The effect of instructional technique on physics students' representational skills.

Patrick B. Kohl and Noah D. Finkelstein

University of Colorado at Boulder

Abstract

In a recent study we showed that physics students' problem solving performance can depend strongly on problem representation[1]. In this paper, we continue that study in an attempt to separate the effect of instructional technique from the effect of content area. We determine that students in a reform-style introductory physics course are learning a broader set of representational skills than those in a more traditional course. We also analyze the representations used in each course studied and find that the reformed course makes use of a richer set of representations than the traditional course and also makes more frequent use of multiple representations. We infer that this is the source of the broader student skills.

I. INTRODUCTION

Our previous work[1] showed that physics students' problem solving success depends strongly on the representation of the problem, corroborating other work in PER[2]. That is, whether one represents a physics problem in terms of words, equations, graphs, or pictures can have a significant impact on student performance on that problem. In this, we are studying students representational skills, which the PER community has been actively investigating for some time. A number of studies have investigated student skills with particular representations of physics concepts and problems[2][3][4][5]. Others have studied how best to explicitly teach students the skills involved in using multiple representations to solve physiscs problems[6][7][8].

In our previous work[1], we also investigated whether providing students with a choice of representational format would help or harm their problem solving performance. That is, we asked whether students are capable of determining which representations are most helpful to them on problems of a given topic. The results from our study were complicated: Student problem solving performance often depended very strongly on whether or not they had a choice of problem format, but the strength and direction of the effect varied strongly with the representation, the subject matter, and the instructor. Such a study considers students' meta-representational skills, and asks what students know *about* representations. Studies on meta-representational competence are relatively rare, but the topic is becoming increasingly popular[9][10][11].

Our first study took place in two different introductory physics courses: Physics 201, taught by a reform-style instructor, and Physics 202, taught by a mostly traditional professor. In either course, we found interesting differences in performance from representation to representation, leading to micro-level questions regarding why and how student performances vary as they do. Previously, we found evidence that subtle contextual features (features of the concept, the representation, and the students' prior instruction) could have significant effects on student performance. We will consider these micro-level questions in more detail in a future paper, where we present the results of detailed problem-solving interviews.

Our first study also found that the character of the 201 data was strikingly different from the character of the 202 data. In particular, there were very strong performance differences

between the students who received a choice of problem format and students who did not in the traditional-style 202 course. Students in the reform-style 201 course showed much smaller differences. This led to more macro-level questions: Was this qualitative difference in performance data a result of the different instructional style, the different content area, or some combination? Also, what was it about the difference in content or instruction that drove the difference in performance? In the present paper we address these questions in two parts. First, we repeat our earlier study in a Physics 202 section taught by the reform-style instructor, allowing a direct comparison with the results from the traditionally taught 202 section. In this manner we determine that the broad differences in performance data are a function more of instructional style than content area. In the second part, we analyze the specific differences in representation use in these classes (in lectures, exams, and homeworks), allowing us to conclude that a pervasive use of multiple representations in a physics course can result in students learning a broader set of representational skills.

A. Previous study - Overview

Here, we present a brief overview of our previous study. We gave students in two on-sequence large-lecture algebra-based introductory physics classes (Physics 201 and 202) homeworks that had four problems, one in each of four different representational formats (verbal, mathematical, graphical, and pictoral). We then gave these students a one-question quiz during their recitation section. This quiz was in one of four formats, and one-half to two-thirds of the sections were allowed to choose which quiz format they would receive without examining the questions ahead of time. The rest of the sections received a quiz in a random format, providing a control group. This process was repeated later in each term for a different topic, providing a total of four trials. The 201 quiz topics were springs and pendulums, while the 202 quiz topics were diffraction and spectroscopy.

We found ample evidence that student problem-solving performance can (but does not necessarily) depend strongly on representation, strengthening results seen elsewhere.[2][12] This naturally leads to the question of why problem solving depends on representation. In some cases, a problem feature was present in one representation but was either not present or not as attractive in another representation. We also observe in interviews that student problem solving strategies depend sensitively on problem representation, a result that we

Quiz subject	Verbal	Math	Graphical	Pictoral
202 Diffraction	**X**	X	0.04	**0.03**
202 Spectroscopy	**0.002**	**0.0001**	**0.0004**	0.001
201 Springs	X	**0.09**	**X**	**0.07**
201 Pendulums	**X**	**X**	**X**	**X**

TABLE I: Statistical significance of the quiz performance differences between the format choice and control groups in the traditional 202 and reform 201 sections. Numbers are p-values using a two-tailed binomial proportion test. X denotes a p-value of greater than 0.10. Bold indicates that the choice group had higher performance than the control group.

will consider in depth in a future paper.

