

Effects of representation on students solving physics problems: A fine-grained characterization

Patrick B. Kohl and Noah D. Finkelstein

Department of Physics, University of Colorado at Boulder, Boulder, Colorado 80309, USA

(Received 31 December 2005; published 26 May 2006)

Recent papers document that student problem-solving competence varies (often strongly) with representational format, and that there are significant differences between the effects that traditional and reform-based instructional environments have on these competences [Kohl and Finkelstein, *Phys. Rev. ST Phys. Educ. Res.* **1**, 010104 (2005); **2**, 010102 (2006)]. These studies focused on large-lecture introductory physics courses, and included aggregate data on student performance on quizzes and homeworks. In this paper, we complement previous papers with finer-grained in-depth problem-solving interviews. In 16 interviews of students drawn from these classes, we investigate in more detail how and when student problem-solving performance varies with problem representation (verbal, mathematical, graphical, or pictorial). We find that student strategy often varies with representation, and that in this environment students who show more strategy variation tend to perform more poorly. We also verify that student performance depends sensitively on the particular combination of representation, topic, and student prior knowledge. Finally, we confirm that students have generally robust opinions of their representational skills, and that these opinions correlate poorly with their actual performances.

DOI: [10.1103/PhysRevSTPER.2.010106](https://doi.org/10.1103/PhysRevSTPER.2.010106)

PACS number(s): 01.40.Fk

I. INTRODUCTION

Student representational and metarepresentational skills have been the subject of a number of studies in physics education research (PER). By “representational skills” we refer to students’ ability to appropriately interpret and apply various representations of physics concepts and problems. These different representations can include verbal, mathematical, graphical, and pictorial formats, though these categories are by no means comprehensive or orthogonal. Much work has been done in PER on student representational skills. Some of these studies investigate students’ handling of particular representations.^{1–3} Others focus on teaching students how to use multiple representations when solving problems.^{4–7} There has, however, been less work studying how student problem-solving performance varies as the problem representation is changed.⁸ A few papers have broadened the study of representations to include metarepresentational skills.^{9–12} Roughly, these studies ask what students know *about* representations, and how that knowledge might affect student performance.

In two previous works,^{13,14} we presented the results of a three-semester study investigating student representational and metarepresentational skills in large-lecture introductory algebra-based physics courses. The first paper¹³ asked two major questions. First, to what extent does student performance on physics problems depend on the representational format of the problem? We found several instances where student performance was significantly different on problems that were essentially isomorphic, but posed in different representations. This supports previous results in PER,⁸ and is consistent with other works in cognitive science (Ref. 15, for example). In our data, it appeared that subtle features of the problems were cueing students to answer differently, and that those features were specific to the particular problem and representation; that is, it was not simply a case of students being generally more competent with one set of representa-

tions than another. Unfortunately, the aggregate nature of the performance data prevented us from making specific inferences about how problem features influenced students’ answers in more than a couple of cases.

The second question was whether performance would improve if students were given a choice of representation. This is a metarepresentational question: Do students know enough about representations and about their own abilities to make productive choices? Our results^{13,14} were complex. The first course studied was the second semester of an algebra-based introductory sequence taught by a traditional professor, Physics 202. In this course, the students who were given a choice of representation (the choice group) often performed either significantly better or significantly worse than students who were assigned a quiz (the control group). Whether the choice group did better or worse varied with representation and topic, but not in a predictable way. In the following two semesters, we performed the study again in Physics 201 and 202 courses taught by a professor who made significant use of PER-based reforms. In these courses, the choice vs control splits were nearly nonexistent.

In the second paper, we attempted to explain the presence or absence of choice vs control performance splits and to describe the representational character of the reform and traditional class environments studied.¹⁴ Analysis of the course components (including lectures, exams, and homeworks) suggested that students in the reformed physics courses were being exposed to and held responsible for using a broader variety of representations. In addition, the quizzes used in the studies asked students to write briefly about which representations they thought that they would perform best with. Student opinions regarding their representational skills appeared to be constant across topics in all three of the courses studied, though performance data suggested that students were not always accurate in selecting the representational format in which they would best perform. These results taken together suggested a possible explanation of the choice vs control splits and lack thereof. If the representation-rich reform

course environment was leading students to develop broader representational skills, then any metarepresentational weaknesses (inaccuracies in assessing one's abilities with regard to a particular problem, for instance) would have less impact. That is, whether students received their preferred representation or not would have less effect on their performance than in the traditional course where, perhaps, student skills were less broad.

The above studies made considerable progress toward addressing the questions asked, but the aggregate nature of the available data was a significant limitation. In this paper, we present data from 16 student interviews that investigate why we observe the above effects in more detail. These interviews had two goals. First, we hoped to gain a deeper understanding of how problem representation affects performance. We address this in two parts. The original aggregate data suggested that oftentimes student responses would be cued by subtle features of the problem or representation at hand. These cueings are much more observable in detailed problem-solving interviews, and we find numerous examples of how particular, representation-dependent features of problems can cue students differently, strengthening the conclusions from the first papers. Next, in watching students solve these problems it became clear that students' solution strategies varied with problem representation. Some prior work in math and science education has investigated the effect of problem representation on student strategies.^{16,17} Koedinger¹⁷ found that young algebra students often chose different problem-solving strategies for problems in different representations (word problems vs symbolic problems, for example). Such representation-dependent strategy variations could begin to explain the different performances we observe in students solving different physics problems. We are unaware of prior work in PER directly investigating the representation dependence of student strategies, though there is substantial work investigating the teaching of representation- and multiple-representation-based strategies⁴⁻⁷ and investigating the differences (often involving representation use) between expert and novice problem-solving strategies.¹⁸⁻²¹ We quantify the variation of student strategy in these interviews, and discuss the effect strategy variation had on student performance. The data suggest that students who were more consistent in their choice of strategy performed better.

Our second goal was to validate our conclusions regarding student assessments of different representations and of their own representational skills (for example, which representations they considered most useful, and why), as these metarepresentational issues were key to our arguments and are interesting in their own right. In our previous studies our data included only written student assessments. In these interviews, we questioned students more thoroughly to investigate and clarify the students' evaluations of representational formats and of their own abilities. Since the interviews took place several weeks after the original in-recitation study quizzes (described in the following methods section), we could also test whether students' opinions of representations and of their own skills are stable. We demonstrate that their opinions are fairly robust and that they do not correlate well with their actual performance, which is consistent with our prior suppositions.

