworked problem vs. declined to rework for each individual problem (see Table I)
  - Students have five days to revise and resubmit
  - Selected problems repeated verbatim on final exam (students do not know this), used across two or all three semesters (see Fig. 1)
- ***Scoring:*** Rubrics devised according to model of QM study [3]
  - # of items differs per problem
  - All problems have two categories of items:
    - Invoke (invoking the correct concepts for the problem)
    - Apply (applying concepts correctly)
  - Inter-rater reliability for sampling of different problems (81.5% initial agreement, 90.8% conferred agreement)

FIG. 1. The three problems chosen for analysis in this study, as well as which unit exam (of three, not including final) each problem was featured in. Included in parentheses is the semester(s) for which each problem was used.

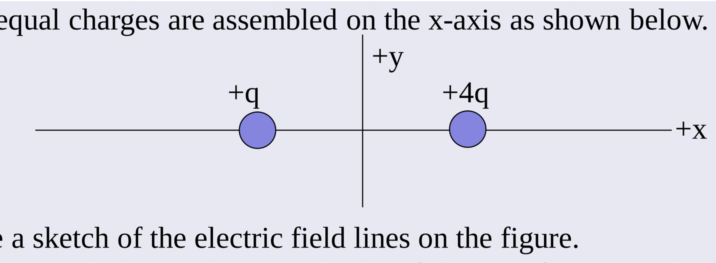
Name	Unit exam featured?	Text
P1	Exam 1 (F19, S21, F21)	Two point charges of equal but opposite charge are separated by a distance d, the +q charge being on the left and -q on the right. If the charges are each moved a distance d/2 away from each other, what is the change in potential energy of the system? Specify whether the potential energy has increased or decreased, and give a conceptual explanation for why this is the case.
P2	Exam 1 (S21, F21)	Two unequal charges are assembled on the x-axis as shown below.  a) Make a sketch of the electric field lines on the figure. b) In the space below make a rough plot of E_x vs. x for points along the x-axis. Don't worry about any numbers, just the general shape of what happens to the x-component of the field – which can be positive or negative depending on whether **E** points to the right or the left.
P3	Exam 2 (S21, F21)	A uniform charged rod with linear charge density +l lies on the z-axis, extending from z = 0 to z = – d as shown. Suppose you want to calculate the potential at point P which lies on the positive z-axis. a) What is the monopole contribution to the potential at point P, V_{mono}? b) In order to get a little more accuracy than the monopole potential, consider the dipole potential: $V \approx V_{mono} + V_{dip}$. (You may or may not have noticed, but the dipole formula does not actually require there to be both positive and negative charges, although typically we think of dipoles in those terms.) What is the dipole contribution to the potential at point P, V_{dip}? c) In order to get even a little more accuracy than the monopole and dipole potentials combined, consider the quadrupole potential: $V \approx V_{mono} + V_{dip} + V_{quad}$. What is the quadrupole contribution to the potential at point P, V_{quad}?

TABLE I. Summary of students who respectively were and were not eligible for reworking problems on each featured problem in the unit exams and final exams. See Fig. 1 for the text associated with each problem ID number.

Semester (total #)	Problem	Eligible? (< 90% on unit exam attempt)			Total eligible attempts (reworked + no rework)
		Yes, reworked	Yes, did not rework	No	
F19 (25)	P1	8	7	10	15
S21 (14)	P1	4	4	6	8
	P2	7	2	5	9
	P3	4	1	9	5
F21 (36)	P1	6	17	13	23
	P2	11	6	19	17
	P3	9	13	14	22

Note: Table I counts students' eligibility, choice to rework separately for each problem
- Counted as "how many students per problem"
- F19: Only had P1 in common with other two sections

TABLE II. Comparisons between groups of students who reworked each problem and students who declined to rework, among the eligible students across all three sampled semesters. Analysis interprets the unit exam attempt as a pretest and the final exam attempt as a posttest. Unit/final exam attempt values are average percentages correct.

	Problem	P1 (F19, S21, F21)		P2 (S21, F21)		P3 (S21, F21)	
	Group	Reworked	Declined	Reworked	Declined	Reworked	Declined
	N	18	28	18	8	13	14
Unit exam attempt	Invoke (%)	60.3	75.9	64.1	72.9	67.7	67.2
	Apply (%)	67.1	73.8	59.3	64.6	54.0	69.9
	Total (%)	63.7	74.8	61.7	68.8	60.9	68.6
Final exam attempt	Invoke (%)	83.8	77.7	87.2	74.5	85.3	84.3
	Apply (%)	89.5	76.7	90.7	74.0	90.8	76.2
	Total (%)	86.6	77.2	89.0	74.2	88.1	80.3
Gain <g>	Invoke	+0.513	+0.211	+0.638	+0.337	+0.509	+0.519
	Apply	+0.621	+0.226	+0.757	+0.412	+0.788	+0.310
	Total	+0.683	+0.276	+0.709	+0.379	+0.629	+0.482

Normalized gains: large effects showing benefit for reworking, seen in all problems:
- P1 ($p < .05$, Overall and Apply)
- P2 (but $p > .05$…)
- P3 ($p < .05$, Apply only)

Why ANCOVA here too?
- Unit exam choice to rework might be covariate
- Shows significance for P2 (overall and Invoke)

TABLE III. Analysis for statistical significance (in p-values) and effect size (in Cohen's d) between groups described in Table II. The comparisons given for the rows for unit exam attempt, final exam attempt, and normalized gains are according to 1-way ANOVA, while the rows for ANCOVA comparisons compare residuals between groups. Statistically significant results (defined as p < 0.05) are in bold. An asterisk indicates that variances were statistically unequal (p < 0.05 on F-tests), which the 1-way ANOVA comparison has taken into account. Note: F19 = Fall 2019; S21 = Spring 2021; F21 = Fall 2021

