

## Implementing and assessing computational modeling in introductory mechanics

Marcos D. Caballero,<sup>1,\*</sup> Matthew A. Kohlmyer,<sup>2,†</sup> and Michael F. Schatz<sup>1,‡</sup>

<sup>1</sup>*Center for Nonlinear Science and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA*

<sup>2</sup>*Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA*

(Received 26 July 2011; published 14 August 2012)

Students taking introductory physics are rarely exposed to computational modeling. In a one-semester large lecture introductory calculus-based mechanics course at Georgia Tech, students learned to solve physics problems using the VPython programming environment. During the term, 1357 students in this course solved a suite of 14 computational modeling homework questions delivered using an online commercial course management system. Their proficiency with computational modeling was evaluated with a proctored assignment involving a novel central force problem. The majority of students (60.4%) successfully completed the evaluation. Analysis of erroneous student-submitted programs indicated that a small set of student errors explained why most programs failed. We discuss the design and implementation of the computational modeling homework and evaluation, the results from the evaluation, and the implications for computational instruction in introductory science, technology, engineering, and mathematics (STEM) courses.

DOI: [10.1103/PhysRevSTPER.8.020106](https://doi.org/10.1103/PhysRevSTPER.8.020106)

PACS numbers: 01.40.Fk, 01.40.gb, 01.50.H–

### I. INTRODUCTION

Computation (the use of a computer to numerically solve, simulate, or visualize a physical problem) has revolutionized scientific research and engineering practice. In science and engineering, computation is considered to be as important as theory and experiment [1]. Systems that cannot be solved in closed form are probed using computation; experiments that are impossible to perform in a lab are studied numerically [2,3]. Yet, as computer usage has grown (e.g., facilitation of homework), most introductory courses have failed to introduce students to computation's problem-solving powers.

Computer usage (as opposed to *computation*) is ubiquitous in large introductory courses, and much work has been done to understand the role of computers in these courses [4]. In contrast, little is known about the role that computers might play as problem-solving tools in introductory courses. For example, given the constraints of a traditional large lecture STEM course, is it possible to provide sufficient instruction in computational problem-solving techniques such that students can apply these tools to new problems? Moreover, what challenges do students face when learning and applying computational

problem-solving techniques, and how might we mitigate those challenges through instruction? In this paper, we describe our efforts to address these questions.

Using computation in introductory physics courses has several potential benefits. Students can engage in the modeling process to make complex problems tractable. Students can also use computation to explore the applicability and utility of physical principles. In a way, students who compute are participating in work that is more representative of what they will do as professional scientists and engineers [5–8]. When constructing simulations, students are constrained by the programming language to certain syntactic structures. Hence, they must learn to contextualize problems in a way that produces a precise representation of the physical model [9,10]. One of computation's key strengths lies in its utility for visualizing and animating solutions to problems. These visualizations can improve students' conceptual understanding of physics [11]. It is important to distinguish between computational modeling and simply writing program statements. Computational modeling is the expressing, describing, and/or solving of a science problem by using a computer. It requires additional steps beyond simple programming, such as contextualizing the programming task and making sense of the physical model. While programming is taught in most introductory computer science courses, computational modeling is not.

We used computation in a large-enrollment introductory calculus-based mechanics course at the Georgia Institute of Technology to develop students' modeling and numerical analysis skills. We built upon previous attempts to introduce computation in introductory physics laboratories [12,13] by extending its usage to other aspects of students' course work. In particular, we taught students to use the VPython programming environment to construct models

\*Current address: Department of Physics, University of Colorado at Boulder, Boulder, CO 80309, USA.

†Current address: Advanced Instructional Systems, Inc., Raleigh, NC 27696, USA.

‡Corresponding author.  
michael.schatz@physics.gatech.edu

*Published by the American Physical Society under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

that predict the motion of physical systems [14]. We describe the design and implementation of homework problems to develop students' computational modeling skills in a high-enrollment foundational physics course (Sec. III). We also provide the first evaluation and explanation of students' skills when they attempt individually to solve a novel computational problem in a proctored environment (Secs. IV, V, VI, and VII). However, our study was not designed to provide evidence of instructional effectiveness but rather to demonstrate the types of common challenges observed in student work at the large scale. We discuss implications for instructional design, considerations regarding student epistemology and the assessment of knowledge transfer, as well as the broader implications of teaching computation to introductory physics students (Sec. VIII).