II. METHODS

A. Previous studies

In the previous studies,^{13,14} we gave students sets of four prerecitation homework problems, one problem in each of four representational formats (verbal, mathematical, graphical, and pictorial). Following the homeworks, students took a one-question multiple-choice recitation quiz which was presented in one of the four formats (students were responsible for solving only one quiz, in one format). Some of the students were allowed to choose between the quiz formats without examining the problems ahead of time. Others were assigned a quiz. This procedure was repeated in each semester for an additional topic, for a total of six trials spread across three classes.

B. Student interviews

We recruited eight students from each of the reform 201 and 202 courses, with one student participating in both 201 and 202 interviews, for a total of 15 unique individuals. Students were solicited for these interviews via course email near the end of the semester and were paid for their time. The student volunteers were primarily from the top two-thirds of the class, with final grades ranging from A to C. Seven of the students were male, and eight were female.

The 201 and 202 populations consisted largely of life-science, kinesiology, and premed students. Physics 201 covered mechanics, while 202 covered electromagnetism, optics, and atomic physics. The courses made heavy use of Peer Instruction,²² personal electronic response systems ("clickers"²³), and well-integrated lecture demonstrations. In a previous study, we established that the exams and lectures in these courses made use of a broader selection of representations than a comparable traditional course, and also explicitly used multiple representations more often.¹⁴ Thus, none of the students in this study were drawn from a strictly traditional course environment.

The interviews were clinical (in the style of Ref. 24), lasted between 30 and 45 minutes, and were videotaped. In the interviews, students were assigned a number of quiz problems identical to those found in the recitation quizzes and were asked to work aloud. The problem order was varied from interview to interview. The interviewer (the lead author) did not provide assistance except for problem clarification, and generally allowed the students to work without interruption except for prompts to think out loud and requests for further explanation of what the students did and why. After the students solved the quiz problems, the interviewer asked the students a number of questions regarding which representations they found most useful and why. Students were not told whether they answered a question correctly until after the interview.

With two subject areas and four representations per class, there were eight problems available in total for each interview (see the appendix of Ref. 13 for a complete set of problems), and a total of 16 problems used in all interviews (eight for 201 students, and eight for 202 students). Students completed anywhere from two to eight problems in the time allowed, with an average of approximately five problems per

interview. Thus, students often solved the same problem in different representations in the same sitting. Students rarely expressed explicit awareness of the isomorphic nature of the problems until prompted at the end of the interview. Nevertheless, giving students all these problems in the same sitting could have led students to be more consistent in their problem-solving approaches than they would have been had they approached each problem uninfluenced by any others. In spite of this, we observed a number of students being inconsistent in their approaches to problems in different representations, and so we do not consider this limitation to be a serious issue.

We analyzed these interviews in three ways. First, we coded each student answer as correct or incorrect. Second, we coded each student's strategy in solving each problem, noting whether it was qualitative or quantitative, which concepts it made use of (energy vs force, for instance), and flagging any special features (analogies to other material, for instance). Third, we flagged instances where students expressed a favorable or unfavorable view of a particular representation. More detail on the interview analysis can be found in the Analysis sections, as can sample codings with interview excerpts.

III. DATA AND ANALYSIS

This presentation and analysis of student interview data has two parts, each addressing one of the two research goals identified in the introduction: (A) the effects of representation on problem solving and (B) investigating students' metarepresentational competence.

A. Effect of problem representation on performance

As noted above, we find two major ways in which problem representation can affect performance. Student problem-solving strategies can vary with problem representation, and students can cue on particular, often representation-dependent problem features when selecting their answers. We do not consider it likely that these two dimensions include all the ways in which representation might affect performance, nor do we consider them perfectly distinct (for instance, it appears that student choice of strategy is often cued by representation-dependent problem features). Nevertheless, these categories (which emerged from our analysis) are useful in organizing the available data. We begin by presenting the data on strategy variation, including several examples of student strategies for reference in later sections.

1. Student problem-solving strategies

Students in our interviews ranged from being very diverse in the strategies they used to being very consistent when confronted with different representations. In this section we examine the selection of strategies employed by students in our interviews.²⁵ As noted above, we coded student strategies according to major problem features such as whether the solution was qualitative or quantitative. The strategy divisions varied from problem to problem; for instance, students'

strategies for solving the spectroscopy problems could be binned according to whether they used an energy approach or a wavelength approach (or, rarely, both), and students' strategies for solving the spring problems could be binned according to whether students worked in terms of force or in terms of energy. We also flagged unusual (within this sample) approaches such as analogies to objects not directly related to the problem (planets, flashlights), or the nonquantitative use of equations as conceptual support (which will be discussed more later).²⁶

We begin with interview excerpts of problem solutions from two students whose strategies were varied (one student who is correct across all representations used and one who is correct on only one representation), and one student whose strategies were consistent. Student names are pseudonyms. In the following, we will indicate a student's grade in parentheses the first time we mention them in an example.²⁷

a. Adam: Varied strategies, mixed success (B-). Adam solved the graphical, mathematical, and pictorial versions of the spectroscopy problems given in Physics 202. In Fig. 1 we show the graphical and pictorial formats of the spectroscopy quiz. Note that the question statements and answer choices map from one format to the other. The mathematical format was similar, and asked students to calculate the difference between the $n=2$ and $n=4$ energy levels given a change in the ionization energy. Adam solved the graphical problem using an analogy to gravity. He states:

"... probably means it's going to be more tightly bound to the nucleus, so the levels are probably going to be lower than they would be if there's a lower ionization level I guess. If we were to reduce the gravity constant so, you know, mass would have less force than it actually does, you'd wind up with wider orbits, I suppose."

Adam chose answer B, which is incorrect and shows an increased spacing, analogous to wider orbits.

Adam then solved the mathematical format. This solution was based on a proportionality argument backed by an equation: The difference in spacing should be that calculable from $E = -13.6/n^2$, but scaled by the factor $11/13.6$ (the ratio of the hypothetical and real ionization energies). Adam performed the calculation and selected the correct answer, C.

For the pictorial format quiz, Adam used a wavelength picture:

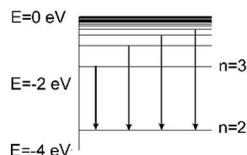
"Well, if you've got a lower ionization energy, that means that the photons that are released when it's ionized are gonna have to have lower energy, which means they have to have a longer wavelength, so we're going to see a spectrum that's redshifted slightly compared to what we have normally."