The effect of providing a choice of problem formats was complex. In some cases giving students a choice of format resulted in significantly increased performance with respect to the random-format control sections. In other cases, the choice of formats resulted in significantly decreased performance. These choice/control differences were often quite strong, and also varied by topic area. A group that chose one format would do significantly better than the corresponding control group for one topic, and then would do significantly worse on the next quiz on a different topic. The performance differences between choice and control groups are shown in Table I, reprinted from our previous paper.[1] Each entry is the statistical significance (p-value, using a 2-tailed binomial proportion test) of the difference between the performance of the choice and control groups for a given combination of quiz topic and quiz format. Bold entries indicate that the sections that had a choice of problem representation outperformed the control groups; unbolded indicates the opposite. An X indicates a p-value greater than 0.10. Notable features include the strength of many of the choice/control splits and the changing direction of the effect across topic. Especially interesting is the fact that the 202 course showed much stronger choice/control splits than the 201 course.

Explaining these results is not trivial. The data do not support a simple explanation such as an alignment of student learning styles with problem representations. We can make progress by attempting to explain the broad differences between the results of the 201 and 202 studies. As we noted in the introduction, these courses were taught by different professors and covered different material, and so the differences observed could conceivably be explained

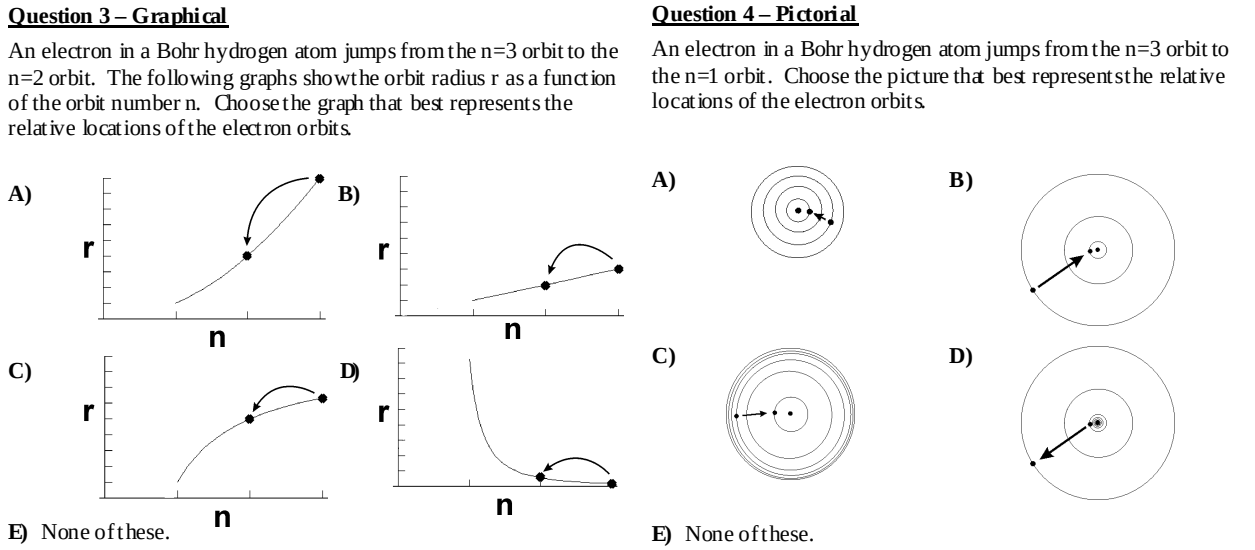


FIG. 1: Nearly-isomorphic homework problems (in graphical and pictoral/diagrammatic formats) regarding Bohr-model electron orbit radii.

by differences in instruction, differences in content, or some combination. We note that the 202 course was taught by a relatively traditional professor, though he used reform techniques such as concept tests[13] and personal response systems[14] to some extent. The 201 course was taught by a more reform-style professor, who made heavier use of "Clicker" questions and well-integrated lecture-demos. These students were a part of the same general population as the 202 students, but were from a different year. Our previous hypothesis was that the much different approach of the reformed 201 course resulted in students having a broader set of representational skills. Thus, whether or not they received their "preferred" representation made less of a difference (positive or negative) in performance. To test our hypothesis by separating out the effect of instruction from the effect of content, we repeated the Physics 202 study in the spring of 2005, when the course was taught by the reform-style professor in charge of the 201 course in the earlier study. We predicted that given the same quizzes, the choice/control splits would be much weaker than they were in the original study with the traditional 202 professor.

