Impact of the second semester University Modeling Instruction course on students' representation choices

Daryl McPadden[1,2] and Eric Brewe[3,4]

[1]Department of Physics, Florida International University, 11200 SW 8th St., Miami, Florida 33199, USA
[2]Department of Physics and Astronomy, Michigan State University, 567 Wilson Rd, East Lansing, Michigan 48824, USA
[3]Department of Physics, Drexel University, 32 S. 32nd Street Philadelphia, Pennsylvania 19104, USA
[4]School of Education, Drexel University, 3141 Chestnut St, Philadelphia, Pennsylvania 19104, USA

(Received 25 May 2017; published 15 November 2017)

Representation use is a critical skill for learning, problem solving, and communicating in science, especially in physics where multiple representations often scaffold the understanding of a phenomenon. University Modeling Instruction, which is an active-learning, research-based introductory physics curriculum centered on students' use of scientific models, has made representation use a primary learning goal with explicit class time devoted to introducing and coordinating representations as part of the model building process. However, because of the semester break, the second semester course, Modeling Instruction-Electricity and Magnetism (MI-EM), contains a mixture of students who are returning from the Modeling Instruction-mechanics course (to whom we refer to as "returning students") and students who are new to Modeling Instruction with the MI-EM course (to whom we refer to as "new students"). In this study, we analyze the impact of MI-EM on students' representation choices across the introductory physics content for these different groups of students by examining both what individual representations students choose and their average number of representations on a modified card-sort survey with a variety of mechanics and EM questions. Using Wilcoxon-signed-rank tests, Wilcoxon-Mann-Whitney tests, Cliff's delta effect sizes, and box plots, we compare students' representation choices from pre- to postsemester, from new and returning students, and from mechanics and EM content. We find that there is a significant difference between returning and new students' representation choices, which serves as a baseline comparison between Modeling Instruction and traditional lecture-based physics classes. We also find that returning students maintain a high representation use across the MI-EM semester, while new students see significant growth in their representation use regardless of content.

DOI: 10.1103/PhysRevPhysEducRes.13.020129

I. INTRODUCTION

Representations (including word descriptions, equations, pictures, diagrams, etc.) form the foundation of communication and problem solving in science. Physics is an especially representation-rich subject, with multiple representations often used or required for a single problem. By casually flipping through any introductory physics textbook, you can easily see a scattering of equations, force diagrams, pictures, and graphs within only a few pages [1,2].

For students, these representations serve as tools to help them learn, understand a problem or phenomena, and evaluate their results by serving a variety of purposes [3,4]. Ainsworth has categorized the multiple purposes into three classes: representations that serve complementary roles, representations that constrain interpretation, and representations that help construct deeper understanding [5]. As an example of representations being used in a complementary role, students can often find several pathways to a solution by using multiple representations, like using a graph rather than an equation or table. Using different representations can be especially productive if there are multiple tasks or if students prefer one representation over another [5] since each of these representations (i.e., graph, table, equations) of a data set highlight different salient features of the situation that may not be obvious from the others. Multiple representations can also aid students by dividing the information and thus reducing cognitive load [6,7]. By offering more than one means to the correct answer and highlighting different or overlapping information, multiple representations have been shown to increase students' ability to solve problems [7–9]. In contrast, constraining representations serve a different purpose; their role is to use familiar tools to help students form and understand a more abstract representation, often through an inherent shared property of the representations [5,10]. This function of multiple representations is to explicitly coordinate between representations and can serve as an

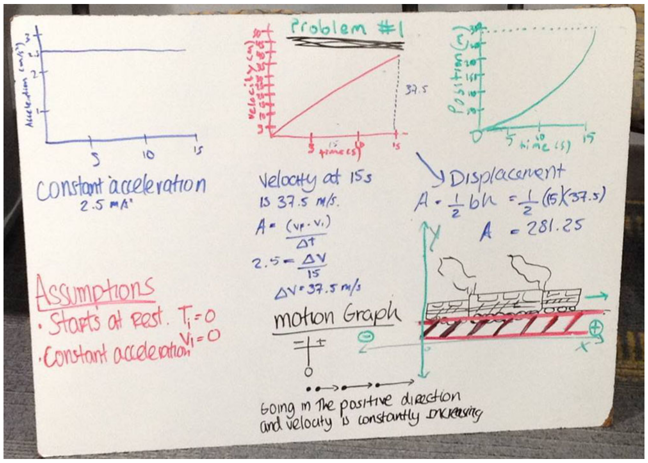

FIG. 1. This is an example of a conceptual model for an accelerating train created by students in the Modeling Instruction–mechanics course. It shows how a conceptual model consists of a coordinated set of representations including: equations, position-velocity-acceleration graphs, a picture, a motion map, an explicit statement of assumptions, and a word description.

evaluative check of an answer. Furthermore, representations often provide concrete, intermediate steps to help students reason abstractly about a problem and move from reading a problem statement to calculating an answer [5,11]. In physics, many representations such as free body diagrams or energy pie charts [12] have also been designed to help students conceptually reason about the problem and construct a deeper understanding of the physical phenomenon. By asking students to create a qualitative representation that then builds into quantitative representations, the intent is to help students ground their calculations and solutions in physical meaning. Given these advantages for students, many educators and researchers in physics and STEM more broadly have recommended using multiple representations in the physics classroom to help students learn [3,13–19].

Within physics and other STEM disciplines, there is extensive research on how students use individual representations. To name a few examples, this includes research on graphs [20–23], equations [24–26], force diagrams [15,27], pictures [28], and energy diagrams [18,29,30]. In comparison, research on students using multiple representations is much more sparse. Rosengrant, Etkina, and Van Heuvelen provide a summary of research on multiple representations in physics education research [4], which indicates that using multiple representations improves students conceptual understanding and ability to solve problems [31,32]. Traxler *et al.* and De Leone and Gire have shown that students in reformed physics classes or inquiry-based classes use more representations than their lecture counterparts [9,33]. De Leone and Gire go beyond that to say that the students who used representations in addition to mathematical representations tended to be more successful in solving the problem [9].

That being said, representation use is a skill that students develop [6,34–36]. Kohl and Finkelstein show that while both expert and novice physics problem solvers use multiple representations, they use them in distinctly different ways: experts are able to move between representations more quickly and are more flexible in their starting point, whereas novices use and jump between more representations in total [37]. If students are not given sufficient time, scaffolding, and practice with representations, they can cause confusion and be ineffective for students. Bakker and Derry succinctly state that "if students ostensibly learn knowledge in the absence of reasoning with representations *that make sense to them*, then this knowledge is likely to become inert and hard to transfer to new situations" (emphasis added) [38]. For example, Heckler showed that students actually did better on certain force problems when they were not prompted to use force diagrams, suggesting students are viewing the representation as a separate task that is unrelated to solving the problem [27]. Building on Heckler's work, Kuo, Hallinen, and Conlin found that prompting for specific diagrams prompts students to follow standard problem solving procedures, rather than deep conceptual reasoning about the problem [39]. This is further supported by studies that show that how a question is formulated can dictate what representations and solution path students take [40–42]. These studies indicate that students do not naturally use the representations as their instructors intend. Therefore, it is important that we teach students how and when to use representations in order for them to use multiple representations effectively.

A. Modeling Instruction

Because representations fill such a critical role in the process of doing science, Modeling Instruction has incorporated explicit class time and instruction on the representations into the curriculum, with their understanding and use as one of the primary learning goals for students in the class. University Modeling Instruction is a research-based, active-learning curriculum that has been developed and refined over the past ten years at Florida International University (FIU) [43] for introductory, calculus-based physics classes, building on the high school Modeling Instruction materials and the modeling theory of instruction [14,17,44,45]. At the university level, there are currently curricula developed for the first and second semester of introductory physics: Modeling Instruction-Mechanics (MI-Mech) and Modeling Instruction-Electricity and Magnetism (MI-EM).

Both of the Modeling Instruction courses are highly reformed, inquiry-based physics classes that use guided activities, experiments, and discussions to help students build a physical understanding of the world around them with conceptual models (as opposed to a physical or replica models). We have defined these conceptual models to consist of a coordinated set of representations (graphs, equations, diagrams, etc.) with the ability to explain, predict, and describe a set of physical phenomena [17,46,47]. For example, when asked to create a model of an accelerating train, we would expect a solution that looks like the one shown in Fig. 1, which contains not only equations, but a picture, word description, motion map, explicit statement about their assumptions, and position-velocity-acceleration graphs.