Based on this argument, Adam chose the correct answer, C.

b. Betty: Consistent strategies, consistent success (B). Betty solved the verbal, mathematical, and graphical versions of the pendulum quiz (see Fig. 2 for the verbal and pictorial problems). In each case, she used the equation for the period of a pendulum to support her reasoning, whether that reasoning involved an exact calculation or not. Regarding the verbal format quiz, Betty said:

Spectroscopy Problem -- Graphical Format

The energy level diagram below shows the electron transitions that lead to the Balmer series:



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 11 eV instead of 13.6 eV. Choose the graph that best represents what the new energy levels would look like.

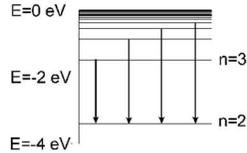
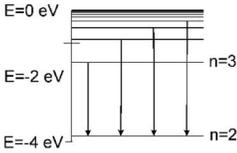
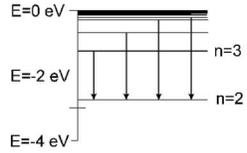
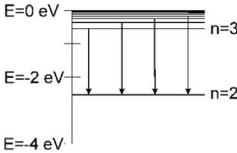
- A) 
- B) 
- C) 
- D) 
- E) Something else.

FIG. 1. (Color) Graphical and pictorial representations of the spectroscopy quiz given in Physics 202.

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.

- A) 
- B) 
- C) 
- D) 
- E) Something else.

“I believe the equation for period is like, $T, 2\pi L/G$ or something [writes $T=2\pi\sqrt{L/G}$]. So if L is under the square root, um, then it would be the square root of that it would be $2L$, so, after one second it should be halfway from where it started.”

Betty’s written calculations coupled with the above indicate that she decided that a quadrupling of the length of the pendulum would halve its period, resulting in the pendulum traveling half as far in a given time interval. She selected A, the correct answer.

Betty began the mathematical quiz by using the same period equation to determine the pendulum’s position, giving no obvious attention to the x vs t equation presented. She then used the v vs t equation provided to correctly determine the sign of the velocity (though in the process she made two offsetting sign errors), and reached the correct answer.

Betty’s solution to the graphical pendulum quiz was extremely analogous to the above solutions, once again using the period equation to make a proportionality argument backed by an equation:

“... extending the pendulum to four times as long, so since the period is, er T is $2\pi\sqrt{L/G}$, four times as long so T is related by being two times as great... If the period is two times as long, at one second it would be half as far, so it would be at this point, zero.”

Again, Betty used this reasoning to select the correct answer, A.

c. Carmen: Varied strategies, consistent success (A).

Carmen solved the mathematical and pictorial versions of the quiz on springs given in Physics 201, shown in Fig. 3. Her approach to the mathematical problem involved little calculation, and was based in part on an energy argument:

Pendulum Problem -- Verbal Format

I set up a pendulum in front of you and pull it back (to your right), and then let it go. The pendulum takes one second to reach the point opposite from where it started.

Now I lengthen the pendulum's string until it is four times as long as it was, with the mass unchanged. I pull the pendulum back to the right again (far enough that the string is at the same angle as before), and let it go. Where is it after one second? Circle the correct answer.

- A) Straight up and down, and moving left.
- B) Opposite from its starting position.
- C) Straight up and down, and moving right.
- D) Back in its starting position.
- E) Somewhere else.

Pendulum Problem -- Pictorial Format

I pull a pendulum back to the position shown below on the left and let it go. It takes one second to swing into the position shown below and on the right.

Start:

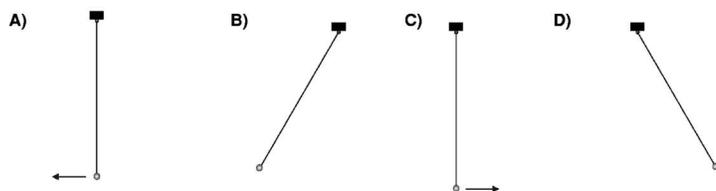


After one second:



Now I change the pendulum so that it is four times as long as before, with the same mass. I pull the pendulum back to the same side to the same angle as before and then let it go.

Select the picture that corresponds to the position of the new pendulum after one second. If the pendulum is straight up and down, select the picture that indicates the correct direction of the motion.



E) None of the above.

Carmen: "... because, okay, because that's where it's at rest, so... lowest energy... I think it's that. I don't know how to explain it.

Interviewer: "What do you mean by 'at rest'?"

Carmen: "Where it's at rest when the ball is hanging."

Carmen's reasoning is difficult to infer precisely, but based on this statement and additional follow-up questions, it appears that she was arguing that the ball will be moving fastest when the spring is at the equilibrium position, where the spring potential energy is lowest. This is the position at which kinetic energy is a maximum when the spring is moving, though she made no explicit mention of kinetic energy. She selected answer B, which is correct.

Carmen solved the pictorial version using a force and acceleration argument, with no mention of energy:

"It makes kind of more sense pictorially, because you know it's stopped here (points at the top of the motion) and that's going to be accelerating, and it'll accelerate until it's at the point it was at rest. And then, the tension of the spring, I don't know what it's called, will start causing it to decelerate, so it's going to be fastest at that point."

Carmen selected the correct answer, B. Notably, the pictorial format of the quiz explicitly depicts stretched and compressed springs, which could perhaps be associated more easily with forces than energies. Also, the mathematical format includes energy equations, but not force equations. While Carmen did not make explicit use of these equations, it is plausible that their presence could have cued an energy argument.²⁸

FIG. 2. Verbal and pictorial representations of the pendulum quiz given in Physics 201.

2. Strategy selection and performance

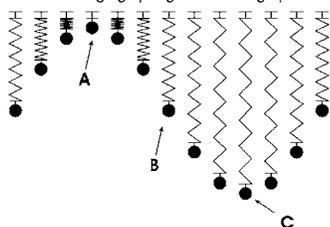
We next ask whether variation in a student's problem-solving strategies is associated with his or her performance. For this, we need to be able to describe a student as being generally varied or generally consistent in their strategy selection. Each student solved problems in two different topic areas. We consider strategies for quiz problems within a particular topic to be different if they have noticeably different qualitative features. Different features include arguing in terms of energy versus force or using an equation-based argument versus a qualitative proportionality argument. If the number of strategies a student employed was greater than half the total number of problems solved within a topic, we designate that student as varied in strategy choice.³⁰ If the student used only two strategies, (one for all representations of each of the two quizzes), we designate the student as consistent. If the student fits in neither category, or uses strategies that are not easily categorized as distinct or similar, we designate the students as mixed state with respect to strategy.