## II. APPROACHES TO IMPLEMENTING COMPUTATION

Since the development of inexpensive modern micro-computers with visual displays, there have been a number of attempts to introduce computation into physics courses. We review these attempts by decomposing them along two dimensions (*openness of the environment* and *size of intended population*) to indicate how our approach fits in with previous work.

Some have worked closely with a small number of students to develop computational models in an *open computational environment*. Historical examples include the Maryland University Project in Physics and Educational Technology [5,6], STELLA [7,8], and the Berkeley BOXER project [9,10]. Open computational environments are analogous to "user-developed" code in scientific research (as opposed to prepackaged or "canned" software). Students who learn to use an open environment have the advantage of viewing and altering the underlying algorithm on which the computational model depends. Moreover, students using an open computational environment might learn to develop new models to solve new problems. It is true, however, that students must devote time and cognitive effort to learning the syntax and procedures of the open environment's programming language. It is most desirable to have students focus on developing the physical model without spending excessive time and effort on the details of constructing working code (e.g., message handling, drawing graphics, garbage collection). It is, therefore, important to consider students' experience (or lack thereof) with programming when choosing an open computational environment.

Others have developed *closed computational environments* for use at a variety of instructional levels. These environments have been deployed in a number of settings ranging from a few students to large lecture sections. Examples of closed environments include Physlets [15] and the University of Colorado's Physics Educational

Technology simulations [16,17]. Closed computational environments are analogous to "canned" code in scientific research. Students can set up and operate the program but do not construct it, nor do they have access to the underlying model or modeling algorithm (a "black-box" environment). User interaction in closed computational environments is often limited to setting or adjusting parameters; a small number of Physlets do allow users to manipulate equations of motion. Closed computational environments are useful because they typically require no programming knowledge to operate, run similarly on a variety of platforms with little more than an Internet browser, and produce highly visual simulations.

It is possible for computational models created in any open environment to be used in a closed manner. Users can be restricted (formally or informally) from viewing or altering the underlying model. Models developed using Easy Java Simulations [18] (EJS), an open-source, freely available Java simulation authoring tool, have been used in a closed manner at a variety of scales and instructional levels [19,20]. However, since it is an open environment, all the features of the physical and computational models in an EJS simulation are accessible to the user. Furthermore, EJS has made authoring high-quality simulations possible for students without much programming experience. Some have proposed teaching upper-division science majors to develop computational models using EJS [21].

VPython [14], an open computational environment, has been used to teach introductory physics students to create computational models of physical phenomena [12]. Typically, students write all the program statements necessary to model the physical system (e.g., creating objects, assigning variables, and performing numerical calculations) using IDLE, VPython's default development environment. The additional details of model construction (e.g., drawing graphics, creating windows, mouse interactions) are handled by VPython and are invisible to the students. VPython supports full three-dimensional graphics and animation without the additional burden to students of learning a great deal of object-oriented programming [22]. Given its roots in the Python programming language, VPython can be a powerful foundation for students to start learning the tools of their science or engineering trade. Moreover, VPython is an open-source, freely available environment accessible to users of all major computing platforms.

The Matter & Interactions (M&I) curriculum [23] introduces computational modeling as an integral part of the introductory physics course. Many of the accompanying laboratory activities are written with VPython in mind, and a number of lecture demonstrations are VPython programs. In the traditional implementation of M&I, the practice of constructing computational models is limited to the laboratory. In a typical lab, students work in small groups to

complete a computational activity by following a guided handout. They pause periodically to check their work with other groups or their teaching assistant (TA). Students' computational modeling skills are evaluated by solving fill-in-the-blank test questions in which they must write the VPython program statements missing from a computational model.

Our approach to teaching computation uses an open environment (in VPython) and builds on our experience with M&I to extend the computational experience beyond the laboratory. We chose to use an open environment to teach computation in order to provide students with the opportunity to look inside the computational "black box" and construct the model themselves. Furthermore, we aimed to teach students how to develop solutions to analytically intractable problems. We chose VPython (instead of, e.g., Java, C, or MATLAB) because it has a number of helpful features for novice programmers, can be used to construct high-quality three-dimensional simulations easily, and is freely available to our students. VPython is also conveniently coupled to M&I, allowing us to leverage our years of experience with teaching M&I. While our implementation builds on our M&I experience, it is not limited to that experience. We describe our implementation philosophy in the next section.