To create their models in class, students work in groups of three to collect data, solve problems, or derive results; they then summarize their findings, solutions, and remaining questions on a shared whiteboard. Once all of the groups have whiteboards, they come together for a large group discussion (24–30 students) or "board meeting" where they discuss what they found, ask questions to their peers, critique each other's work, and come to consensus about what they learned from the activity. Rather than lecturing, the role of the instructor in these classes is to plan the activities, guide the discussion, and facilitate student work. At Florida International University, MI-Mech and MI-EM have been implemented in 30 and 75 student formats and have been shown to be successful on multiple measures [48–52].

For each semester of Modeling Instruction, the curriculum is centered around 4–5 general conceptual models based on a class of phenomena, rather than 10–12 textbook chapters [17]. As described by Brewe, the course generally follows a model building cycle, which repeats approximately every two weeks: introduction and representation of a phenomenon, coordination of representations, application of the model to a variety of contexts, abstraction and generalization of the model, and, finally, refinement of the model [17]. This means that in class, Modeling Instruction students are using multiple representations early and often, adding representations over the course of the semester as they become needed and applicable to the model that students are building. While traditional physics classes (lecture, lab, recitation courses) certainly use, require, and teach representations as a part of the curriculum, it is rarely an explicit learning goal for the course or given specific instruction. Often a detailed discourse about how to use the representation, its components, its limitations, and how the representation relates to previous topics is skipped or assumed to be covered in a prerequisite course.

As introductory physics is a two semester sequence at FIU, there is a significant mixing of students that occurs at the semester break between MI-Mech and MI-EM. Not all of the students who take the MI-Mech course are able to take the MI-EM (e.g., due to course or work conflicts, major requirements, or class size limitations). Thus, the MI-EM course is comprised of a mixture of students: students who are returning to Modeling Instruction from MI-Mech with a consistent exposure to multiple representations through extensive practice using and talking about the representations in their models and students who are new to the Modeling Instruction environment—meaning they most likely took a lecture-based mechanics course and have not had this kind of experience with representations.

Because the Modeling Instruction environment heavily emphasizes and relies on multiple representations, we are particularly interested in studying how students engage with representations in the Modeling Instruction courses. In this paper, we begin to address this by examining the number and variety of representations that Modeling Instruction students would choose in mechanics and EM contexts, the impact of the second semester (MI-EM) on students' representation choices, and how students' familiarity with the Modeling Instruction class (whether they are new or returning to modeling in the second semester) impacts their representation use.

B. Research questions

This paper will address three primary research questions: (i) What are the differences between new and returning students' representation choices, particularly in how many and which representations they choose? (ii) How do students' representation choices vary based on physics content (mechanics or EM questions)? (iii) How does the MI-EM course impact both new and returning students' representation choices over the course of the semester?

II. METHODS

A. Problem Solving and Representation Use Survey (PSRUS)

To answer these research questions, we designed a modified card sort survey about students' choice of representations on various physics problems, which we called the Problem Solving and Representation Use Survey (PSRUS). The survey is 25 questions, with 13 questions covering EM topics and 12 questions covering mechanics topics. In addition to the questions, students were given a list of 15 representations (plus a write-in "other" option) covered in the Modeling Instruction class. The listed representations (shown in Fig. 2) were graphs, equations, words, assumptions, motion map, picture, system schema, force diagram, momentum vectors, energy pie charts, energy bar charts, circuit diagram, right hand rules (RHR), potential lines, and field lines. (The entire PSRUS is provided in the Supplemental Material [53] for reference.)

For each question, students were asked to choose what representations they would use if they were going to solve that problem. For example, one problem asks "A frisbee is thrown straight up in the air with an initial speed of 3 m/s. Find the maximum height the frisbee reaches." Students could respond with any number of the representations that they would use on the problem. For instance, a student may choose equations, a system schema, a picture, and energy pie charts to solve this problem or any other combination of the 16 listed representations. Note that we did not ask

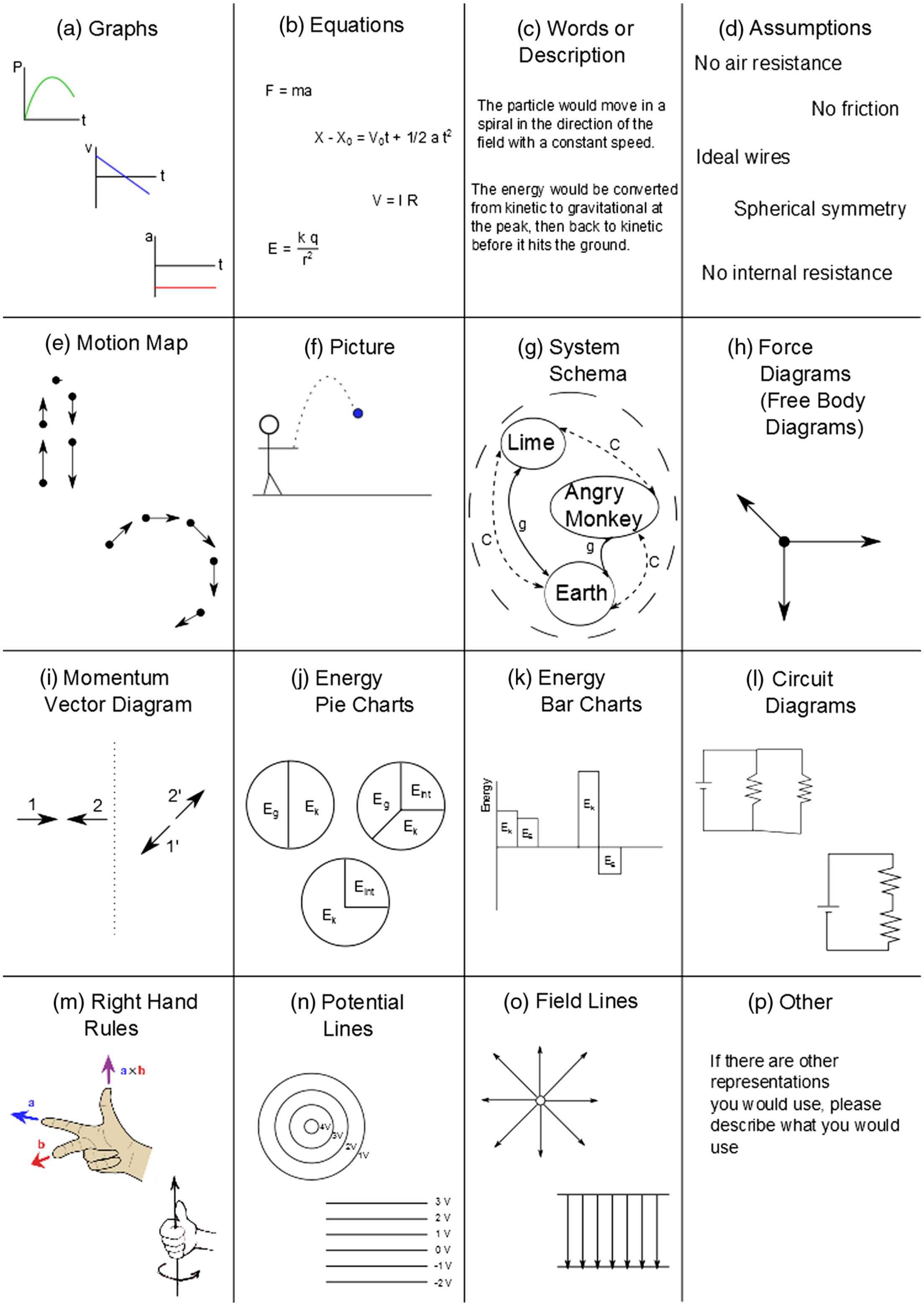

FIG. 2. The list of representations that were given to students along with the Problem Solving and Representation Use Survey, including: (a) graphs, (b) equations, (c) words or description, (d) assumptions, (e) motion map, (f) picture, (g) system schema, (h) force diagrams (or free body diagrams), (i) momentum vector diagram, (j) energy pie charts, (k) energy bar charts, (l) circuit diagrams, (m) right hand rules, (n) potential lines, (o) field lines, and (p) other.

students to solve the 25 problems, only to list the representations they would use if they were.

We chose the listed representations for the survey by identifying the representations explicitly covered in the Modeling Instruction courses (both mechanics and EM) through examination of the curricular materials, instructor's guides, and conversations with MI-Mech and MI-EM instructors about their expectations of what should be contained in a student's solution. In addition, we intentionally included an *other* option so students could add their own answers to the listed representations. The questions we designed for the PSRUS were meant to cover a broad range of content covered in the Modeling Instruction courses from both mechanics and EM, with each question open to the possibility of using multiple representations. We also designed many parallel questions between the Mechanics and EM questions (i.e., a "Find the Newton's third law pairs" question in both mechanics and EM contexts—questions 4 and 20 in the Supplemental Material [53]). In contrast to most Modeling Instruction problems that ask students to "make a model," we asked students to find a specific quantity in each of the questions so as to not artificially lead students to listing all of the representations.