Of the eight Physics 201 students interviewed, seven solved four or more quiz problems. One solved only two problems, and will not be considered in this analysis (designation by the above standards is impossible). Of these seven students, two (including Betty from the above) were designated as consistent. Five were designated as varied (including Carmen). None were mixed state. The two consistent students solved a total of 11 problems, answering nine correctly, for an 82% success rate. The five varied students solved a total of 24 problems, answering 13 correctly, for a 56% success rate.

Of the eight Physics 202 students interviewed, all but one solved six or more problems; the other solved four. Of these eight students, one was designated as consistent. Five were

Spring Problem -- Pictorial Format

A ball on a hanging spring is oscillating up and down as shown in the following snapshots.



At which point is the ball moving the fastest?

- A) The ball is moving fastest at point A.
- B) The ball is moving fastest at point B.
- C) The ball is moving fastest at point C.
- D) The ball is moving fastest at some other point (not A, B, or C).

Spring Problem -- Mathematical Format

A ball is hanging from a spring at rest at $y = 0$ cm. The spring is then compressed until the ball is at $y = 5$ cm, and is then released so that the ball oscillates. Up is in the positive- y direction. At which point y is the ball moving fastest? Note that

$$K = \frac{1}{2}mv^2 \quad U_{spring} = \frac{1}{2}k(y - y_0)^2 \quad \text{and} \quad U_{gravity} = mgy$$

where y_0 is the unstretched length of the spring.

- A) $y = -5$ cm
- B) $y = 0$ cm
- C) $y = +5$ cm
- D) The ball is moving fastest at some other point.

FIG. 3. Pictorial and mathematical representations of the quiz on springs given in Physics 201.

designated as varied (including Adam). Two were mixed state. The consistent student solved six of six problems correctly, for a 100% success rate. The five varied students solved 16 of 32 problems correctly, for a 50% success rate. When we average together the data for the two classes, we see that the consistent students have a success rate of 88% over 17 problems, and the varied students have a 52% success rate over 56 problems. The standard deviations for the averaged data are 11% and 22%, respectively. These average success rates are different at a $p=0.007$ significance level using a two-tailed binomial proportion test, and indicate that of the students studied here, those who are consistent in their problem-solving strategies are outperforming those who are not. We do not consider this an obvious result; one could easily imagine that students who tailor their solutions to the representation at hand would perform better than those that do not. We also cannot be completely certain as to why the more consistent students are performing better (nor can we claim that this pattern would hold in general, outside this sample), though it may be that these students have a better abstract understanding of the problem at hand and thus are less sensitive to representation-dependent problem features. The major data regarding strategy variation are summarized in Table I.

3. Representation-dependent cueing

In previous work,¹³ we claimed that subtle, representation-dependent features of a problem can (but do not necessarily) have a significant impact on student success. The aggregate nature of the prior data limited us to only a few examples of this representation-dependent cueing. Here, we bolster our earlier claim with additional examples made

TABLE I. Success rate for students using a variety of problem-solving strategies on different representations of a problem versus students consistently using the same strategy. Numbers in parentheses indicate the number of problems solved by students in a category. Data are presented for each class and for both classes averaged together. The performance difference in the ‘‘Overall’’ category is significant at the $p=0.007$ level.

	Varied	Consistent
Physics 201	56% (24)	82% (11)
Physics 202	50% (32)	100% (6)
Overall	52% (56)	88% (17)

possible by the student interviews, and begin to organize these examples into categories that emerged in the analysis. Note that these categories are intentionally narrow. We are not trying to claim that these observed behaviors are ubiquitous; rather we wish to establish their existence for some combinations of representation, topic, and student.

a. Literal interpretation of graphs and language. In several cases, students appeared to interpret problem features overliterally, drawing inappropriate conclusions or making inappropriate associations that led to incorrect answers. This could include literal reading of actual language, or literal interpretation of other representations (for instance, assuming that a graph feature that is positioned lower on the graph indicates a smaller quantity). Two examples of these literal readings follow. As we saw above, Adam solved three different representations of the spectroscopy quiz problems. He handled the mathematical and pictorial problems expertly (in the opinion of the interviewer), but was incorrect in the case of the graphical problem. Adam’s solution to this problem used an analogy to gravity, which was followed by the comparison of energy levels to planetary orbits. We note that the graphical representation was the only representation of this problem that was oriented vertically, with higher-numbered levels being placed physically higher than lower-numbered levels. Adam appeared to cue on this: ‘‘probably means it’s going to be more tightly bound to the nucleus so the, levels are probably going to be lower than they would be if there’s a lower ionization level I guess.’’ We also note that in another part of the interview (not shown) Adam stated that this material connects strongly with what he knows in chemistry about electron ‘‘orbitals.’’ We speculate that here, a student who is otherwise quite skilled with the material at hand was prompted to respond incorrectly by a literal interpretation of the vertical arrangement of the levels, which was perhaps reinforced by an association between the word orbitals and the concept of planetary orbits. Student Doug (B-) exhibited a similar pattern. Doug solved the pictorial version of the spectroscopy quiz first. He expressed uncertainty, but used appropriate reasoning and selected the correct answer. Later, he solved the graphical version, exhibiting correct reasoning at first:

‘‘If the electric charges are weaker, and the um electrons aren’t held as closely to the nucleus, to the atom, that prob-

ably means the distance between... going from n , $n=1$ to $n=2$ is shorter. ... In doing this I've been trying to, um, think it through like chemistry to help me kinda understand ... I've been using $n=1$ $n=2$ as orbital levels, so as you um increase, your n , to higher energies. So if it's lower energy, then, the amount of energy to go from $n=3$ to $n=2$ or 2 to 1 would be smaller."

Doug then appeared to make a "lower means lower" type of error when choosing between the available answers:

"I don't think the right answer is on here, to be completely honest (circles none of the above). Because I think that you would have all of your lines slightly smooshed closer and *at a lower energy*. [italics added]"

Despite having a fairly complete grasp of the problem, and despite the energy axis being clearly labeled, Doug focused on the physical position of the lines in the representation, and concluded that lower-energy lines needed to be *below* the higher-energy lines. Note that there is some potential for confusion here: Since the energy of these bound states is negative, it would be technically possible to use the word "lower" to refer to states that occupy a lower position on the graph. However, student explanations and answer choices indicate this was not their intent.