Spectroscopy Problem – Pictorial Format

The Balmer series of spectral lines is shown below, as seen through a spectrometer:

Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus and the ionization energy is 13 eV instead of 13.6 eV. Choose the picture that best represents what the new spectrum would look like.

B) 

Spectroscopy Problem -- Verbal Format

Consider the Balmer series of spectral lines from hydrogen gas. Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus. This means that the ionization energy for the electron will be smaller. What will happen to the Balmer lines that we see?

B) The spectral lines will all shift to shorter wavelengths (toward the bluer colors).

FIG. 2: Problem statement and second answer choice for the verbal and pictoral format quizzes. The other distractors match up between the formats as well.

II. STUDY: SEPARATING THE EFFECTS OF CONTENT AND INSTRUCTION

A. Methods

1. *Study homeworks and quizzes*

The study was conducted in the same way, using the same materials, as the 202 study in the earlier work (see the brief recap above or [1]). In Figure 1 we see two example homework problems, in graphical and pictorial formats. In Figure 2, we see an example of the verbal and pictorial format spectroscopy quizzes. The problem statements and one distractor are shown for each. Note that in all quizzes, the statements and distractors mapped from one quiz representation to the next. The homework problems could not be made completely isomorphic since the students had to do all of them, but they were made as similar as was reasonable. For instance, some problems asked students how the radius of a Bohr-model electron's orbit varied with the orbital number. The different questions would be the same except for which particular transition (n = 3 to n = 2, for example) was under consideration and the representation.

Reform Course, Spring 2005	Verbal format	Mathematical	Graphical	Pictoral
202 Diffraction/Interference HW (N = 332)	0.44	0.36	0.39	0.46
202 Spectroscopy/Bohr HW (N = 341)	0.63	0.60	0.55	0.48

TABLE II: Fraction of students answering a homework problem correctly, broken down by representational format and topic. Standard errors vary but are on the order of 0.02.

Reform Course	Control (random format) group				Choice group			
Spring 2005	Verbal	Math	Graphical	Pictoral	Verbal	Math	Graphical	Pictoral
202 Diff.	0.19 (46)	0.35 (46)	0.14 (46)	0.18 (44)	0.15 (16)	0.57 (34)	0.13 (37)	0.21 (77)
202 Spec.	0.59 (46)	0.39 (46)	0.57 (42)	0.54 (46)	0.41 (17)	0.32 (25)	0.49 (37)	0.52 (89)
202 Diff. (dist)	0.33	0.45	0.44	0.45	0.26	0.22	0.41	0.32

TABLE III: Quiz performance of students from the random-format recitation sections (left) and from the recitations sections that had a choice of formats (right). The number of students taking a quiz is in parentheses. The quiz topics are diffraction and spectroscopy. Standard errors vary and are not shown. The last line indicates how many students chose a particular distractor on the diffraction quiz, as discussed in the text.

B. Data

In Table II, we see the performance of the reformed Physics 202 students on the pre-recitation homeworks. Notably, the difference between the graphical and pictorial questions on the Bohr model homework is much smaller than it was in the traditional 202 course (0.77 vs. 0.62).

In Table III, we see the performance of the students on the diffraction and spectroscopy recitation quizzes, broken down by representation and by whether the students were in a format choice or control group. Performance variation across representation was generally weaker in this course than it was in the traditional 202 section.[1] We should also note the very low success rates on the diffraction quiz. The two 202 courses placed different focuses on the different subtopics available, and the reform section spent very little time on double finite-width slit diffraction. Student comments and performance suggest that most students treated this as a double infinitesimal-width slit problem. One of the distractors is correct for such a problem, and student selection of this distractor is noted in the (dist) line of Table

Quiz subject	Verbal	Math	Graphical	Pictoral
202 Diffraction, Reform	**X**	0.06	**X**	X
202 Spectroscopy, Reform	X	X	X	X

TABLE IV: Statistical significance of the quiz performance differences between the format choice and control groups in the reformed 202 section. Numbers are p-values using a two-tailed binomial proportion test. X denotes a p-value of greater than 0.10. Bold indicates that the choice group had higher performance than the control group.