After the PSRUS was developed, we administered the survey to a test panel, which consisted of faculty, graduate, and undergraduate students, and asked for their feedback. In order to address content and face validity in our survey [54], we asked faculty and graduate students to judge the appropriateness of the questions for introductory physics students and whether the questions were appropriate for measuring a broad range of students' representation use. Substantive validity, which asks if the intended audience is able to read, respond, and engage with the instrument in the intended manner [54], was addressed by administering the survey to undergraduate students and asking for their feedback, particularly on the wording of the questions and instructions and on the time required for survey. In addition, our panel consisted of people who had a wide range of exposure to Modeling Instruction: there were developers of Modeling Instruction, students and instructors who had previously taken or taught Modeling Instruction courses, and students and instructors with no exposure to the Modeling Instruction courses. So this suggests that the PSRUS is interpretable and substantively valid [54] for people who have been in Modeling Instruction before and those who haven't. Furthermore, our test panel was diverse, including representation from multiple genders, multiple ethnic and racial groups, and bilingual or non-native English speakers. Since the survey was given at a large, Hispanic-serving institution, it was particularly important that the survey be fair and understandable by members of these populations [54].

The feedback from this test panel was used to make substantial edits to the PSRUS. For instance, we cut an "explanation" section from each question due to the length of time required for the survey, and we changed the wording of the instructions to clarify our intent that students were not supposed to solve the problems. In addition, we made changes to wording and questions based largely on the faculty and graduate student feedback about whether the problems were appropriate.

Finally, we created two forms of the survey, which were given out randomly in the class—one which ordered the EM questions first and one which ordered the mechanic questions first. When we compared student responses from each version of the test, we found no differences between the two versions for all questions combined, mechanics questions only, or EM questions only; thus, PSRUS satisfies alternate-form reliability [54].

Note that the purpose of this survey was to determine what representations students in a Modeling Instruction course thought would be useful in their problem solving and what tools they tended to rely on. For this reason, we did not in any sense try to grade this survey or look for "correct" answers. Though there are some cases, for instance, using circuit diagrams on the frisbee problem, that would indicate students not taking the survey seriously. In addition, we did not attempt expert-novice comparisons on these data because it has been shown that steps that are useful for experts are not the same as those for novices [37,55]. Furthermore, it has been pointed out that expert-novice comparisons as a methodology can perpetuate inequity in the field [56].

B. Data collection

We gave the PSRUS survey pre- and postinstruction to the MI-EM class in the Spring 2015 semester at a large, Hispanic-serving, R1 institution. The survey was administered in class, but there was no grade or extra credit given for completion. A total of 69 students took the presurvey, 60 students took the postsurvey, and 58 of those students were matched pre-post (which are the data used in the following analyses). Of the 58 matched students in the MI-EM course, 30 students were returning from the MI-Mech course and 28 students were new to the Modeling Instruction environment (meaning they took the lecture-based mechanics course at the same institution or had transfer credit for their mechanics physics requirement). These groups of students are referred to as the returning and new students, respectively.

C. Data analysis

1. Distributions

From the survey data, we counted the number of representations that each student chose on all of the questions divided by the total number of representations they could have chosen. We report this as a fraction rather than a pure number because there is a differing number of questions for the mechanics and EM contexts. This gives a

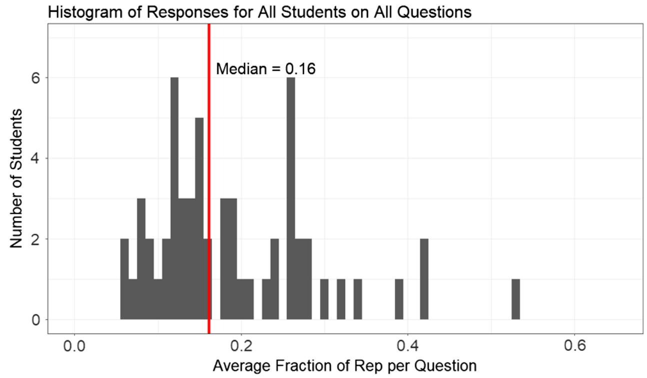

FIG. 3. The distribution shows the average fraction of representations chosen on all questions combined by all the students in the MI-EM class on the Pre-PSRUS. The median is marked at 0.16 or between 2 and 3 representations per question.

single number for each student that represents the average fraction of representations chosen per question. Again, this fraction is the fraction of the *total* listed representations (the 16 shown in Fig. 2), not a measure of "correctness." Based on this fraction, the distribution for a group of students can then be made. An example is shown in Fig. 3 showing the distribution for an average fraction of representations used on all the questions (both mechanics and EM) in the pre-PSRUS by all students (both new and returning).

This data set is not normally distributed, with a floor on the left side of the distribution and a right-skewed tail. Based on the format of the PSRUS, this is expected. We expect students to respond with at least one representation on each question, which accounts for the floor of 0.06. We would also expect that while most of the questions lend themselves to multiple representation, not all of the representations are applicable to every question, thus accounting for the right-skew tail. This type of distribution has implications for the statistical tests and averages that we report. Since the distribution is not (and would not be expected to be) normal, we used nonparametric statistical tests and effect sizes, which are expanded upon in the following sections. We also report median values as the measure of center for these distributions because the mean is more drastically affected by the right-skewed tail.

In a similar manner, we created distributions that were separated by student group (returning or new) and by content area (mechanics, EM, or all) for both the pre- and post-PSRUS data. As expected, all of these distributions held the same shape as that shown in Fig. 3: non-normal, left floor, and right-skewed tail.

2. Wilcoxon-signed-rank and Wilcoxon-Mann-Whitney tests

To compare these nonparametric distributions, we used Wilcoxon-signed-rank (WSR) tests and Wilcoxon-Mann-Whitney (WMW) tests, which are the respective nonparametric equivalents to paired and unpaired t tests that require no assumptions about the shape of distribution. We used paired WSR comparisons for all pre vs post and mech vs EM comparisons since the same students comprised these populations, whereas unpaired WMW was used for all new vs returning comparisons as these are independent populations. Both WSR and WMW tests are based on a rank-sum procedure and thus are able to distinguish between shape and placement differences of distributions; however, when the distributions have the same shape, the WMW and WSR tests can be interpreted as a difference in medians [57–59]. Since all of our distributions have the same general shape, we will interpret differences by the WSR and WMW tests as differences in the medians of the distributions.

Given that we ran multiple comparisons (9 pre vs post, 6 new vs returning, 6 mech vs EM), we used an alpha level for each test of 0.002 38, based on the Bonferroni correction of

$$\alpha_{B1} = \frac{0.05}{21} = 0.002\,38, \qquad (1)$$

which gives a 95% confidence in the overall study based on 21 individual tests. The Bonferroni correction is often cited as the most conservative correction for multiple tests [60]; however, it is the best correction for both family-wise error rates and per-family error rates on type 1 error [61].

3. Cliff's delta

To compare individual representations (rather than an average fraction as was done in the WSR and WMW tests), we used Cliff's delta as a nonparametric effect size measure, which makes no assumptions about the shape of the distributions. Cliff's delta is calculated from comparing the number of data points from distribution X that are larger than the number of data points from distribution Y divided by number of data points in each distribution as shown in Eq. (2), with a correction if the distributions are paired [62].

$$\delta = \frac{\#(X_i > Y_j) - \#(X_i < Y_j)}{n_x \times n_y}. \qquad (2)$$

In contrast to Cohen's d which has no upper limit, the delta effect size is limited to be between -1 and 1. Cohen's interpretation of small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$) effects [63] then correspond to delta valves of $\delta = 0.147$, $\delta = 0.33$, and $\delta = 0.474$, respectively [64].

Since we used Cliff's delta to compare the 16 representations, which each had their own effect size, we again used a Bonferroni correction to calculate the confidence intervals for each effect as shown in Eq. (3):

$$\alpha_{B2} = \frac{0.05}{16} = 0.003\,125. \qquad (3)$$

Thus, the reported confidence intervals on each effect size are 99.7% confidence intervals so that the overall comparison of the set of 16 representations has a confidence of 95%.