These overliteral readings call to mind the "what you see is what you get" (WYSIWYG) knowledge element proposed by Elby,³¹ where students interpret a representation in the simplest, most literal way possible (a bump on a graph corresponds to a hill), even if further reflection demonstrates that they "know" the material well. This WYSIWYG element is a representational analog of the phenomenological primitives (or p-prims) described by diSessa,³² which include such basic reasoning elements as "lower means lower."

b. Feature pairs versus feature sets. Students solving the graphical or pictorial versions of the spectroscopy quiz usually did so in one of two modes. They viewed the problem either in terms of the relative positions of pairs of features (for example, the spacings between particular energy levels or spectral lines), or in terms of the positions of the sets of features as a whole (noting that the entire set of energy levels compressed or expanded, or that the entire set of spectral lines redshifted or blueshifted). We describe this as a focus on feature pairs versus entire feature sets. Students solved a total of five pictorial spectroscopy quiz problems, and seven graphical spectroscopy problems. In four of the five pictorial problems, student language made it clear (for example, "I'd probably pick the same set of spectral lines but at a lower energy") that they were focusing on feature sets, while in one case the student used a mixed approach. In four of the seven graphical problems, students appeared to be focused on feature pairs (for instance, saying " n_3 to n_2 , it's gonna be, it's gonna be smaller because uh, it's only 11 electron volts instead of 13.6") with one case of a student focusing on feature sets and two cases of students using mixed approaches. These numbers are too small for statistics to be used comfortably, but they suggest that the graphical representation of the problem might cue a different class of strategies than the

pictorial representation. While each representation has a set of discrete features, we note that the discrete nature of the energy level diagram is emphasized by the arrows indicating transitions, the lack of background clutter, and the fact that the steps between levels have physical meaning that is discussed in class, while the steps between spectral lines in the pictorial format do not. It is thus perhaps reasonable that students would be likely to focus on a pair of discrete energy levels and their relative positioning while using the graphical representation, and treat the spectral lines as a single feature (or set of features) moving against a background when using the pictorial representation. This is in some ways similar to the results of Mestre *et al.*³³ who found that showing students videos of pairs of balls rolling on tracks versus videos of individual balls tended to trigger different readout strategies, in which students considered either absolute or relative ball motions.

c. Presence or absence of equations. The presence of equations, not surprisingly, caused some of the interview subjects to attempt calculation-based solutions when they were unnecessary. The quiz problem on springs displayed in Fig. 3 could be (and was) solved in a number of ways, and was relatively easy compared to the other problems. Students Emma (B) and Mindy (A-) solved the pictorial version of this quiz quickly and correctly. Emma inferred the velocity of the ball from the size of the frame-to-frame change in the ball's position. Mindy recalled from lecture that balls on springs are moving fastest as they pass through their equilibrium position. Both Emma and Mindy solved the mathematical version of the spring quiz incorrectly. Mindy did not know what to do with the available equations and gave up after several minutes. Emma set the kinetic and potential energy equations equal to each other to derive $v = \sqrt{2gy}$, inferring from that the ball is moving faster when its y coordinate is higher. She then selected answer choice C. In each case, the student arguably was using equations without thinking about why they were using them, a metarepresentational failure frequently observed by Schoenfeld³⁴ (and, most likely, by any practicing instructor). Also, Sabella³⁵ notes that students take fundamentally different approaches when engaging in problems they perceive as qualitative or quantitative. From this perspective, Emma and Mindy can be seen as incorrectly judging problems to be strictly quantitative, and engaging in inappropriate solution strategies as a result.

In contrast to the above, some students were more expert-like in their handling of equations. Most notable were those that used an equation qualitatively to support the reasoning involved in solving a nonmathematical representation of a problem. By "qualitatively" we mean that the student used the equation without performing complete calculations. For an example, see Betty's solution to the pendulum problems in the earlier section on strategy selection. There, she used the formula $T = 2\pi\sqrt{L/G}$ to find that quadrupling the length of a pendulum doubled the period, but she did not calculate a specific number for T at any time. There were ten instances of students using mathematics qualitatively to support their solutions of nonmathematical representations, spread out

over six students and five different combinations of topic and representation, with each topic represented except for the quiz on springs. The students solved the problem correctly in nine of these instances (90% correct). These six students had an overall success rate on all problems of 80%, compared to a success rate of 46% for the other ten students. This difference is statistically significant at the $p=0.002$ level. While the samples presented here are small, it appears that the more successful students (at least with regards to these tasks) are the ones capable of using (or willing to use) mathematics as a conceptual support, in addition to any calculation-based uses. This result is again consistent with Sabella³⁵ and others¹⁸ who have shown that expert problem solvers integrate qualitative and quantitative approaches more often than novices when solving physics problems.

B. Student assessment of representations and of themselves

In our original study of this subject,¹³ we were able to examine student assessments of their own representational skills and preferences in aggregate, through comments solicited on the study quizzes. We found two notable results. First, students were generally fairly consistent in their representational preferences and assessments. Students that provided comments on both quizzes given over the course of the semester usually claimed to be good at the same representations on each quiz. Second, students' actual performance on the quiz and homework problems correlated poorly with their assessments of their own skills. The robustness of their opinions and the lack of correlation between self-assessments and performance was a key feature of our previous arguments.¹⁴ The second major goal of the current study is to further validate the conclusions of previous papers using the more detailed information available through interviews. Below, we present data supportive of each of the two results just discussed.

1. Student consistency

In each interview, we asked students which representations they preferred to work in, and why. We also invited students to make any comments about the representations themselves that they wished. Follow-up questions were posed as needed to clarify student responses. We then compared students' interview assessments with the comments we requested on their original recitation quizzes. Of the 15 students interviewed (one was interviewed twice), 14 students had provided comments on the recitation quizzes. Of these, 12 were consistent in their assessments across the quizzes and interviews. We count as consistent any student whose quiz and interview statements (which were separated by several weeks) considered the same representations to be favorable and/or unfavorable. This standard for consistency includes students who favorably rate one representation on one quiz, favorably rate another representation on another quiz, and then describe each of those favorably during the interview. The other two students were somewhat consistent, but not completely. One student stated a preference for one representation (pictorial) on the recitation quizzes and stated a preference for two others (math and verbal) during the inter-

views; we consider this stance to be inconsistent, though not directly contradictory. Another student was consistent in her assessments of the math, verbal, and graphical representations, but indeterminate in her evaluation of the pictorial representation.

