APPENDIX

A. Course-Scale Learning Goals

1. **Math/physics connection:** Students should be able to translate a physical description of a junior-level electromagnetism problem to a mathematical equation necessary to solve it. Students should be able to explain the physical meaning of the formal and/or mathematical formulation of and/or solution to a junior-level electromagnetism problem. Students should be able to achieve physical insight through the mathematics of a problem.

2. **Visualize the problem:** Students should be able to sketch the physical parameters of a problem (e.g., E or B field, distribution of charges, polarization), as appropriate for a particular problem.

3. **Organized knowledge:** Students should be able to articulate the big ideas from each chapter, section, and/or lecture, thus indicating that they have organized their content knowledge. They should be able to filter this knowledge to access the information that they need to apply to a particular physical problem, and make connections/links between different concepts.

4. **Communication.** Students should be able to justify and explain their thinking and/or approach to a problem or physical situation, in either written or oral form.

5. **Problem-solving techniques:** Students should be able to choose and apply the problem-solving technique that is appropriate to a particular problem. This indicates that they have learned the essential features of different problem-solving techniques (eg., separation of variables, method of images, direct integration). They should be able to apply these problem-solving approaches to novel contexts (i.e., to solve problems which do not map directly to those in the book), indicating that they understand the essential features of the technique rather than just the mechanics of its application. They should be able to justify their approach for solving a particular problem.

   …5a. Approximations: Students should be able to recognize when approximations are useful, and use them effectively (eg., when the observer is very far away from or very close to the source). Students should be able to indicate how many terms of a series solution must be retained to obtain a solution of a given order.

   …5b. Series expansions: Students should be able to recognize when a series expansion is appropriate to approximate a solution, and complete a Taylor Series to two terms.

   …5c. Symmetries: Students should be able to recognize symmetries and be able to take advantage of them in order to choose the appropriate method for solving a problem (eg., when to use Gauss' Law, when to use separation of variables in a particular coordinate system).

   …5d. Integration: Given a physical situation, students should be able to write down the required partial differential equation, or line, surface or volume integral, and correctly calculate the answer.

…5e. Superposition: Students should recognize that – in a linear system – the solutions may be formed by superposition of components.

6. **Problem-solving strategy:** Students should be able to draw upon an organized set of content knowledge (LG#3), and apply problem-solving techniques (LG#4) to that knowledge in order to organize and carry out long analyses of physical problems. They should be able to connect the pieces of a problem to reach the final solution. They should recognize that wrong turns are valuable in learning the material, be able to recover from their mistakes, and persist in working to the solution even though they don't necessarily see the path to the solution when they begin the problem. Students should be able to articulate what it is that needs to be solved in a particular problem and know when they have solved it.

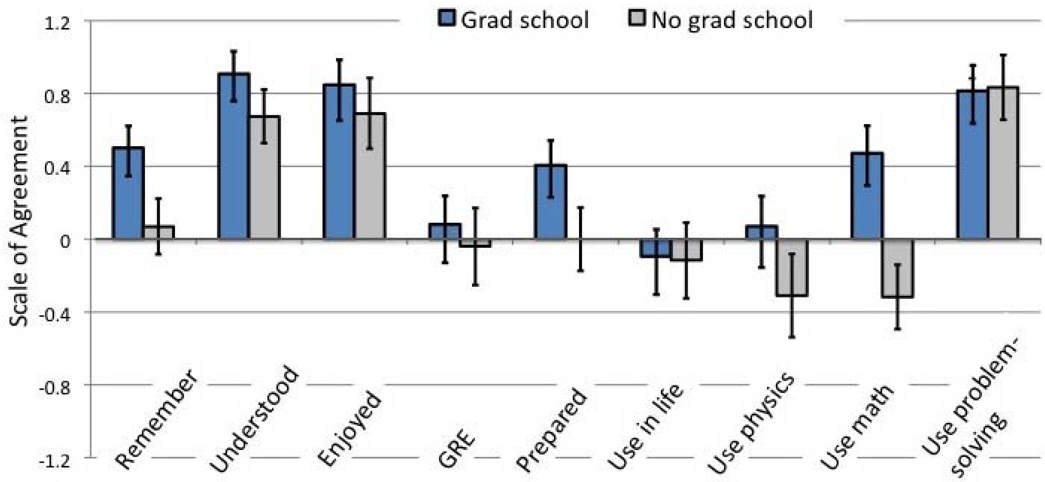
7. **Expecting and checking solution:** When appropriate for a given problem, students should be able to articulate their expectations for the solution to a problem, such as direction of the field, dependence on coordinate variables, and behavior at large distances. For all problems, students should be able to justify the reasonableness of a solution they have reached, by methods such as checking the symmetry of the solution, looking at limiting or special cases, relating to cases with known solutions, checking units, dimensional analysis, and/or checking the scale/order of magnitude of the answer.