		P1 (F19, S21, F21)		P2 (S21, F21)		P3 (S21, F21)	
		p-values	Cohen's d	p-values	Cohen's d	p-values	Cohen's d
Unit exam attempt, 1-way ANOVA	Invoke	0.073*	0.621	0.22	0.535	0.95	0.025
	Apply	0.27	0.340	0.59	0.235	0.18	0.536
	Total	**0.026**	**0.698**	0.38	0.379	0.36	0.360
Final exam attempt, 1-way ANOVA	Invoke	0.40	0.258	0.13	0.660	0.90*	0.047
	Apply	**0.029**	**0.680**	0.065	0.820	0.12*	0.609
	Total	0.10	0.502	0.089	0.752	0.32*	0.380
Gain <g>, 1-way ANOVA	Invoke	0.07	0.559	0.14	0.643	0.96	0.021
	Apply	**0.011**	**0.799**	0.12	0.691	**0.014**	**1.017**
	Total	**0.0066**	**0.861**	0.11	0.713	0.39	0.339
ANOVA-based ANCOVA	Invoke	0.144		**0.029**		0.92	
	Apply	**0.016**		0.052		**0.022**	
	Total	**0.012**		**0.040**		0.13	

Does incentivized reworking of problems help students perform better on final?

- Normalized gains: Statistically significant overall for P1, not for P2, Apply only for P3
  - However, appear to be moderate-large effect size and/or statistical differences on unit exam score
- ANOVA-based ANCOVA (addressing unit exam score as covariate): Significant overall for P1 and P2, not overall for P3 (Apply only)
  - Covariate seems to matter for P2 (the conceptual problem) in particular

How does the intervention appear to help for different kinds of problem structures?

- P1: Apply, but not Invoke (more of an effect for algorithm-based part than conceptual-based part)
- P2: Invoke, but not Apply (primarily conceptual problem)
- P3: Apply, but not Invoke (primarily algorithmic problem)

Results

- ***Students benefit more from reworking problems for partial credit***
  - ANCOVA: effect not merely due to unit exam score (i.e. students aren't "primed" by choosing to rework)
- ***Different effects for different problem structures***
  - "Invoking" improved for primarily conceptual problem (P2)
  - "Applying" improved for primarily algorithmic problem (P3) and problem that had both conceptual and algorithmic parts (P1)

Future Work

- ***This poster is portion of more detailed approach*** [10]
  - Qualitative think-aloud protocol for interviews of selected students who reworked and declined (in progress): expert-novice differences? [1,8]
  - More detailed statistical analysis planned: explore degree of unit exam score vs. reworking problems
- ***Future plans: potential to look across upper-division courses*** [11]
  - Observable differences between how physics majors look at EM1 and QM1?
- ***Mason sabbatical, Fall 2022*** – look beyond exams

Acknowledgements

- Helpful conversations for methodological design and rubric checking:
  - Chandralekha Singh and Robert P. Devaty, Department of Physics and Astronomy, University of Pittsburgh
- Qualitative interpretation (future plans):
  - Jessica M. McCardell, Department of Curriculum and Instruction, University of Idaho Coeur d'Alene
- Statistical analysis checking, advice, interpretation:
  - Philip White, Department of Statistics, Brigham Young University

References

1. A. Mason and C. Singh, "Do advanced physics students learn from their mistakes without explicit intervention?" Am. J. Phys. 78(7), 760-767 (2010).
2. C. Henderson and K. Harper, Quiz corrections: "Improving learning by encouraging students to reflect on their mistakes," Phys. Teach. 47, 581 (2009).
3. B. Brown, A. Mason, and C. Singh, "Improving performance in quantum mechanics with explicit incentives to correct mistakes," Phys. Rev. PER **12**, 010121 (2016).
4. A. Collins, J. S. Brown, and S. E. Newman, "Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics," in Knowing, learning, and instruction: Essays in honor of Robert Glaser, edited by L. B. Resnick, pp. 453-494 (1989).
5. V. P. Dennen and K. J. Burner, "The cognitive apprenticeship model in educational practice." In *Handbook of research on educational communications and technology, 3rd edition,* Ed.s J. M. Spector, M.D. Merrill, J. van Merriënboer, and M. P. Driscoll, L Erlbaum Associates, pp. 425-439 (2008).
6. R. Sayer, A. Maries, and C. Singh, "Quantum interactive learning tutorial on the double-slit experiment to improve student understanding of quantum mechanics." Phys. Rev. Phys. Educ. Res. 13(1), 010123 (2017).
7. E. Marshman and C. Singh, "Investigating and improving student understanding of the expectation values of observables in quantum mechanics." Eur. J. Phys. 38(4), 045701 (2017).
8. A. Mason & C. Singh, "Surveying graduate students' attitudes and approaches to problem solving." Phys. Rev. ST – PER 6(2), 020124 (2010).
9. e.g. A. Maries, S.-Y. Lin, and C. Singh, "Challenges in designing appropriate scaffolding to improve students' representational consistency: The case of a Gauss's law problem." Phys. Rev. Phys. Educ. Res. 13(2), 020103 (2017).
10. e.g. Q. Ryan, B. Wilcox, and S. Pollock, "Student difficulties with boundary conditions in the context of electromagnetic waves." Phys. Rev. PER 14(2), 020126 (2018).
11. e.g. B. Wilcox, M. D. Caballero, C. Baily, H. Sadaghiani, S. Chasteen, Q. Ryan, and S. Pollock, "Development and uses of upper-division conceptual assessments." Phys. Rev. ST – PER 11(2), 020115 (2015).