2. Correlation with performance

We have found that student performances do not improve in general when students are given a choice of representational format in which to work.^{13,14} This result suggests that students' assessments of their own representational skills are not very accurate. In the interviews, we found that most of the students were quite consistent regarding which representations they preferred and/or thought they would do best at. Including interview problems, pre-recitation homeworks, and recitation quizzes, these students have solved a large selection of study problems in different representations (an average of 12 problems per student). To correlate student performances with their assessments, we must be able to define whether a student has rated a particular representation favorably, unfavorably, or neutrally/not at all. Student responses are almost always unambiguously favorable or unfavorable. Two sample quotes [from students Doug (B-) and Tina (B-/C+)] are

“Sometimes the verbal ones are worded in ways which are hard to think about. Math ones tend to be straightforward”

“Given the choice, pictorial and mathematical is preferred. I hate graphical.”

We analyzed the responses of 13 students. The two students whose interview and quiz remarks were inconsistent were discarded, and the student who only provided interview assessments (and not quiz assessments) was included. For each of these students, we noted whether he or she evaluated a representation favorably, unfavorably, or neutrally/not at all. We then divided the set of problems that each student completed into subsets according to whether the student evaluated the representation favorably or otherwise, and found their average performance on each subset. By comparing student performance on each subset, we could describe each student as having done better or worse on problems in their preferred representation. As an example, student Nate (C) rated the mathematical and graphical representations favorably. He rated the pictorial representation unfavorably, and did not discuss the verbal representation in the quizzes or in the interview. We see examples of these ratings in the following subsection, which includes example remarks from Nate. Nate solved eight mathematical and graphical problems and answered three correctly. He solved seven verbal and pictorial questions, and answered three correctly. Thus overall, he performed worse (though only slightly) on problems that were in his preferred formats. Six of the 13 students analyzed performed better on their favored representations than on other problems. The other seven performed worse. Thus, we see no correlation between their problem-solving success and their representation assessments. This result is consistent with the aggregate performance data in earlier papers,^{13,14} where providing students a choice of

problem representation did not produce a consistently negative or positive effect.

3. Sample comments

In this section we present sample comments from two students, one from Physics 201 and one from Physics 202.

a. Betty: Physics 201 (B). Betty was consistent in her assessments of the different representations, and was also accurate in her assessment in that she performed better overall on the representations that she claimed to do better at. Betty was randomly assigned recitation quiz problems. In response to the quiz question “Of the four problem formats you saw on the prerecitation homework, which do you think you would do best at, given the choice? Why?,” Betty wrote “Mathematical—I relate to equations.” She expanded on this in the interview:

“Well, for me personally, I really um do poorly on conceptual questions, and I’ll do a lot better if I have numbers and I can use equations and figure things out in that way.”

Later, she said

“I don’t like [the verbal quiz] at all because, I don’t like it when you make me draw my own picture because sometimes I misunderstand what I’m reading and I draw the picture incorrectly which, affects how I do the whole problem. [the graphical quiz] really appeals to me because, if I’m looking at waves I like to see this [the graph], and then I really like calculus so, I like to see like if you’re telling me take the derivative of a graph, I can like draw it out for myself. ... Ideally, if I was doing a problem, I’d like them to have pictures and equations, because that would probably help me the most.”

In two instances, Betty used mathematics to support non-mathematical quiz representations in the manner described earlier, which supports her claims that she is particularly comfortable working with equations. Also notable is the fact that she answered the mathematical version of the spring quiz incorrectly due to an incorrectly labeled picture that she drew (her reasoning was correct), which is also consistent with her assessment.

b. Nate: Physics 202 (C). Nate was consistent in his assessments of the different representations, and performed slightly better on those he rated unfavorably or neutrally than on those he preferred. In response to the recitation quiz question, he wrote:

“Mathematical because that is all we do with CAPA [a web-based homework system³⁶] and I have gotten used to it.”

In the beginning of the interview, Nate was asked why he preferred the mathematical formats. His response was

“Kinda what I got used to during physics. Almost all of the tests use math.”

Later in the interview, Nate is asked to provide input on all of the representations he has seen:

“I guess, the pictures I find hardest because there’s a lot goin’

on with them. Um, the graphs are a little more manageable, just, I dunno, I guess there’s, you know, there’s numbers on the page and actual information given to you while this [the pictorial quiz] is kind of an interpret the information kind of thing. When I do like the math best is if you know the equation you just plug it in and you’re set.”

Betty and Nate were fairly typical in the length and style of their remarks.

IV. DISCUSSION

This paper set forth two goals. First, we planned to further investigate the means by which problem representation affects performance. We have a number of examples available here and in previous work.¹³ Second, we hoped to validate our conclusions regarding student assessments of different representations and of their own representational skills.

A. Effect of problem representation on performance

We have a number of examples available here and in previous work¹³ of how representation affects performance. The representation effects are complex and appear to depend on a number of things. Performance can be influenced by the particulars of the representation and how it is implemented, as in the cases where the vertically oriented graphical spectroscopy problem triggered a “lower is lower” misinterpretation that was not observed with the horizontally oriented pictorial spectroscopy problem. Performance can also be influenced by prior student knowledge, including what topics a student has been taught in conjunction with the material at hand, or what other material a student feels comfortable in relating to the problem. An example is the case of Adam, who cited his chemistry experience when he drew an analogy between electron orbitals and planetary orbits. This analogy appears to have reinforced an error of the “lower is lower” variety. It is also likely that class norms and expectations,³⁷ such as whether students have been taught to draw pictures in support of particular problems, can play a role. For example, we see here instances of students unnecessarily using equations without a clear understanding of why, and we speculate that this stems from the (not unreasonable) expectation that a mathematically framed physics question will have a quantitative solution.

The above findings are complicated by the fact that most real situations will involve a combination of the factors discussed. Adam’s error probably stems from both a misapplication of his prior chemistry knowledge and a p-prim-like³² interpretation triggered by the specific features of the graphical problem representation. Emma’s inappropriate derivation and use of $v = \sqrt{2gy}$ for the sake of solving the mathematical spring problem likely started with a course expectation, and may have been strengthened by the fact that this equation is commonly used in a different mechanics context, making it appear familiar once produced. (This interpretation is speculation, as the interview did not probe this association directly.) This complex dependence of performance on representation, student knowledge, and course norms is consistent with what we have observed in previous work.¹³ Of course,

not all students exhibited the same sensitivity to changes in problem representation. We found that many students used a variety of solution strategies in their interviews while some used very few. Furthermore, those who used fewer strategies appeared to outperform those who used many. It is possible that students who are more comfortable and competent with a topic are guided more by the topic than the particular representation of that topic; that is, these students may have a better grasp of the abstract concept behind the representation and cue on the deep rather than the surface problem features.¹⁹

These results are reminiscent of previous results in PER that show significant differences in student approach and performance on qualitative versus quantitative problems. Such work includes that of Mazur,²² Sabella,³⁵ and the numerous University of Washington studies on student difficulties with various topics.³⁸ Indeed, some of the mathematical-format problems used in this study were quantitative in nature, in contrast to the more qualitative problems in the other formats (and, sometimes, in other questions using the mathematical format). However, the study here is focused on finer-grained problem divisions: Most of the problem representations were isomorphic, meaning that their intended solutions all included similar blends of quantitative and qualitative work. Substantial differences in performance and approach emerged in spite of this coarse similarity, which some may find surprising. While it is possible to use the broader characterization of qualitative or quantitative problems, we observe in this study that the focus on representational format provides insight into some of the factors that influence whether students choose to approach problems qualitatively or quantitatively. We also note that the converse may be true: If students view problems as quantitative or qualitative in nature, that can drive how they use representations (successfully or not).