8. **Intellectual maturity:** Students should accept responsibility for their own learning. They should be aware of what they do and don't understand about physical phenomena and classes of problem. This is evidenced by asking sophisticated, specific questions; being able to articulate where in a problem they experienced difficulty; and take action to move beyond that difficulty.

9. **Maxwell's Equations.** Students should see the various laws in the course as part of the coherent field theory of electromagnetism; ie., Maxwell's equations.

10. **Build on Earlier Material.** Students should deepen their understanding of introductory course material. I.e., the course should build on earlier material.

B. Alumni survey



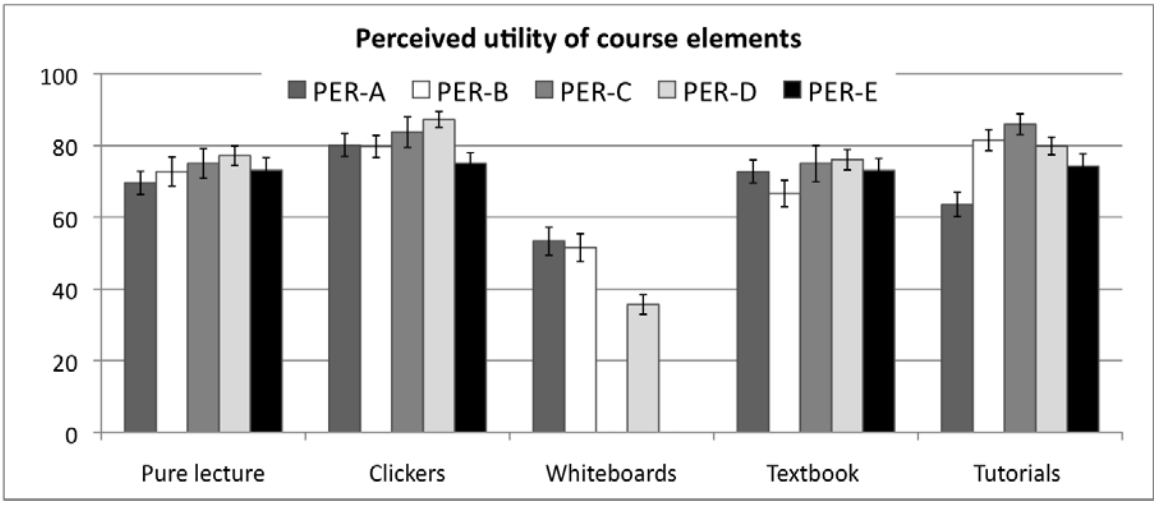
Alumni survey results. (color online) Alumni were asked to answer on the basis of their graduate degree program (if ever enrolled) or current job (if never enrolled in graduate school). Questions were rated on a scale of 1-5 (strongly disagree to strongly agree), and then converted to a scale of -2 (strongly disagree) to +2 (strongly agree) by subtracting 3 from the overall average. Questions were as follows: (1) I remember what I learned in PHYS301, (2) I understood the material in PHYS301, (3) I enjoyed PHYS301, (4) PHYS301 prepared me well to take the GRE (if applicable), (5) PHYS301 prepared me well for my job or graduate school, (6) I use something I learned in PHYS301 in my life *outside* of my primary job or graduate research, (7), I use the *physics* I learned in PHYS301 in my primary job or graduate research, (8) I use the *math* I learned in PHYS301 in my primary job or graduate research, (9) I use the *problem-solving techniques* or approaches that I learned in PHYS301 in my primary job or graduate research.