### III. DESIGN AND IMPLEMENTATION OF COMPUTATIONAL HOMEWORK

We aimed to develop an instructional strategy that helps computation permeate the course, but which does not require that students have previous programming experience. At Georgia Tech, half of our introductory mechanics students have no programming experience; the other half have had some introduction to programming in high school or in their previous two semesters at Georgia Tech [24]. The Georgia Tech "introduction to programming" course is typically taught using MATLAB, not VPython. Furthermore, our implementation had to be easily deployable across large lecture sections, the setting in which most introductory calculus-based courses are taught. Our philosophy held that students should learn computation by altering their own lab-developed programs to solve slightly modified problems outside of the lab. This design philosophy was informed by what research scientists do quite often: they write a program to solve a problem and then alter that program to solve a different problem that is of interest to them. We envisioned developing computational activities that would start with guided inquiry and exploration in the laboratory and finish with independent practice on homework. In this vision, students would work with TAs in the laboratory to develop a program that solves a problem. Students would then modify that program on their own to solve a different problem on their homework. In our implementation, 11 out of a total of 13 mechanics laboratories contained a

computational modeling activity. Students solved 13 computational modeling homework problems (one per week) which were embedded in their regular homework (~ problems per week).

The class of problems that becomes available to students who have learned computation is large and diverse; we chose to focus our efforts on teaching students to apply Newton's second law iteratively to predict motion. Students taking a typical introductory mechanics course would learn several motion prediction equations emphasizing kinematics, a way of describing the motion without explicitly connecting changes in the motion to forces (i.e., dynamics). These kinematic formulas are quite limited in their applicability; students can only apply them to problems in which the forces are constant. This can confuse students when they are presented with a situation where such formulas do not apply [25]. Students taking most introductory mechanics courses are limited to solving problems that are analytically tractable. Furthermore, the special case of constant force motion is usually the capstone of motion prediction in an introductory mechanics course. Some courses might teach students to analytically determine the motion of an object subjected to some integrable function  $a(t)$  or subjected to linear ( $F \sim v$ ) air drag in one dimension, but neither case demonstrates the full predictive power of Newton's second law. By contrast, computation allows instructors to start first and foremost with Newton's second law and emphasize its full predictive power. Students can numerically model the motion of a system as long as they are able to develop a physical model of the interactions and express it in the computational environment. The numerical integration technique used to predict motion is a simple algorithm.

As a concrete example of our design, we show a midsemester laboratory activity and homework problem in which students modeled the gravitational interaction between two bodies. In this example, students develop a VPython program that models the motion of a craft as it orbits Earth (Fig. 1). Students later make a number of modifications to this program to solve a new problem on their homework. This example is useful because it illustrates both the level of sophistication we expect of students and the types of alterations we ask students to make on their homework.

In groups of three, students wrote a program in the laboratory to model the motion of a low-mass craft as it orbits Earth (Fig. 1). In VPython, they created the objects (lines 4–6), assigned the constants and initial conditions (lines 8–10, 12–13, and 15–17), and set up the numerical integration loop (lines 19–29). The program statements in this loop included those which calculated the net force (lines 21–23) and updated the momentum (using Newton's second law) and position of the craft (lines 25–26). When developing their physical model, students discussed the

```

1  from __future__ import division
2  from visual import *
3
4  craft = sphere(pos = vector(10e7,0,0), color = color.white, radius = 1e6)
5  Earth = sphere(pos = vector(0,0,0), color = color.blue, radius = 6.3e6)
6  trail = curve(color = craft.color)
7
8  G = 6.67e-11
9  mcraft = 1500
10 mEarth = 5.97e24
11
12 vcraft = vector(0,2400,0)
13 pcraft = mcraft*vcraft
14
15 t = 0
16 deltat = 60
17 tf = 365*24*60*60
18
19 while t < tf:
20
21     r = craft.pos-Earth.pos
22     rhat = r/mag(r)
23     Fgrav = -G*mEarth*mcraft/mag(r)**2*rhat
24
25     pcraft = pcraft+Fgrav*deltat
26     craft.pos = craft.pos + pcraft/mcraft*deltat
27
28     trail.append(pos = craft.pos)
29     t = t + deltat
30
31 print 'Craft final position: ', craft.pos, 'meters.'
```

**Initial Conditions****Force Calculation****Newton's Second Law****Position Update**

FIG. 1 (color online). Under the guidance of their TAs, students wrote the VPython program above in the laboratory. This program modeled the motion of a craft (size exaggerated for visualization) orbiting Earth over the course of one “virtual” year. To construct this model, students had to create the objects and assign their positions and sizes (lines 4–6), identify and assign the other given values and relevant initial conditions (lines 8–10, 12–13, and 15–17), calculate the net force acting on the object of interest appropriately (lines 21–23), and update the momentum and position of this object in each time step (lines 25–26).

motion of Earth; it experiences the same (magnitude) force as the space craft. Students reasoned that its change in velocity is negligible. Hence, students did not include the motion of Earth in their VPython program. When writing this program in the lab, students could seek help from TAs at any time. The accuracy of the students’ completed code was checked by their TAs. After completing the lab, students had written a VPython program which could model the motion of the craft moving around Earth for any arbitrary amount of time.