B. Student metarepresentational skills

Our second goal was to further support the claim that student assessments are relatively constant over time and across topic, and that these assessments do not generally correlate well with their performance. In comparing student quiz comments to their interview responses, we see that student opinions are indeed fairly consistent, at least over the course of the semester being studied. Furthermore, the performance data support our claim that these opinions are not particularly accurate in general. We must note here that while these student opinions regarding representations appear to be robust over the course of the study, we cannot take this as evidence that students generally come into a physics course with well-formed assessments of representations and of their own representational skills. In some cases they certainly do—Tina expressed such a strong dislike of graphs that the opinion must have existed before she answered the study questions—but we suspect that in some cases students were being asked questions such as these for the first time, and that they were generating their opinions on the spot. Of course, neither do our data allow the opposite conclusion, that students do not generally come into a physics course with well-formed opinions.

Metarepresentational activities are not a part of a standard physics course, and we are curious as to what impact such activities could have. Studies have shown that attention to metalevel skills (including an explicit focus on epistemological issues) can have a positive impact on student performance in math³⁴ and physics³⁹ courses. Might explicit in-class attention to the uses and drawbacks of different physics representations in different classes improve student representational skills, or at least improve their own self-awareness? Tina was the only student to participate in an interview in both Physics 201 and 202, meaning that she had more formal opportunities to engage in metarepresentational reflection than the other students. Perhaps not coincidentally, she was the only student to explicitly challenge her own representational assessments, as seen near the end of the second interview:

Tina: “I picked the picture because pictures usually help, they usually make things easier. But not really... now that I’m looking at all of them. These two [indicates the mathematical and verbal spectroscopy problems]”

Interviewer: “Like, in your prior experience, you’ve found problems that have pictures to”

Tina: “Yeah, because they’re visual, I mean, you could see what was going on, but the information is much more straightforward in these [verbal and mathematical quizzes]. This one [the pictorial problem] I gotta kinda and figure out that this is wavelength; I need to hear it straightforward like in these. So. Yeah, I dunno, I always picked pictures because I thought it’d be easier and I don’t think it ever was.”

Interviewer: “Really? Even in the past?”

Tina: “Well, I mean like, until school got hard [laughs].”

Tina went on to suppose that pictures are useful when used in conjunction with other representations, such as written descriptions, equations, and graphs. It would appear that she moved from viewing pictures as intrinsically useful to viewing them as one of many tools that need to be used together in order to be most effective. This is, in our opinion, significant metarepresentational progress, and we find it satisfying to see that given the opportunity a student can make such gains. Also note that the metarepresentational failures observed in this study involve only a subset of student metarepresentational skills. Other studies have found evidence of significant metarepresentational strengths in students, especially with respect to their ability to generate new representations,^{10,40} so the picture of student metarepresentational competence is far from bleak.

C. Instructional implications

While this study did not have the goal of developing or testing new instructional materials or techniques, we can speculate as to the instructional implications of our results. We have found several cases in which student performance was significantly affected by which representation was used, and that impact could be tied to surprisingly small

representation-dependent problem features that either cued an answer directly or changed the student's overall strategy. Instructor awareness of this sensitivity to representation would likely be productive. This sensitivity would also be relevant in test construction, though we note that PER-based assessments (the FCI,^{41,42} FMCE,⁴³ and BEMA,⁴⁴ for example) generally contain a variety of representations.

We also found that students in our study were not very successful when selecting between different representations to work in, perhaps because of their limited experience with such decisions. This result suggests that if an activity provided students with similar representational freedom (that is, the freedom to select between canonical representations or create new ones), such an activity would probably need to be guided to be more helpful. Or, as an alternative, students would need preparatory metarepresentational instruction. We are not aware of any such activities in wide use today, but with the increasing interest in multiple representation use and metacognition, we can envision such an activity becoming more common in the future.

Finally, since this study and the previous work^{13,14} begin to specify *how* representation and performance are related, they suggest the possibility of developing a more complete theoretical understanding of how students handle representations when solving problems. An appropriate theory would (among other things) describe the manner in which students use representations and multiple representations, and would discuss how instructional environment can influence performance and use, which would no doubt have practical application. The data obtained so far provide potential starting points for such theoretical development, which we will reserve for future work.

V. CONCLUSIONS

In this paper and two previous ones, we have attempted to investigate student representational skills at many levels. We

have examined the impact of small-scale representation-dependent problem features, the variation of problem-solving strategy with representation, the role of instructional environment (representationally rich vs sparse), and student metarepresentational skills, with consideration of both aggregate data and in-depth interviews. We find two major results. First, student performance on physics problems does depend on problem representation, but the dependence is complex. Particular combinations of representation, topic, and student experience can result in much different performances, often as a result of different strategy selection. This finding suggests that it might be quite difficult to infer whether or not students understand a concept based on an assessment presented in only one representation of that concept. Second, students (when asked) will form consistent opinions regarding which representations they handle best, but these opinions correlate poorly with their actual performance. Since traditional introductory physics courses usually specify the representations to be used on a problem, this metarepresentational failure may not significantly impede student performance. On less-constrained (and more realistic) physics problems, lack of metarepresentational skills might be more significant. To our knowledge this dependence has not been studied formally. The results so far indicate that such a multilayered approach is necessary to begin to paint a picture of how physics students learn and use representations. In future work, we hope to broaden and strengthen our results and to develop the theoretical tools necessary to produce a more coherent understanding of student representational and metarepresentational skills.

ACKNOWLEDGMENTS

This work was supported in part by the NSF and by Colorado PhysTEC. We give special thanks to the rest of the Physics Education Research group at the University of Colorado at Boulder, and thanks also to P. Beale, T. Munsat, and J. Peterson for their cooperation and aid.