C. Detailed information about courses in the study

		PER-A	PER-B	PER-C	PER-D	PER-E	STND-A	STND-B
Pedagogy		Research-based transformations					Traditional lecture	
Instructor		PER1 + Non-PER1	Non-PER2	PER2	Non-PER2	Non-PER3	Non-PER4	Non-PER5
Course N		48	37	22	56	46	41	41
Major (%)	PHYS	27	49	50	55	57	39	54
	EPEN	48	22	36	30	20	34	15
Females (% of class)		27	26	25	16	11	25	7
Ave lecture attendance (% of class)		86	77	94	77	76	73	69
Ave students attending a tutorial (% of class)		30	42	44	37	38	N/A	N/A
Ave students attending a help session (% of class)		30	Unknown	58	Unknown	Unknown	Unknown	N/A
FCQ	Instructor	85	97	98	95	97	87	90
	Course	80	90	92	87	85	85	85
Demographics								
Cumulative GPA		3.1	3.1	3.1	3.2	3.2	3.1	2.9
Physics GPA		3.0	3.1	3.0	3.2	2.9	2.9	2.9
BEMA	Post-102	61	58	69	58	55	60	55
	Post-301	66	63	71	63	N/A	64	N/A
CUE results								
		55.0	59.6	60.1	55.6	53.2	39.8	46.2

Course demographics. (color online) Courses involved in this study, not in chronological order. PER 1 and 2 are different PER faculty. Non-PER 1-4 are different non-PER faculty. “EPEN” = Engineering Physics, and “PHYS” = Physics. Attendance is an average of the attendance on the days that the FCQ and CUE were administered, and the clicker attendance scores (where applicable). Students who missed two exams and/or did not take the final exam were excluded from study, and students who took the course more than once (without failing/dropping) were included only in the first enrollment. “FCQ” = Faculty Course Questionnaire given at the end of the semester, given out of 100%. “Instructor” = “Rate this instructor compared to all your other university instructors.” “Course”= “"Rate this course compared to all your other university courses.” Cumulative and Physics GPA are calculated prior to the start of PHYS301. “BEMA” = Basic Electricity and Magnetism Assessment, given as a Post-test after introductory physics (PHYS 102) and Junior E&M I (PHYS 301).

D. Student attitudes towards elements of the course



Perceived utility of course elements. Student responses to the question, "Please rate how useful for your learning each of the following class activities are in Phys 301 (Mark N/A if you did not attempt to use this resource for help)". Axis represents average strength of utility calculated by a weighted average, such that "very useful" is weighted as "4" and "completely useless" as "0", normalized to 100%, and then averaging over all student responses in that course. Thus, a larger number indicates a stronger perception of utility of that element.

E. Details of the Multiple Regression

In order to do determine the effect of tutorials on student exam and CUE scores when background variables are taken into account, we model these outcome variables according to the following equations:

$$OUTCOME = b_0 + (\sum_{k=1}^{N} b_k \times VAR_k) + (b_{TUT} \times TUTORIAL)$$

where *OUTCOME* is either the CUE post-test score or the z-score of the average of the three course exams, *TUTORIAL* represents the percent of tutorials attended throughout the term, VAR_k are the background variables that are included in the model, b_k are the coefficients for each term. The value of b_{TUT} is the coefficient for the *TUTORIAL* variable, and gives the relative impact of attending the tutorials on *OUTCOME,* all other factors being equal. Variables are entered into the model manually, and background variables, *VAR*, are entered until a model with a high R^2 and the fewest possible background variables is obtained. Then the variable *TUTORIALS* is added. The statistical difference between models is calculated using the F-test.

Results of the regression are shown below in Table 2. In all models, the only background variables that enter into the regression were the calculated GPA in all previous physics courses ("PHYS GPA"; including both pre-requisites and electives) and the BEMA score from after introductory physics. Variables considered that did not enter into the model were: Pre-requisite math courses, GPA in all prior math courses, cumulative GPA, CUE pre-test, and lecture attendance (i.e., their variance was accounted for by the inclusion of PHYS GPA or BEMA). Because only some students have BEMA scores, the inclusion of this variable reduces N significantly (and may bias the sample). Thus, we only include the BEMA as a predictor for students who have BEMA scores, and present those models separately.

	CUE Model 1A	**CUE Model 1B** (with tutorials)	**CUE Model 2A**	**CUE Model 2B** (with tutorials)	**Exam Model 1**	**Exam Model 2**
Population	All students	All students	Students with BEMA	Students with BEMA	All students	Students with BEMA
N	156	156	87	87	192	103
Model based statistics						
Multiple R^2	0.23	0.26	0.40	0.46	0.46	0.60
F statistic	47.24	27.08†	58.38	36.77††	166.93	156.3
Residual standard error	15.26	15.04	13.01	12.41	0.77	0.66
	b_k	b_k	b_k	b_k	b_k	b_k
Predictors						
Phys GPA	0.48**	0.45**			0.68**	0.78**
BEMA			0.64**	0.63**		
Tutorials		0.17*		0.24**		