4. Box plot comparisons

In cases where the Cliff's delta comparisons are not meaningful or there is complete overlap in confidence intervals, we use box plots to compare differences in the 16 representations. The box plots are able to show differences not only in median but also in the shape of the distribution, which is lost information in the effect size analysis. However, in using box plots, we lose the statistical significance of the comparison. While we could use WMW tests to compare the 16 representations, the shape of the distributions are no longer the same when comparing each representation (rather than all representations combined). This means that a significant difference by the WMW could be due to a difference in the shape of the distributions or a difference in the placement of the distributions, with no way to tell which is the significant factor. Thus, we opt to present box plots without statistical significance in these cases because it best presents the information of interest, but these should not be used as an extensive, generalizable comparison.

III. RESULTS AND DISCUSSION

A summary of the average representation choice per question by content and by group of students is shown in Fig. 4 for the pre- and post-PSRUS. The average percentage of representations per question chosen ranges from 13.4% to 30.5%, or somewhere between 2 and 5 representations per question. This indicates that students are neither selecting only one representation repeatedly nor selecting all the representations consistently, but are being thoughtful in their representation choices. The features and conclusions from this plot will be discussed in the subsequent sections.

A. Baseline—new vs returning

When comparing new vs returning students, we see significant differences ($p < \alpha_{B1}$ by WMW) between students' representation choices on AllPre, MechPre, and MechPost (all other new vs returning comparisons are not significant—shown in Table 3 in the Supplemental Material [53]), which is consistent with our previous results [65]. This tells us that students are entering the MI-EM course making different representation choices, primarily in mechanics content. While this is not entirely surprising given the differences in prior instruction and representation emphasis, the mechanics precomparison in particular provides a good baseline for differences between representations in Modeling Instruction and traditional lecture courses.

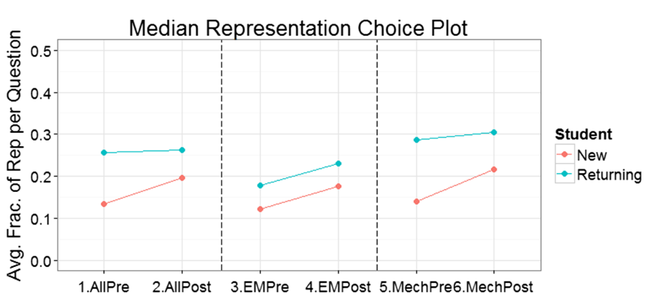

FIG. 4. The median fraction of representations that students chose per question separated by content area and group of students.

As is shown in Fig. 4, column 5, returning students on average use 28.7% of the available representations per question. This means that returning students are saying they would use 4–5 representations on each mechanics question. In contrast, new students say they would use 14.1% of the available representations on average, which is between 2 and 3 representations.

When we break it down by the individual representations, we see more differences between new and returning students. The Cliff's delta effect size for each representation on mechanics questions from the pre-PSRUS is shown in Fig. 5. In this comparison, a positive effect means that the representation is favored by returning students and a negative effect means that the representation is favored by new students.

From the effect sizes, we see that there are several large positive effects, meaning there are several representations that are highly favored by returning students over new students. These include some representations that we would expect to be favored by returning students, such as energy pie charts [12] and system schema [31], which were developed for and almost exclusively used in Modeling Instruction courses. However, there are also several representations including assumptions, graphs, and motion maps that we might expect to be universal (used by both lecture and Modeling Instruction course) that we see are favored by returning students. This means that, as a baseline, returning students use more representations because not only do they know more representations, but they see more traditional representations as more widely applicable.

We also see no difference between new and returning students' use of equations or force diagrams in predata. In the MI-Mech curriculum, two of the largest modifications from the typical mechanics content organization are that (i) the MI-Mech curriculum introduces energy before forces and (ii) there is a large focus in the course materials on conceptual understanding and use of representations before introducing equations [17]. The fact that we see no difference between new and returning students on equations and force diagrams indicates that these changes to the MI-Mech curriculum are not detrimental to students' understanding of when these representations are appropriate.

B. Returning students

When we separate students responses by whether they were new or returning to Modeling Instruction, we see two

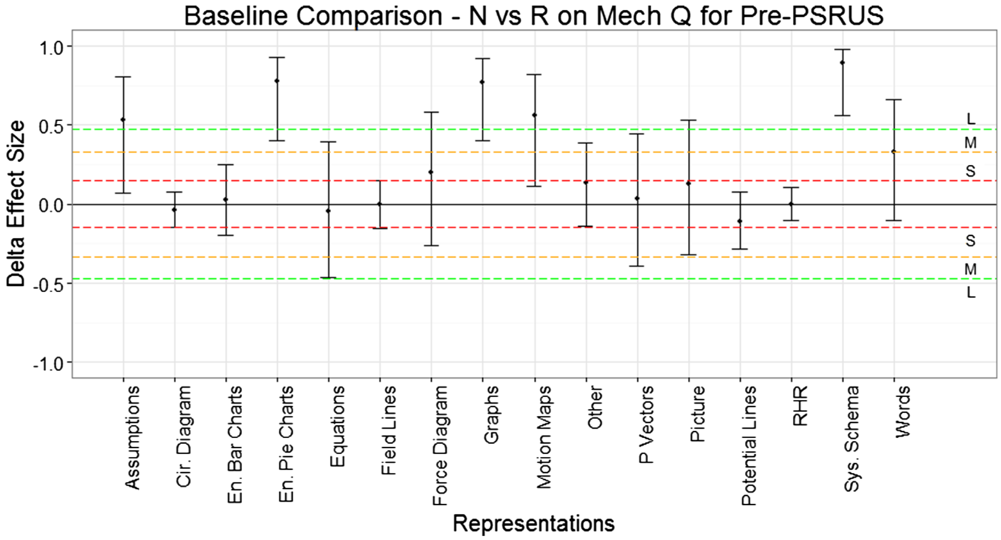

FIG. 5. Cliff's delta effect size for each representation comparing new and returning students on mechanics questions from the Pre-PSRUS with 99.7% confidence intervals (from α_{B2}). Positive effects show a representation that is favored by returning students; negative effects show those favored by new students.

distinct patterns in their representation choices. Starting with returning students, we see a high level of representation use that is maintained from pre to post in MI-EM, which is particularly evident in columns 1–2 (All questions combined) in Fig. 4 where there is virtually no change in the average number of representations that students choose per question. When we split the responses by content area, we do see an increase from pre to post in both EM (17.8%–23.1%) and mechanics (28.7%–30.5%) questions, but neither of these shifts are significant at the α_{B1} level by WSR tests (See Table 1 in Supplemental Material [53] for reference).

When examining the individual representations, we see similar results. The effect sizes for the representations chosen on mechanics questions are shown in Fig. 6. In the comparison of pre to post, a positive effect size means the representation was chosen more often in the postsurvey whereas a negative effect size means the representation was chosen more often in the presurvey.

From Fig. 6, we see that most of the representations have small or negligible effects, meaning there were no changes from pre to post in the use of these representations. We see medium, positive effects in energy pie charts, pictures, and system schema, which suggests that MI-EM may help students see the applicability of these representations in the mechanics content; however, we do not have the statistical power to conclusively say so. The one representation that is significant at the α_{B2} level is force diagrams, which has a large, positive effect. This tells us that over the second semester of Modeling Instruction, students are increasing their use of force diagrams on content that was not introduced that semester (though force diagrams were used).

The results of Fig. 6 are important because they tell us that returning students are maintaining, if not increasing, their representations on content from the previous semester.

Likewise, the effect sizes of returning students' representation choices on EM questions are shown in Fig. 7. Again, a positive effect means the representation was favored in the post-PSRUS, and a negative effect means the representation was favored in the pre-PSRUS. As before, many of the representations have negligible or small effects. The only representation with a significant effect at α_{B2} is the right hand rule, which is a huge, positive effect. Energy bar charts, energy pie charts, force diagrams, and system schema have medium, positive effects, which suggests that students have increased their use of these representations in the post survey; however, their confidence intervals are large enough to cross zero. Interestingly, circuit diagrams has a medium, negative effect (though not significant), which suggests that students thought that circuit diagrams would be more useful at the beginning of the semester than they did at the end.

Figure 7 tells us that while there is some growth in returning students use of representations on EM problems, they are maintaining the number of representations from pre to post. This could indicate that returning students are approaching new material with a "try all representations" strategy, even in unfamiliar contexts. However, while maintaining representation use is good on both mechanics and EM problems, it is concerning that we do not see more growth of representations in EM, particularly in field lines and potential lines, which are two of the new representations introduced in MI-EM. This result may indicate that more scaffolding is needed around these representations in the curriculum.