III.

In Table IV, we show the statistical significances (p-values using a two-tailed binomial proportion test) of the differences between the choice and control groups on the different topics and formats. Note that these data are essentially the same if considers the distractor mentioned above for the diffraction quiz (the math p-value changes from 0.06 to 0.03; all others remain insignificant). When comparing these data to those in Table I, we see that they are very similar in character to the reformed 201 course, and much different in character than the traditional 202 course. This suggests that the choice/control splits (or lack thereof) are associated more closely with the instructor and course environment than with the content area. Thus, we chose to analyze these environments in more detail to with the goal of finding any differences in representational content and use that may be present.

III. STUDY: ANALYSIS OF COURSE ENVIRONMENTS

The second part of this study involves characterizing the representation use in each of the classes under consideration: The reform 201, the traditional 202, and the reform 202. These courses had many components, including lecture, lab/recitation, exams, and homeworks. The homeworks included problems assigned and graded through a web-based automated system (LON-CAPA[15]), and long-answer hand-graded problems. In comparing the courses, we judged the labs/recitations and CAPA-based homeworks to have very similar representational character. We thus focus our analysis on the lectures, exams, and long-answer homeworks. This provides two views of the class. We see how the use of physics representations was modeled for them (the lectures), and how they were held responsible for using physics representations themselves (homeworks and exams).

A. Methods

1. Analysis of course lectures

We videotaped several lectures from each of the three courses. The lectures covered the topics from the quizzes and related material. We chose three lectures from each course for analysis, with each set of lectures spread over different topics.

In order to quantify the different uses of representations in the different lectures, we divided each tape into one-minute sections. For each segment, we noted which representations were used significantly according to the following rubric:

- Verbal: Writing sentences expressing an idea or concept on the board; presenting and explicitly referring to a slide with verbal-only content for the sake of the point at hand (words surrounding mathematics are not counted).
- Mathematical: Writing equations; explicitly referring to equations for the sake of the point at hand; doing mathematics. Writing numerical data by themselves is not counted (for example, simply writing h = 6.636x10^{-34} Js does not count).
- Graphical: Drawing or modifying a graph; explicitly referring to a graph for the sake of the point at hand.
- Pictorial: Drawing or modifying a picture; explicitly referring to a picture for the sake of the point at hand.
- Physical demonstration: Carrying out a physical demonstration.

Note that for lectures, we have added the representational category "Physical demonstration." We also noted which intervals include Clicker questions. Finally, any interval in which more than one representation was used was tallied as a Multiple Representations interval (the Clicker category did not count for this purpose).

Because the professor is talking during nearly every part of a lecture, we did not count that towards the use of verbal representations. This is an example of a broader feature of this study: the privileged position of the verbal representation. Essentially every aspect of the course had some verbal component (even math problems include explanatory text), and so we necessarily have stricter standards as to what counts as verbal representation

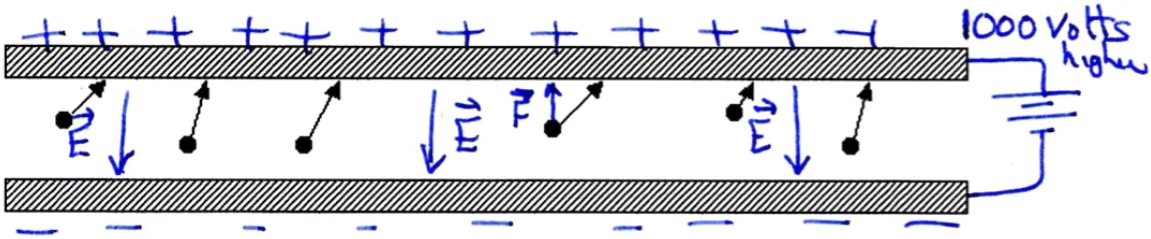
use compared to the other categories. These standards are explained in every category of analysis.

Once a lecture was coded, we calculated the fraction of the lecture that showed use of each of the representational categories. We then averaged across the three lectures from each class to obtain an average representational content for those courses' lectures. To test for inter-rater reliability, two individuals coded lectures working from a written description of the coding standard. Results were extremely similar; the average difference between raters on each category was 1.3%.