TABLE 3. Multiple regression models to determine impact of tutorials on CUE and exam scores. All F statistics are significant at p<0.0001 value. Coefficients reported are significant at the p<0.05 (*) and p<0.01 (**) level; if a coefficient is not reported, then it did not enter into the model as a significant predictor. The y-intercept (b_0) is insignificant for all models, and thus is not reported. Significant differences from the previous listed model is designated by †, p<0.05 and †† p<0.01. Note that we report results on tutorial attendance as a continuous variable (i.e., percentage of tutorials attended), but similar results were obtained when using a binary variable (number of tutorials ≤ 3 or >3) or trinary variable (percent of tutorials <20%, between 20% and 50% inclusive, or >50%).

F. Gender Differences

Did males and females experience different outcomes of the course transformations? At the introductory level, there is a well-documented gender gap in course performance. Female students tend to score more poorly on both conceptual measures such as the BEMA[i] and on traditional exams[ii]. There is mixed evidence as to whether interactive engagement in introductory courses reduces the gender gap[iii] or not[i,iv,v]. This gender gap at the introductory level at CU was shown to be related primarily to the female students' lower scores on a conceptual exam at the beginning of the introductory course[ii]. Because the students in upper-division courses are among the strongest, most committed, and best prepared students from the lower-division courses, we did not expect to see strong gender differences in student performance. However, we do find a gender gap on many of the measures in this course, which appears to be ameliorated by the course transformations. We find evidence that female students enter PHYS301 less well-prepared than male students, and receive lower grades in PHYS301, and that the PER-based courses may have helped to close this gender gap, primarily by improving scores on exams. The analysis is found below.

Pre-course measures: In the current study, we find that the gender gap in pre-course measures persists at the upper-division. BEMA scores (after introductory physics) of women in the PER-based courses were lower than those of men (see Table 4). Female students' CUE pre-test scores are also lower than male students. Females score lower on these performance measures *despite* higher grades in previous courses compared to male students as determined by their GPA in prior physics and math courses, pre-requisite courses, and cumulative GPA.

Post-course measures: CUE. The gender gap on BEMA scores persists post-PHYS301 (for that subset of students who were given the BEMA after PHYS301). This is not surprising, as the course transformations did not strongly affect introductory level conceptual understanding.[vi] The gender gap on CUE score, however, is *not* present at the end of the course: Men and women do *not* have significantly different CUE post-test scores in the PER courses. (Gender differences on the post-course CUE in STND-A are not significantly different, but it is difficult to come to clear conclusions based on this result due to low N and the fact that this is one isolated course). Thus, the course transformations may assist in closing the gender gap on conceptual assessments directly related to the course content. This may be due, in part, to a greater participation on the part of the women in the innovative aspects of the course. Women attended more lectures, and slightly more of the optional tutorials (non-significant).

Post-course measures: Exams. To provide a more robust comparison between PER and STND courses, we also examined course grades, as a proxy for course exam scores, in PER courses and the past 10 semesters of traditional lecture-based courses. In this analysis we also found evidence for a gender gap on course exam scores in PHYS301 taught with standard instruction, which is closed in the PER-based courses. Previous work[vii] has shown that female students tend to score more poorly than male students on course exams, but have stronger homework scores, resulting in similar course grades. We suggest that this is also the case in several previous iterations of PHYS301 taught with a standard approach. In the PER-based courses, however, the gender gap on exam scores appears to be eliminated. Below, we explain the evidence for these statements.

We analyzed data from the past 10 semester of PHYS301 prior to the course transformations, and found that female students in PHYS301 received significantly lower course grades – a trend that holds when looking at each semester individually or the 10 semesters on average. Do these differences arise from homework or exam scores, or both? Without homework and exam data for these early courses, we must infer from our data on STND-A and PER courses. We find that female students receive homework scores that are *higher* than the men in both the STND-A and PER courses. It is reasonable to assume that this is also the case in the earlier iterations of PHYS301. Thus, the female students' *lower* course grades in the past five years of PHYS301 are likely due to lower exam scores (rather than homework scores). This conclusion may be supported by the fact that we do indeed observe such a pattern in STND-A course, where female students' exam z-scores are 0.38 *less* than male students on average (though this difference is statistically insignificant.) We note, however, that this number is based on a very small sample size (N=10 female, N=20 male), So, we conclude that female students' exam scores are likely historically lower than men's in PHYS301 taught using standard instruction.