Furthermore, we see large differences between the number of representations that returning students use on mechanics questions compared to EM questions. On the presurvey, returning students choose on average 17.8% of

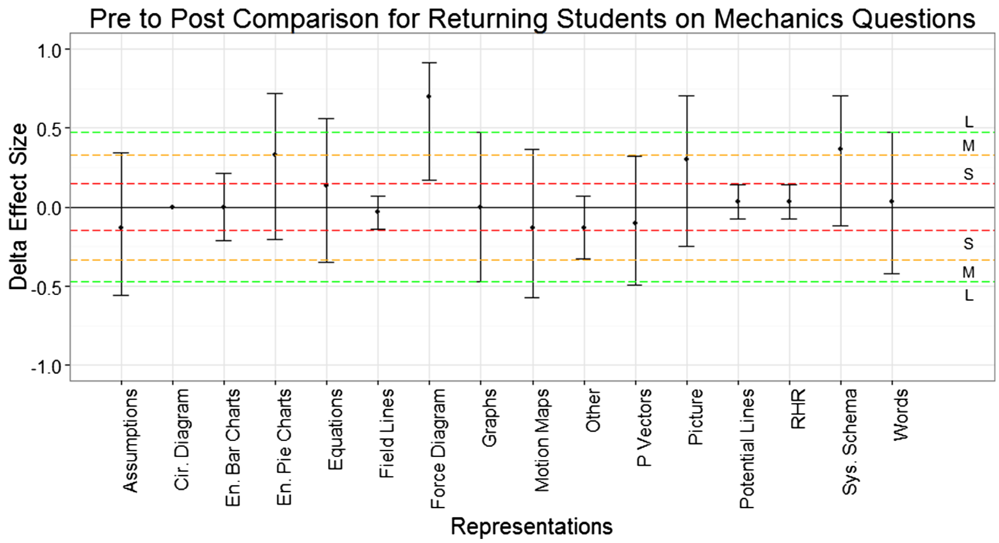

FIG. 6. Cliff's delta effect size for each representation comparing pre to post representation choices on mechanics questions from returning students with 99.7% confidence intervals (from α_{B2}). Positive effects show a representation that is favored in the post-PSRUS; negative effects show those favored in the pre-PSRUS.

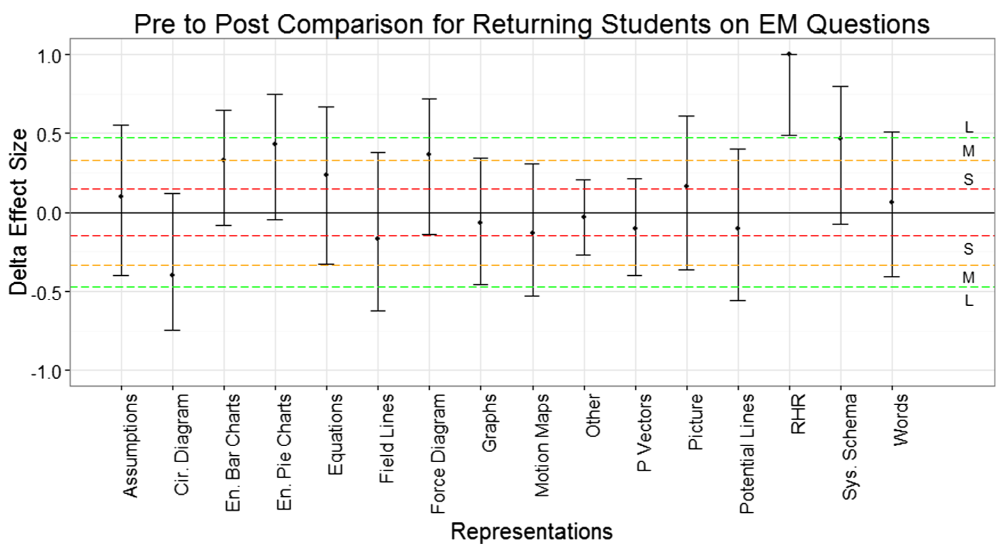

FIG. 7. Cliff's delta effect size for each representation comparing pre to post representation choices on EM questions from returning students with 99.7% confidence intervals (from α_{B2}). Positive effects show a representation that is favored in the post-PSRUS; negative effects show those favored in the pre-PSRUS.

the representations per question for EM and 28.7% for mechanics, which is a significant difference by WSR test at α_{B1}. This kind of difference at the beginning of the semester is not shocking, but the difference is persistent to the end of the MI-EM semester. Even on the postsurvey, returning students choose significantly fewer (by WSR at α_{B1}) representations on EM questions (23.1%) than on mechanics questions (30.5%). (See Table 2 in the Supplemental Material [53] for reference.)

When we compare the individual representations on mechanics and EM questions, there are clear distinctions between what returning students see as useful on mechanics problems and what they see as useful on EM problems. Figure 8 shows the effect size analysis for the returning students on the post-PSRUS, looking at differences in representation use between mechanics and EM questions. In this comparison, a positive effect means the representation was favored on EM questions, and a negative effect means it was favored on mechanics questions.

As would be expected, the representations that are favored on EM questions (large, significant, positive effects) are those that were introduced in the MI-EM course: circuit diagrams, energy bar charts, field lines, potential lines, and right hand rules. Words, pictures and assumptions have confidence intervals that cross zero, thus cannot be categorized as either mechanics or EM (though

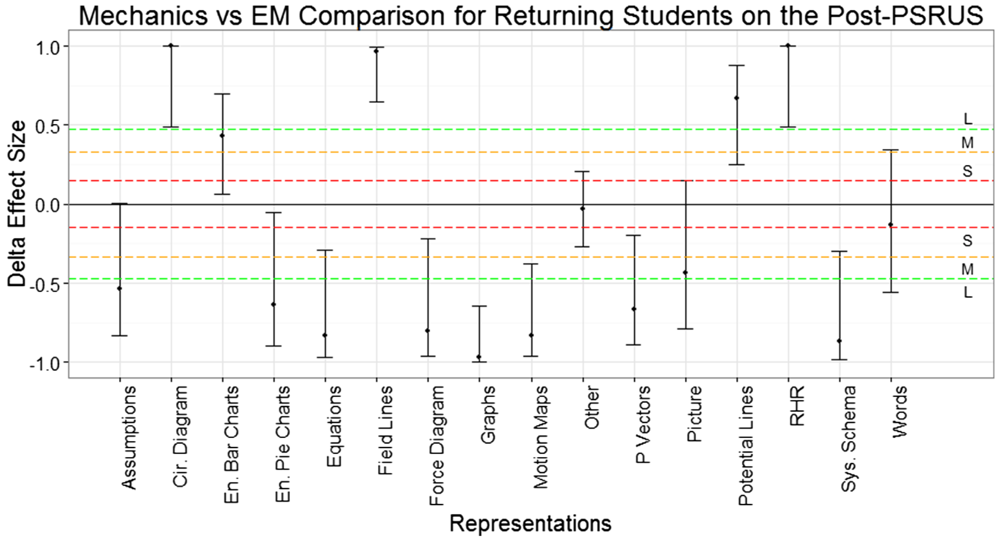

FIG. 8. Cliff's delta effect size for each representation comparing mechanics to EM representation choices on the post-PSRUS from returning students with 99.7% confidence intervals (from α_{B2}). Positive effects show a representation that is favored on EM questions; negative effects show those favored on mechanics questions.