2. Analysis of exams

Each of the three courses considered here had three exams, not including the final. The final exam took place after all the recitation quizzes and homeworks, and thus could not have had an effect on student study performance. We quantified the fraction of each of these exam that could be described as verbal mathematical, graphical, and pictorial in representation. We also quantified the fraction of each exam that explicitly required the use of multiple representations. The procedure was as follows: On an exam, each problem is designated as having a verbal, mathematical, graphical, or pictorial component or some combination thereof. The total fraction of the exam composed of a particular representation was defined as the fraction of points possible that came from problems belonging to that representation. Problems with more than one representational component had their points counted in full towards each of the relevant representations; no effort was made to weight the components (for instance, we did not designate a problem as 80% pictorial and 20% mathematical). Both the problem presentation and the intended solution were considered. For example, a problem that requires a student to read information off of a provided graph in order to do a numerical calculation is designated mathematical/graphical. If it is a ten point problem, ten points are assigned to both the mathematical and graphical categories. Thus, an exam that is rich in representations can have more than 100 points of representational content assigned to it in this analysis. Any problem that explicitly involves more than one representation has its points counted towards a Multiple Representations category as well. Once we characterized each exam in terms of its representational content, we calculated the average representational content of the exams in each of the courses.

An electrostatic air filter uses electric fields to attract charged particles of dust, pollen, etc. to one of two electrically charged metal plates.



The dust particles are first electrically charged by a small electric discharge and then blown along with the air left to right between the charged plates using a fan. The two plates are wired to a voltage supply so the top plate is 1000 volts higher in voltage than the bottom plate. The two plates are spaced 1 cm apart. The particles are strongly attracted to the top plate and stick there. They can later be removed by just wiping the top plate with a cloth.

a) How are the particles charged, i.e. are they positive or negative? Explain.

Negatively charged since they are attracted to the positive plate.

b) Indicate on the figure the direction on the electric field between the two plates and determine the magnitude of that electric field.

E points downward; away from positive, toward negative

$$|\vec{E}| = \left|\frac{\Delta V}{\Delta x}\right| = \frac{1000\,V}{0.01\,m} = 100{,}000\,\frac{V}{m} = 1\times 10^{5}\,\frac{V}{m}$$

FIG. 3: Example exam problem with pictorial, graphical, and verbal components. The problem is from a reform 202 exam.

The representational categories are defined as follows:

- Verbal: For multiple choice questions, the available answers are phrases that are conceptual or qualitative in nature. For long-answer questions, there is an "Explain your answer" or similar written component.
- Mathematical: The problem requires numerical or algebraic calculation, manipulation, or interpretation, or other significant and explicitly quantitative reasoning.
- Graphical: A graph presents relevant information, or students have to construct a graph. Diagrams that have labeled axes have a graphical component.
- Pictorial: Students must draw or modify a picture, or a picture contains needed information, or a picture meaningfully depicts the relevant physical situation. A picture of a CD accompanying a "plug 'n chug" problem about the wavelength of a

CD-reading laser would not fit this standard.

In Figure 3 we see an example exam problem with solution. Part A of this problem was judged by the above standards to have verbal and pictorial components. Part A was worth 11 points (out of 100), and so there were 11 points of verbal component and 11 points of pictorial component for the sake of the averaging. Part B was judged to have mathematical and pictorial components.

3. Analysis of long-answer homeworks

In addition to web-based CAPA problems (formulaic in nature for all classes studied here), students were assigned several additional homeworks requiring more in-depth work. Here, we consider those from the two 202 classes (traditional and reformed) for the sake of direct comparison. We use essentially the same coding scheme as with the exams: The representational content of each assignment is broken down and weighted according to point value. Then, the content of all assignments is averaged together. The reform 202 course had eight homeworks of this sort, and the traditional 202 had fifteen.