In the transformed PER-based PHYS301 courses, however, female and male students' exam z-scores (and course grades) are statistically *equivalent* (difference = 0.07). Thus, the course transformations appear to be closing a previously unreported gender gap in upper-division exam scores.

	Course type	Female	Male
BEMA scores (after introductory physics)	PER	54.5%**	65%
CUE pre-test scores	PER	26%*	34%
GPA in prior math courses	PER	3.3*	2.9
GPA in prior physics courses	PER	3.3*	3.0
GPA in pre-requisite courses	PER	3.3*	3.1
Cumulative GPA on entering PHYS301	PER	3.3*	3.1
Lecture attendance	PER	89%⁰	79%
PHYS301 homework z-score	PER	+0.39**	-0.1
Exam z-scores	PER	0.06	-0.01
PHYS301 Course Grade	PER	2.9	2.7
CUE Post-test	PER	53.5%	56.9%
CUE Post-test	STNDA	36.5%	41.5%
PHYS301 Course Grade (historic)	10 past STND	2.76	3.01
PHYS301 homework grade (STND-A course)	STND-A	+0.55**	-0.17

Table 4. Gender differences in PHYS 301. In PER courses, student N for females ranges from 29-39, and N for males ranges from 106-161, depending on available data, with the exception of homework scores (N = 10 female, N=30, male). Historic STND course grades are taken from the 10 previous semesters, and include N=64 female and N=303 male. STND-A course is comprised of N=10 female and N=30 male, and CUE scores were available for N=9 female and N=18 male. Significant differences are denoted by ** at $p<0.001$ level, * at $p<0.05$ level, and ⁰ for $p<0.07$. STND-B is not included in this analysis due to the lack of women (N=2) and lack of available grades at time of analysis.

It is hard to determine how broadly we may apply these results. Looking across a broad spectrum of physics and mathematics courses, female students receive similar or higher course grades than men, typically by about 0.2 grade points. However, in investigating results from an external institution (STND-6; see Table 2 in main paper), we find a similar trend to that observed in PHYS301: Female students had lower course grades (71% vs. 77%), and lower CUE pre-test (37.1 vs 49.2) and post-test scores (33.7 vs 47). Because of very low numbers of female students (14 out of a class of 130), these results are only marginally significant ($p<0.1$).

Thus, the course transformations appear to have a positive effect in closing the gender gap in performance on both conceptual measures targeted to PHYS301, as well as more traditional measures of assessment. These findings may be limited to the content area of PHYS301, which appears to pose more difficulty for female students than do other courses, in which they excel. These results are of particular note since our student population of physics majors is very different from the introductory students studied previously, and because of the documented loss of female physics majors at the end of an undergraduate degree[viii].

---

[i] S. J. Pollock, N.D. Finkelstein and L.E. Kost, "Reducing the Gender Gap in the Physics Classroom: How sufficient is interactive engagement?" Phys. Rev. ST Physics Ed. Research, **3**, 010107-1-4 (2007).
[ii] L.E. Kost, S.J. Pollock, & N.D. Finkelstein, "Characterizing the Gender Gap in Introductory Physics," Phys. Rev. ST Physics Ed. Research, **5**(1), 010101-1-14 (2009).
[iii] M. Lorenzo, C. Crouch and E. Mazur, "Reducing the gender gap in the physics classroom," Am. J. Phys., **74**, 118-122 (2006).
[iv] E. Brewe, V. Sawtelle, L. H. Kramer, G.E. O'Brien, I. Rodriguez and P. Pamela, "Toward equity through participation in Modeling Instruction in introductory university physics," Phys. Rev. ST Physics Ed. Research, **6**, 010106-1-12 (2010).

[v] J. Docktor and K. Heller, “Gender differences in both Force Concept Inventory and introductory physics performance,” PERC Proceedings 2008, *AIP Conference Proceedings,* **1064**, 15-18 (2008).
[vi] S. J. Pollock and S. V. Chasteen, "Longer term impacts of transformed courses on student conceptual understanding of E&M," PERC Proceedings 2009, *AIP Conference Proceedings,* **1179**, 237-240 (2009).
[vii] L.E. Kost, S.J. Pollock, & N.D. Finkelstein, "Characterizing the Gender Gap in Introductory Physics," Phys. Rev. ST Physics Ed. Research, **5**(1), 010101-1-14 (2009).
[viii] http://www.nsf.gov/statistics/women/