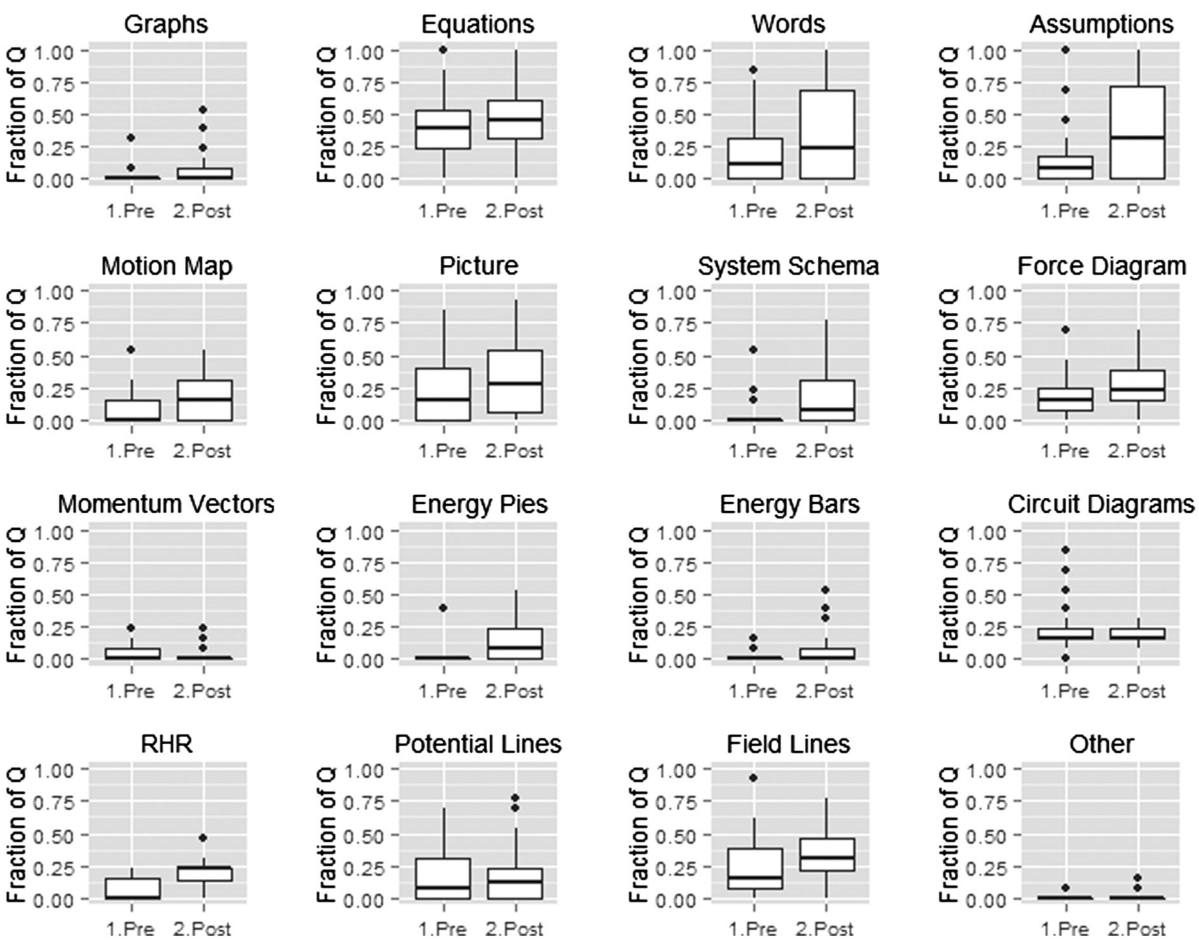

FIG. 9. Box plots showing the side-by-side comparison for the frequency of each representation in the pre- and post-PSRUS on EM questions for new students. The number of questions that students chose that representation is presented as a fraction of all EM questions on the vertical axis.

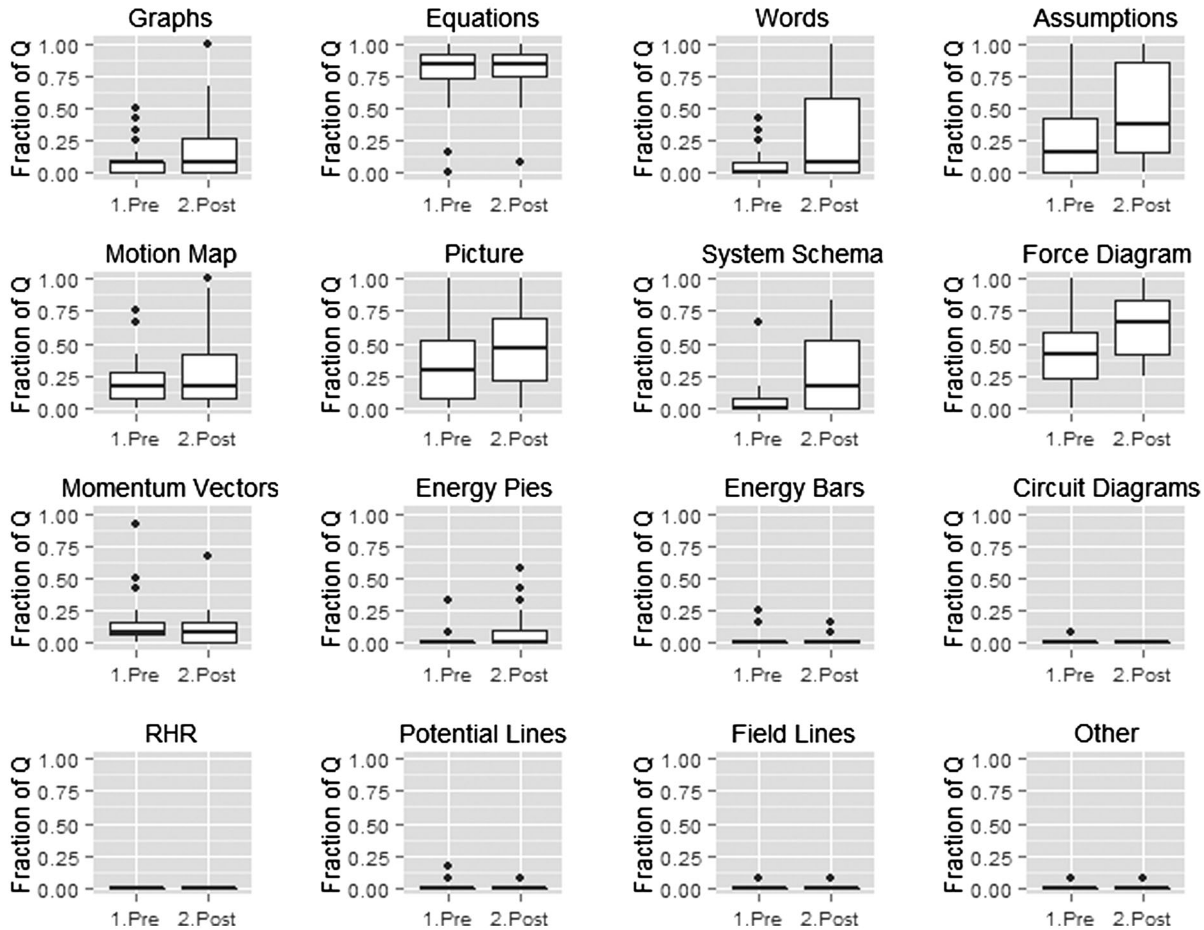

FIG. 10. Box plots showing the side-by-side comparison for the frequency of each representation in the pre- and post-PSRUS on mechanics questions for new students. The number of questions that students chose that representation is presented as a fraction of all mechanics questions on the vertical axis.

we do recognize that both assumptions and pictures have large, negative effect sizes). With the exception of the other category, the rest of the representations were highly favored on mechanics questions as evidenced by the large, significant, negative effects.

While some representations are certainly more useful in EM or mechanics contexts (i.e., circuit diagrams are generally exclusive to EM contexts), there are many representations that we wouldn't expect to be so polarized. For example, equations, graphs, and force diagrams have some of largest negative effects, but from an expert perspective are certainly useful tools in EM. Any introductory textbook in EM will present Maxwell's equations as the foundation of EM, will show force diagrams with example problems involving the electric or magnetic force and will use electric field vs position or magnetic flux vs time graphs to highlight the relevant relationships [1,2]. However, we see that the returning students in MI-EM are using these representations more heavily in mechanics. This indicates that they are not seeing the applicability of these representations in the new EM contexts.

C. New students

In contrast to the returning students, the new students in MI-EM do see significant growth in the number of representations that they use over the course of the semester. For all questions combined, new students, on average, increase their representation choices per question from 13.4% to 19.6%, which amounts to a gain of one representation per question. This increase is shown from column 1 to 2 in Fig. 4. When we split the questions by content, we see very similar patterns. EM representations increase from 12.3% to 17.6% and mechanics representations increase from 14.1% to 21.6% (shown in columns 3–4 and 5–6 in Fig. 4, respectively). All of these gains are significant by WSR at α_{B1} (see Table 1 in Supplemental Material [53] for reference).

In itself, this is an important result. In addition to impacting students use of representations in the EM content that is taught in the class, new students in MI-EM are increasing and backtracking representations to physics content that was not explicitly covered in the semester. This indicates that Modeling Instruction is not only teaching conceptual physics knowledge but is influencing the tools that students use to problem solve regardless of content. Furthermore, we see relatively even gains in representations between mechanics and EM questions for new students, unlike the returning students who favor representations in mechanics. By the WSR tests, there were no significant differences between EM and mechanics in either the pre- or post-PSRUS.