B. Data

1. Lecture content

In Figure 4, we see the representational content in the reformed Physics 201, reformed 202, and traditional 202 lectures according to the standards described previously. Differences exist between all three sets, suggesting (not surprisingly) that both instructor and content have a significant effect on representation use. That is, a particular instructor will use a different selection of representations in different courses, and in a particular course, different instructors will use different selections of representation. Most relevant to us is the comparison between the reform and traditional sections of 202. The reform section shows a broader selection of representations, with the verbal, math, graphical, and pictorial fractions summing to 1.04 versus 0.83 in the traditional section. We also see more use of multiple representations (0.35 versus 0.22), and much more use of interactive Clicker questions (0.51 versus 0.23). The Clicker category does not describe a representation in itself; rather, it

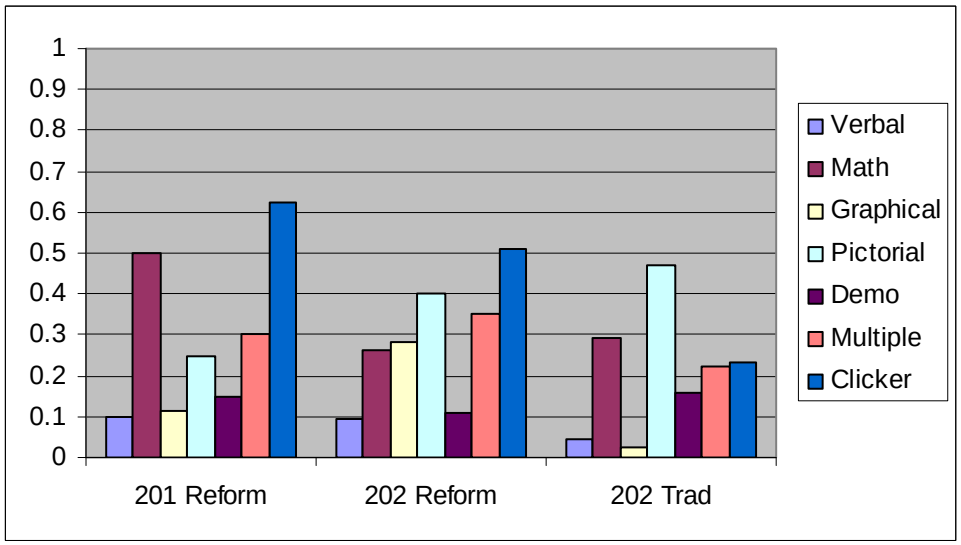


FIG. 4: Representational content of the lectures for the reformed 201, reformed 202, and traditional 202 courses. "Multiple" category indicates use of multiple representations. Clicker category indicates fraction of class time involving questions that used a personal response system.

tells use something about *how* representations are used in the course.

2. Exam content

In Figure 5, we show the representational content of the exams in the reformed Physics 201, reformed 202, and traditional 202 courses. These data show the average across all exams in each course, excluding the final exam. We see the fraction of the exam problems (weighted according to their point value) that were verbal in nature, mathematical, graphical, and pictorial. We also see the fraction of the exam problems that required explicit use of multiple representations.

It is clear that the exams from the reform sections of 201 and 202 made use of a broader selection of representations than the traditional 202 section. Perhaps most striking is the difference in the proportion of multiple-representations problems, with 52% of the reform 201 and 74% of the reform 202 exam problems making use of multiple representations, and 26% of the traditional exam problems doing so.

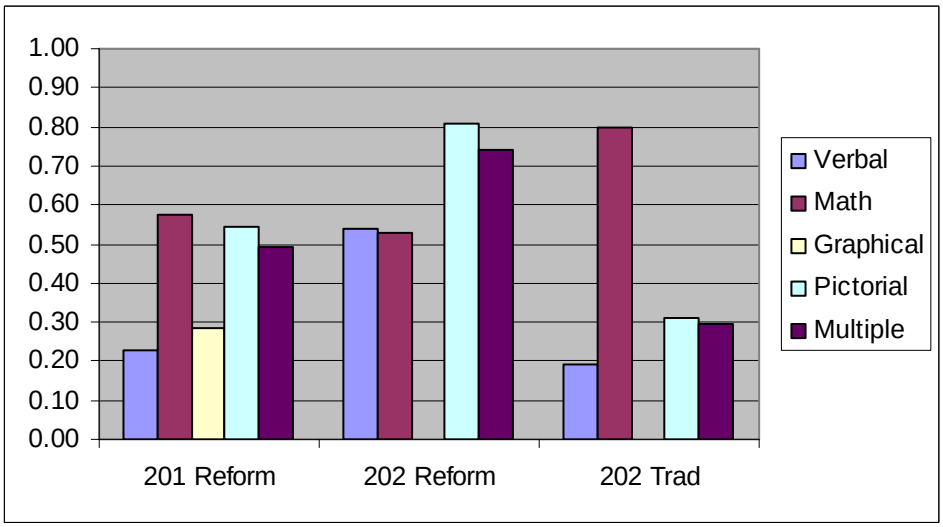


FIG. 5: Distrubution of representations used in the exams in the three courses studied here. Also includes fraction of the exam problems that required explicit use of multiple representations.