¹R. J. Beichner, Testing student interpretation of kinematics graphs, *Am. J. Phys.* **62**, 750 (1994).

²F. M. Goldberg and J. H. Anderson, Student difficulties with graphical representations of negative values of velocity, *Phys. Teach.* **27**, 254 (1989).

³L. 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).

⁴A. Van Heuvelen and X. Zou, Multiple representations of work-energy processes, *Am. J. Phys.* **69**, 184 (2001).

⁵A. Van Heuvelen, Learning to think like a physicist: A review of research-based instructional strategies, *Am. J. Phys.* **59**, 891 (1991).

⁶R. J. Dufresne, W. J. Gerace, and W. J. Leonard, Solving physics problems with multiple representations, *Phys. Teach.* **35**, 270 (1997).

⁷P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping, Parts i and ii, *Am. J. Phys.* **60**,

627 (1992).

⁸D. E. Meltzer, Relation between students' problem-solving performance and representational mode, *Am. J. Phys.* **73**, 463 (2005).

⁹A. A. diSessa and B. L. Sherin, Meta-representation: an introduction, *J. Math. Behav.* **19**, 385 (2000).

¹⁰A. A. diSessa, D. Hammer, B. L. Sherin, and T. Kolpakowski, Inventing graphing: Meta-representational expertise in children, *J. Math. Behav.* **10**, 117 (1991).

¹¹A. A. diSessa, in *Symbolizing, Modeling and Tool Use in Mathematics Education*, edited by K. Gravemeijer, R. Lehrer, B. van Oers, and L. Verschaffel (Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003).

¹²F. Reif, Understanding and teaching important scientific thought processes, *Am. J. Phys.* **63**, 17 (1995).

¹³P. 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).

- ¹⁴P. B. Kohl and N. D. Finkelstein, The effect of instructional environment on physics students representational skills, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010102 (2006).
- ¹⁵J. Zhang and D. Norman, Representations in distributed cognitive tasks, *Cogn. Sci.* **18**, 87 (1994).
- ¹⁶R. E. Mayer, Different problem-solving strategies for algebra, word, and equation problems, *J. Exp. Psychol. Learn. Mem. Cogn.* **8**, 448 (1982).
- ¹⁷K. R. Koedinger and M. J. Nathan, The real story behind story problems: Effects of representations on quantitative reasoning, *J. Learn. Sci.* **13**, 129 (2004).
- ¹⁸J. H. Larkin, in *Mental Models*, edited by D. Gentner and A. Stevens (Lawrence Erlbaum, Mahwah, NJ, 1983).
- ¹⁹M. Chi, T. H. Glaser, and E. Rees, in *Advances in the Psychology of Human Intelligence*, edited by R. J. Sternberg (Lawrence Erlbaum, Hillsdale, NJ, 1982), Vol. 1.
- ²⁰D. P. Maloney, in *Handbook of Research on Science Teaching and Learning*, edited by D. Gabel (MacMillan, New York, 1994).
- ²¹L. Hsu, E. Brewster, T. M. Foster, and K. A. Harper, Resource letter RPS-1: Research in problem solving, *Am. J. Phys.* **72**, 1147 (2004).
- ²²E. Mazur, *Peer Instruction: A User's Manual* (Prentice-Hall, Upper Saddle River, NJ, 1996).
- ²³<http://www.h-itt.com/>
- ²⁴A. A. diSessa, A. Elby, and D. Hammer, in *Intentional Conceptual Change*, edited by G. Sinatra and P. Pintrich (Lawrence Erlbaum, Mahwah, NJ, 2003).
- ²⁵Note that we use the term "strategy" broadly, and use it to include the problem features that a student chooses to focus on in addition to the overall solution plan.
- ²⁶Strategy categorization is necessarily somewhat subjective, but in nearly all cases, the categories used were unambiguous (either students use quantitative calculations or they do not; either they mention wavelength or they do not). Thus, we had only one researcher code and verify these data.
- ²⁷Our analysis focuses primarily on student performance on study tasks, and not on their in-class performance. Nevertheless, this information may be of interest.
- ²⁸Interestingly, no student made reference to both force and energy when solving any one problem. This is consistent with the findings of Sabella and Redish (Ref. 29) who demonstrate that even advanced students can struggle when moving from force to energy pictures and vice versa.
- ²⁹M. S. Sabella and E. F. Redish (unpublished).
- ³⁰For example, a student that solves three representations of each quiz, for a total of six representations, would be designated varied if he or she used four or more strategies in total. This could be either two strategies for each quiz topic, or three for one and one for the other.
- ³¹A. Elby, What students' learning of representations tells us about constructivism, *J. Math. Behav.* **19**, 481 (2000).
- ³²A. A. diSessa, in *Constructivism in the Computer Age* edited by G. Forman and P. B. Pufall (Lawrence Erlbaum, Hillsdale, NJ, 1988).
- ³³J. P. Mestre, T. C. Thaden-Koch, R. J. Dufresne, and W. J. Gerace, in *Research on Physics Education, Proceedings of the International School of Physics "Enrico Fermi," Course CLVI*, edited by E. F. Redish and M. Vicentini (IOS, Amsterdam, 2004).
- ³⁴A. Schoenfeld, in *Cognitive Science and Math Education*, edited by A. Schoenfeld (Erlbaum, Mahwah, NJ, 1987).
- ³⁵M. S. Sabella, PhD. thesis, University of Maryland, 1999.
- ³⁶<http://www.lon-capa.org>
- ³⁷N. D. Finkelstein, Learning physics in context: A study of student learning about electricity and magnetism, *Int. J. Sci. Educ.* **27**, 1187 (2005).
- ³⁸<http://www.phys.washington.edu/groups/peg/pubs.html>
- ³⁹A. Elby, Helping physics students learn how to learn, *Am. J. Phys.* **69**, S54 (2001).
- ⁴⁰F. S. Azevedo, Designing representations of terrain: A study in meta-representational competence, *J. Math. Behav.* **19**, 443 (2000).
- ⁴¹D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, *Phys. Teach.* **30**, 141 (1992).
- ⁴²M. H. Dancy and R. Beichner, Impact of animation on assessment of conceptual understanding in physics, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010104 (2006).
- ⁴³R. K. Thornton and D. R. Sokoloff, Assessing student learning of Newton's laws: The force and motion conceptual evaluation, *Am. J. Phys.* **66**, 228 (1998).
- ⁴⁴L. Ding, R. Chabay, B. Sherwood, and R. Beichner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010105 (2006).