When we examine the individual representations that students chose on EM questions, we see that the new students are both gaining the new representations emphasized in MI-EM course and using familiar representations more frequently. The box plots in Fig. 9 summarize the changes for new students in each representation from pre to postsemester. Overall, we see increases in the representations introduced in MI-EM, particularly in RHR and field lines as evidenced by the large jump not only in the median but also in the inner quartile range (IQR). We also see that the representations that are unique to Modeling Instruction (energy pie charts and system schema) are used more frequently in the post. There is a large change in the spread of the IRQ for energy pie charts and system schema, accompanied by a increase in the median value. Beyond that we see large shifts in the spread of the distributions for the more familiar representations as evidenced by a larger IRQ in the postdata; this says that by the end of the semester new students are using words, assumptions, pictures, and force diagrams more frequently on EM problems.

By looking at how new students respond to the mechanics questions, we can see what representations students are backtracking and transferring out of the class content. These results are summarized in the box plots shown in Fig. 10. We do see new students using representations that they would have learned in the MI-EM course on mechanics questions—primarily the system schema, which shows a large spread in the distribution and an increase in the median for the post survey. However, most of the large increases in representations on the mechanics questions are coming from those that they are already familiar with. There are large overall shifts in both the median and IQR to the distributions of pictures, force diagrams, and assumptions as well as changes to the shape of the distributions in graphs, words, and momentum vectors as evidenced by the larger IQR's.

This suggests that MI-EM is helping new students recognize the applicability of the representations that they already know, which they are then able to apply to the new EM context as well as backtracking to the more familiar mechanics context.

IV. CONCLUSIONS

To answer research question (i), we compared new and returning students' representation choices on the PSRUS mechanics and EM questions. We found a significant difference on the mechanics presurvey questions that was persistent to the postsurvey. In particular, the difference between new and returning students on the mechanics presurvey is a good baseline for comparing how Modeling Instruction and traditional lecture courses teach representations. From the Cliff's delta analysis (in Fig. 5), we showed that not only do returning students use more novel tools like energy pie charts and system schema on mechanics questions, but they also use familiar representations like graphs, motion maps, and assumptions more frequently. We also see no differences in students' use of equations or force diagrams, despite a qualitative-first, energy-first curriculum. This result is consistent with Traxler, De Leone, and Gire who found that students from inquiry based physics classes use more representations than their lecture counterparts [9,33]; however, we have moved beyond their results by looking at which representations Modeling Instruction students use more and in which contexts.

To answer research question (ii) and (iii), we looked at how new and returning students representation choices changed from pre to post. We found that returning students maintain high levels of representation use, with the only significant growth of representations in right hand rules for EM questions and force diagrams for mechanics questions. This suggests that returning students have a high retention of representations and are not changing their views of when the representations are applicable. However, we do not see significant growth in representations on EM questions for returning students even though MI-EM curriculum introduces five new representations and builds on those from mechanics, nor do we see returning students reaching similar frequency of representation use that they do on mechanics questions (addressing research question 3). While we would not expect exactly the same representation use in mechanics and EM contexts, the fact that returning students heavily favor general representations like equations, graphs, and pictures in mechanics contexts is concerning (addressing research question 2). This indicates that the MI-EM curriculum is not integrating and emphasizing the representations in a consistent manner to the MI-Mech course.

In contrast, the new students in MI-EM see significant growth in their representation choices from the beginning to the end of the semester (addressing research question 3), regardless of mechanics or EM content (addressing research question 2). The fact that the new students see even growth between the content areas further supports the idea that there are differences in how representations are taught between MI-Mech and MI-EM. Even more importantly, we are seeing that the new students in MI-EM are backtracking what they have learned about representations to physics content that was not taught in the course. This suggest that MI-EM is impacting students' problem solving skills in addition to their content knowledge and providing them with more tools to use in that process. Furthermore, we see that the gains in representations come from both gaining new representations (i.e., energy pie charts and system schema) and using familiar representations more frequently (i.e., words, force diagrams, pictures) from Figs. 9 and 10. That being said, the new students never reach the level of representation use that the returning students have. On mechanics questions, there was a persistent, significant difference between number of representations that the new and returning students chose. On EM questions, we see that the number of representations the new students end with is about the same as the number of representations that the returning students start with. This indicates that the MI-EM curriculum is not doing a good job of catching the new students up to speed with the returning students.

These results suggest that MI-EM is meeting its goal of increasing the number and variety of representations that they use in physics. Its students have access to more representations and use more traditional representations in physics more frequently. This result is consistent with what other researchers have found for reformed, inquiry based physics courses [9,33]. Regardless of the content being taught, Modeling Instruction is impacting how students problem solve and the tools that they rely on to do so. However, we also see significant differences between new and returning students' representation choices, which indicates that these two groups of students are having different experiences in the class. From these results, curricular changes to MI-EM are indicated, particularly to incorporate representations consistently across the two semesters and to transition the new students into Modeling Instruction more effectively.

This study also has wider implications for physics instructors and researchers. Particularly for reformed classes, we need to pay attention to how much we are building on students' experiences in previous courses. This is especially true of the second course in the sequence, like MI-EM, which may require more introductory scaffolding than is given. Furthermore, this study exemplifies a course evaluation based on a noncontent related learning goal. While multiple representation use is a core tenant of Modeling Instruction, previous measures of success would not have been able to address whether students found these representations applicable. We encourage instructors and curriculum developers to assess their classes on learning goals that move beyond content knowledge.

ACKNOWLEDGMENTS

We thank the PER group at FIU for their input on this project, particularly Dr. Geoff Potvin. Dr. Rachel Scherr, Dr. Jayson Nissen, and the PER group at University of Colorado-Boulder also provided valuable feedback on this study. This work was supported by NSF Grants No. DUE 1140706 and No. DGE 1038321.

[1] R. W. Chabay and B. A. Sherwood, *Matter and Interactions*, 3rd ed., edited by S. Johnson (John Wiley & Sons, Inc., Hoboken, 2011).

[2] H. D. Young, R. A. Freedman, A. L. Ford, and F. W. Sears, *University Physics: with Modern Physics*, 14th ed. (Pearson Addison Wesley, San Francisco, 2004).

[3] J. J. Kaput, Linking representations in the symbol systems of algebra, in *Research Issues in the Learning and Teaching of Algebra*, edited by S. Wagner and C. Kieran (Macmillan, New York, 1989), pp. 167–194.

[4] D. Rosengrant, E. Etkina, and A. Van Heuvelen, An overview of recent research on multiple representations, AIP Conf. Proc. **883**, 149 (2007).

[5] S. Ainsworth, The functions of multiple representations, Comput. Educ. **33**, 131 (1999).

[6] T. Fredlund, C. Linder, J. Airey, and A. Linder, Unpacking physics representations: Towards an appreciation of disciplinary affordance, Phys. Rev. ST Phys. Educ. Res. **10**, 020129 (2014).

[7] R. Cox and P. Brna, Supporting the use of external representations in problem solving: The need for flexible learning environments, J. Interact. Learn. Res. **6**, 239 (1995).

[8] H. J. M. Tabachneck, K. R. Koedinger, and M. J. Nathan, Towards a theoretical account of strategy use and sense making in mathematical problem solving, in *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, edited by A. Ram and K. Eiselt (Erlbaum, Hillsdale, 1994), pp. 836–847.

[9] C. J. De Leone and E. Gire, Is instructional emphasis on the use of non-mathematical representations worth the effort?, AIP Conf. Proc. **818**, 45 (2005).

[10] S. Ainsworth, D. Wood, and C. O Malley, There is more than one way to solve a problem: evaluating a learning environment that supports the development of children's multiplication skills, Learn. Instr. **8**, 141 (1998).

[11] D. L Schwartz, The emergence of abstract representations in dyad problem solving, J. Learn. Sci. **4**, 321 (1995).

[12] E. Brewe, Energy as a substancelike quantity that flows: Theoretical considerations and pedagogical consequences, Phys. Rev. ST Phys. Educ. Res. **7** (2011).

[13] R. J. Dufresne, W. J. Gerace, and W. J. Leonard, Solving physics problems with multiple representations, Phys. Teach. **35**, 270 (1997).

[14] D. Hestenes, Modeling methodology for physics teachers, AIP Conf. Proc. **399**, 935 (1996).