3. Homework content

In Figure 6 we see the distribution of representations used in the reformed and traditional 202 course homeworks. The distributions are very similar, with the reform course making somewhat greater use of verbal representations and the traditional course making greater use of pictorial representations. The data are similar enough that we can make no claims regarding significant differences. However, we will take this opportunity to point out that our representation coding scheme for the homeworks says little about how these different representations are used, which is likely relevant to student performance. For example, four of the eight reform 202 homeworks had students perform estimations, while none of the traditional homeworks did. While it is our opinion that quite generally representations were used more effectively in the reform course, we will mainly restrict our arguments to the more quantifiable representational content.

IV. DISCUSSION

Our original hypothesis appears to be validated: The reform 202 course shows choice/control performance splits that are much more consistent with the reform 201 data

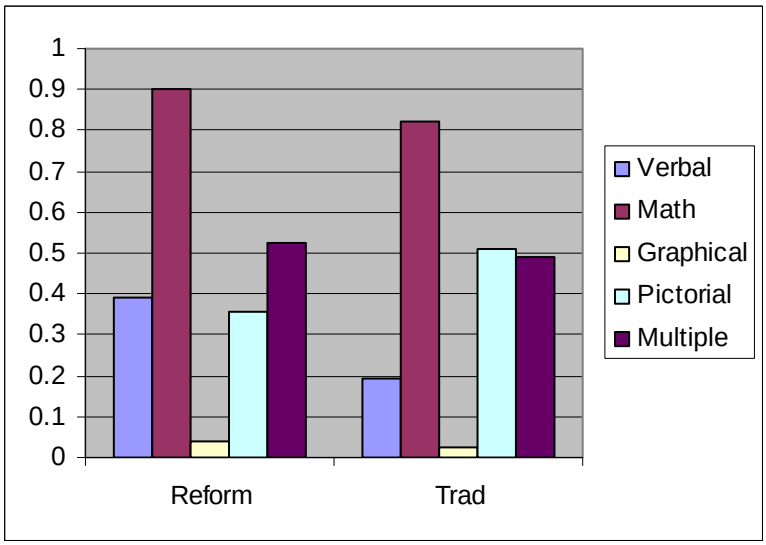


FIG. 6: Distrubution of representations used in the long-format homeworks in the two Physics 202 courses studied here (reform and traditional). Also includes fraction of the homework assignments that required explicit use of multiple representations.

than with the traditional 202 data. We thus conclude that these choice/control splits are associated more with instructional environment than content area. The course analysis data demonstrate that major components of the class (in particular the lectures and exams) were strikingly different in their representational character, with the reform content being richer and using multiple representations more frequently. While we also believe that the reform course made more productive use of these representations, we will make no attempt to quantify that in this paper.

This richer use of representations in-class is consistent with the notion that these students are learning a broader set of representational skills, which could explain the choice/control splits or lack thereof. With this broader set, working in a preferred representation as opposed to a disliked one has less impact on performance. It is also perhaps significant that the reform 202 section shows generally smaller performance variations across representation within the choice and control groups. Further, this interpretation allows for the fact that the direction of the choice/control splits can vary across topic. The previous study and student interviews not analyzed here show that students have definite opinions regarding their abilities and the relative worth of different problem representations and that these opinions are constant across topics. However, it would appear that their problem-solving

skills in different representations are not constant across topics. This is in itself something of a meta-representational failure, though an understandable one, as physics students are not often asked to engage in such meta-representational tasks.

This argument regarding superior learning of representational skills would perhaps be stronger if absolute student performance on the quizzes and homeworks was better in the reform 202 course than the traditional. However, we must note that for the sake of replication, the exact questions that were designed for the traditional 202 course were given to the students in the reform 202 course. These courses placed different emphases on different particular subtopics, and so we would not necessarily expect students to equally well on any specific task. Relative student performance on the different representations and across choice and control groups should not have this problem, and so we have focused our attention mostly on relative performances.

We do not believe that students in the reform sections were necessarily learning better *meta*-representional skills. It is quite conceivable that these students were no better than the traditional 202 course at assessing their own abilities and evaluating the different representations available to them, and that their broader set of representational skills made any meta-representational failures less significant. Of course, neither do the data allow us to conclude that the reform 202 students were *not* learning better meta-representational skills. To our knowledge, no one has yet undertaken the research task of measuring meta-representational competence, though some work has characterized student meta-representational skills[9][11].

In our first paper on this topic[1] we suggested that a complete analysis would require attention to context at different levels. The first paper focused primarily on small-scale, specific context, demonstrating how particular problem features, topics, and representations could affect student performance. In this paper we have taken an explicit look at the effect of instructional environment, which involves a broader level of context as well as that broad level's influence on the specific context. At the small-scale, instruction changes what resources and skills students bring with them when solving a particular problem. Here, we have inferred that the reform 202 students brought with them a more complete set of representational skills, making them less vulnerable to variations in problem representation. At the broader level, we find the norms, expectations, and attitudes associated with a course. The course analysis presented here demonstrates that the reform course explicitly valued the use of a variety of representations, often together. Such multiple representation use

was demonstrated for the students on a regular basis in the lectures, and they were held accountable for such use in their own work, especially on the exams. It is quite likely that this norm helped drive the development of students' representational skills, tying together the different levels discussed here.

V. CONCLUSION

Our results suggest that instructional environment can play a significant role in developing student representational skills as they apply to problem solving. Pervasive use of different representations and of multiple representations appears to have broadened students' representational skills. Unfortunately, without assessment tasks that are less sensitive to the topic coverage, we are limited in our ability to conclude that student representational skills are stronger in a reform-style class in an absolute sense; all we are comfortable in asserting is that student problem solving skills varied significantly less from representation to representation. However, we consider it plausible that in becoming broader these students' skills have also become more developed overall. This suggests that if instructors wish to make increased representational facility a primary goal of their course, they can do so effectively so long as their efforts span all (or at least most) aspects of a course.

Acknowledgments

This work was supported in part by an NSF Graduate Fellowship and by the LEAP program. Special thanks to the rest of the Physics Education Research group at the University of Colorado at Boulder. Thanks also to Noah Podolefsky and Drs. Beale, Munsat, and Peterson for their cooperation and aid.

---

[1] P. B. Kohl and N. D. Finkelstein. Student representational competence, self-assessment, and problem solving in physics. *Physical Review Special Topic: Physics Education Research*, submitted.

[2] D. E. Meltzer. Relation between students' problem-solving performance and representational mode. *Am. J. Phys.*, 73:463, 2005.

[3] R. J. Beichner. Testing student interpretation of kinematics graphs. *Am. J. Phys.*, 62:750, 1994.

[4] F. M. Goldberg and J. H. Anderson. Student difficulties with graphical representations of negative values of velocity. *The Phys. Teach.*, 27:254, 1989.

[5] L. C. McDermott, M. L. Rosenquist, and E. H. Van Zee. Student difficulties in connecting graphs and physics: Examples from kinematics. *Am. J. Phys.*, 55:503, 1987.

[6] A. Van Heuvelen and X. Zou. Multiple representations of work-energy processes. *Am. J. Phys.*, 69:184, 2001.

[7] A. Van Heuvelen. Learning to think like a physicist: A review of research-based instructional strategies. *Am. J. Phys.*, 59:891, 1991.

[8] R. J. Dufresne, W. J. Gerace, and W. J. Leonard. Solving physics problems with multiple representations. *The Phys. Teach.*, 35:270, 1997.

[9] A. A. diSessa and B. L. Sherin. Meta-representation: an introduction. *Journal of Mathematical Behavior*, 19:385, 2000.

[10] A. A. diSessa, D. Hammer, B. L. Sherin, and T. Kolpakowski. Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, 10:117, 1991.

[11] A. A. diSessa. Students' criteria for representational adequacy. In K. Gravemeijer, R. Lehrer, B. van Oers, and L. Vershaffel, editors, *Symbolizing, Modeling and Tool Use in Mathematics Education*. 2003.

[12] K. R. Koedinger and M. J. Nathan. The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences*, 13:129, 2004.

[13] E. Mazur. *Peer Instruction: A user's manual*. Prentice Hall, Upper Saddle River, NJ, 1996.

[14] http://www.h-itt.com/.

[15] http://www.lon-capa.org.