[15] D. Rosengrant, A. Van Heuvelen, and E. Etkina, Free-body diagrams: Necessary or sufficient?, AIP Conf. Proc. **790**, 177 (2005).
[16] J. I. Heller and F. Reif, Prescribing effective human problem-solving processes: Problem Description in Physics, Cognit. Instr. **1**, 177 (1984).
[17] E. Brewe, Modeling theory applied: Modeling Instruction in introductory physics, Am. J. Phys. **76**, 1155 (2008).
[18] A. Van Heuvelen and X. Zou, Multiple representations of work-energy processes, Am. J. Phys. **69**, 184 (2001).
[19] M. Van Sommeren, P. Reimann, H. Boshuizen, and T. de Jong, *Learning with Multiple Representations* (Pergamon, Amsterdam, 1998).
[20] R. J Beichner, Testing student interpretation of kinematics graphs, Am. J. Phys. **62**, 750 (1994).
[21] 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).
[22] S. N. Friel, F. R. Curcio, and G. W. Bright, Making sense of graphs: Critical factors influencing comprehension and instructional implications, J. Res. Math. Educ. **32**, 124 (2001).
[23] D.-h. Nguyen and N. S. Rebello, Students' understanding and application of the area under the curve concept in physics problems, Phys. Rev. ST Phys. Educ. Res. **7**, 010112 (2011).
[24] B. L. Sherin, How students understand physics equations, Cognit. Instr. **19**, 479 (2001).
[25] T. J. Bing and E. F. Redish, Analyzing problem solving using math in physics: Epistemological framing via warrants, Phys. Rev. ST Phys. Educ. Res. **5**, 020108 (2009).
[26] D.-h. Nguyen and N. S. Rebello, Students' difficulties with integration in electricity, Phys. Rev. ST Phys. Educ. Res. **7**, 010113 (2011).
[27] A. F. Heckler, Some consequences of prompting novice physics students to construct force diagrams, Int. J. Sci. Educ. **32**, 1829 (2010).
[28] J. H. Larkin and H. A. Simon, Why a diagram is (sometimes) worth ten thousand words, Cogn. Sci. **11**, 65 (1987).
[29] R. E. Scherr, H. G. Close, E. W. Close, and S. Vokos, Representing energy. II. Energy tracking representations, Phys. Rev. ST Phys. Educ. Res. **8**, 020115 (2012).
[30] R. E. Scherr, H. G. Close, S. B. McKagan, and S. Vokos, Representing energy. I. Representing a substance ontology for energy, Phys. Rev. ST Phys. Educ. Res. **8**, 020114 (2012).
[31] B. E. Hinrichs, Using the system schema representational tool to promote student understanding of Newton's Third Law, AIP Conf. Proc. **790**, 117 (2004).
[32] N. D. Finkelstein, W. K. Adams, C. J. Keller, P. B. Kohl, K. K. Perkins, N. S. Podolefsky, S. Reid, and R. Lemaster, When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment, Phys. Rev. ST Phys. Educ. Res. **1**, 010103 (2005).
[33] A. L. Traxler, J. V. Mahadeo, D. McPadden, and E. Brewe, Multiple representations and epistemic games in introductory physics exam solutions, *Proceedings of the Physics Education Research Conference, 2014, Minneapolis, MN* (AIP, New York, 2014), pp. 247–250.
[34] M. Yerushalmy, Student perceptions of aspects of algebraic function using multiple representation software, J. Comput. Assist. Learn. **7**, 42 (1991).
[35] Y. Uesaka, E. Manalo, and S. Ichikawa, What kinds of perceptions and daily learning behaviors promote students' use of diagrams in mathematics problem solving?, Learn. Instr. **17**, 322 (2007).
[36] A. A. diSessa and B. L. Sherin, Meta-representation: An introduction, J .Math. Behav. **19**, 385 (2000).
[37] P. B. Kohl and N. D. Finkelstein, Patterns of multiple representation use by experts and novices during physics problem solving, Phys. Rev. ST Phys. Educ. Res. **4**, 010111 (2008).
[38] A. Bakker and J. Derry, Lessons from inferentialism for statistics education, Math. Thinking Learning **13**, 5 (2011).
[39] E. Kuo, N. R. Hallinen, and L. D. Conlin, When procedures discourage insight: epistemological consequences of prompting novice physics students to construct force diagrams, Int. J. Sci. Educ. **39**, 814 (2017).
[40] M. De Cock, Representation use and strategy choice in physics problem solving, Phys. Rev. ST Phys. Educ. Res. **8**, 020117 (2012).
[41] D. E. Meltzer, Relation between students' problem-solving performance and representational format, Am. J. Phys. **73**, 463 (2005).
[42] P. B. Kohl and N. D. Finkelstein, Student representational competence and self-assessment when solving physics problems, Phys. Rev. ST Phys. Educ. Res. **1**, 010104 (2005).
[43] University Modeling Instruction, http://univ-modelinginstruction.com/.
[44] D. Hestenes, Toward a modeling theory of physics instruction, Am. J. Phys. **55**, 440 (1987).
[45] I. A. Halloun, *Modeling Theory in Science Education*, 1st ed. (Springer, Netherlands, 2006).
[46] S. Manthey and E. Brewe, Toward university modeling instruction-biology: Adapting curricular frameworks from physics to biology, CBE Life Sci. Educ. **12**, 206 (2013).
[47] P. S. Oh and S. J. Oh, What Teachers of Science Need to Know about Models: An overview, Int. J. Sci. Educ. **33**, 1109 (2011).
[48] E. Brewe, V. Sawtelle, L. H. Kramer, G. E. O'Brien, I. Rodriguez, and P. Pamelá, Toward equity through participation in Modeling Instruction in introductory university physics, Phys. Rev. ST Phys. Educ. Res. **6**, 010106 (2010).
[49] E. Brewe, A. Traxler, J. De La Garza, and L. H. Kramer, Extending positive CLASS results across multiple instructors and multiple classes of Modeling Instruction, Phys. Rev. ST Phys. Educ. Res. **9**, 020116 (2013).
[50] V. Sawtelle, E. Brewe, and L. H. Kramer, Positive impacts of modeling instruction on self-efficacy, AIP Conf. Proc. **1289**, 289 (2010).
[51] R. M. Goertzen, E. Brewe, and L. Kramer, Expanded markers of success in introductory university physics, Int. J. Sci. Educ. **35**, 262 (2013).
[52] I. Rodriguez, G. Potvin, E. Brewe, and L. H. Kramer, The Impacts of Modeling Physics in Upper-Level Courses: The Persistence of Males and Females, *Proceedings of the Physics Education Research Conference, 2014, Minneapolis, MN* (AIP, New York, 2014), pp. 219–222.

[53] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.13.020129 for the complete Problem Solving and Representation Use Survey and for the Wilcoxon-signed-rank and Wilcoxon-Mann-Whitney significance tables.

[54] K. A. Douglas, A. Rynearson, S. Purzer, and J. Strobel, Reliability, validity, and fairness: A content analysis of assessment development publications in major engineering education journals, Int. J. Eng. Educ. **32**, 1960 (2016).

[55] 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**, 020103 (2017).

[56] A. L Traxler, X. C Cid, J. Blue, and R. Barthelemy, Enriching gender in PER: A binary past and a complex future, Phys. Rev. Phys. Educ. Res. **12**, 020114 (2016).

[57] F. Wilcoxon, Individual Comparisons by Ranking Methods, Biom. Bull. **1**, 80 (1945).

[58] H. B. Mann and D. R. Whitney, Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to The Annals of Mathematical Statistics. ® www.jstor.org, Ann. Math. Stat. **18**, 50 (1947).

[59] M. P. Fay and M. a. Proschan, Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules, Stat. Surv. **4**, 1 (2010).

[60] A. Y Gordon, G. Glazko, X. Qiu, and A. Yakovlev, Control of the mean number of false discoveries, Bonferroni and stability of multiple testing, Ann. Appl. Stat. **1**, 179 (2007).

[61] A. V Frane, Are Per-Family Type I Error Rates Relevant in Social and Behavioral Science?, J. Modern Appl. Statistical Methods **14**, 12 (2015).

[62] N. Cliff, Dominance statistics: Ordinal analyses to answer ordinal questions., Psychol. Bull. **114**, 494 (1993).

[63] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Academic Press, New York, 1969), p. 23.

[64] J. Romano, J. D. Kromrey, J. Coraggio, J. Skowronek, and L. Devine, Exploring methods for evaluating group differences on the NSSE and other surveys: Are the t-test and Cohen's d indices the most appropriate choices?, *Proceedings of the Annual meeting of the Southern Association for Institutional Research* (2006), pp. 14–17, [arXiv:1011.1669v3], http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.595.6157&rep=rep1&type=pdf.

[65] D. McPadden and E. Brewe, Network Analysis of Students' Representation Use in Problem Solving, *Proceedings of the Physics Education Research Conference, 2014, Minneapolis, MN* (AIP, New York, 2015), p. 227.