In the week following the lab, students solved a computational homework problem in which they used the computational model that they had written in lab to solve a modified problem. Students were asked to alter their initial conditions to predict the position and velocity of the craft after a specified time. To solve this problem successfully, students had to identify and make changes to their initial conditions (lines 4, 6, 9–10, and 12) and their integration time (line 17). In addition, students had to add

an additional print statement (after line 31) to print the final velocity of the craft.

Computational homework problems were deployed using the WebAssign course management system, which facilitated the weekly grading of students’ solutions. To create the homework problem, we numerically integrated several hundred initial conditions and stored the solutions, including final quantitative and qualitative results. Each student was assigned a random set of initial conditions corresponding to a particular set of results. Randomization ensured a high probability that each student received a unique problem. Students used their assigned initial conditions and wrote additional statements to answer the questions posed in the problem. Students entered numerical answers into answer fields and selected check boxes to answer qualitative questions. On these weekly assignments, only students’ final results were graded; their code was uploaded for verification purposes but not graded. Grading programs for structure and syntax

at this large scale require additional work by TAs who are already charged with a number of other teaching and grading tasks. Computational homework problems were generally completed in the week that followed the associated laboratory activity.

To facilitate student success and to help students learn to debug their programs, each assignment contained a *test case*—an initial problem for which the solution (i.e., the results from the numerical integration) was given. When writing or altering any computer program, even experts can create programming errors (bugs). Learning to debug programs is part of learning how to develop computational models. This test case ensured that a student's program worked properly and helped to instill confidence in students who might otherwise have been uncomfortable writing VPython programs without the help of their group members or TAs. After a student checked their program against the *test case*, they completed the *grading case*—a problem without a given solution.

In keeping with our overall design philosophy, most of the homework problems that students solved had similar designs as the aforementioned example. In particular, students built a computational model in the laboratory and independently used that fully functioning model to solve a modified problem on their homework. On the first four homework assignments (of which the previous example is the fourth), students made only a few modifications to their programs, altering their initial conditions and adding a new print statement. In the next several labs, students learned to model more complicated systems (e.g., three-body gravitational problem, spring dynamics with drag) while learning new algorithms such as decomposing the net force vectors into radial and tangential components. Students also learned to represent these force components as arrows in VPython. On the homework problems associated with these labs, students still used their lab-developed programs to solve new problems by changing initial conditions and representing new quantities with arrows.

The last two homework problems that students solved were not related to the laboratory exercises; we intended to emphasize the utility of learning to predict motion using Newton's second law. To solve these problems, students wrote all the statements missing from a partially completed code to predict the motion of two interacting objects. These were interactions that students had not seen before (e.g., the anharmonic potential and the Lennard-Jones interaction). In these problems, we omitted the appropriate initial conditions and the statements that numerically integrated the equations of motion. Students had to contextualize the word problem into a programming task and produce a precise representation of the problem in the VPython programming environment. With regard to programming tasks, students had to do no more than identify and assign variables and implement the usual motion prediction algorithm for these two problems. A similarly designed

problem was used as an evaluative assignment and is discussed in detail in Sec. IV.

#### IV. EVALUATING COMPUTATIONAL MODELING SKILLS

Students performed as well on computational modeling homework problems as they did on their analytic homework. We found no statistical difference between students' performance on their computational homework and their performance on analytic homework using a rank-sum test (analytic 84.6% versus computational 85.8%) [26]. A side effect of the additional practice was an overall increase in student performance on handwritten fill-in-the-blank programming questions on exams. However, this result did not indicate what fraction of students were able to solve these computational homework problems without assistance. While randomizing initial conditions between each student's realization ensured that students' solutions differed with high probability, working programs could be distributed easily from student to student by Email. We note that the distribution of students' programs might not be deleterious; students who received these programs must still have read and interpreted the program statements to enter in their initial conditions, make changes to the force law, or print additional quantities. This is a more complex interaction than simply plugging numbers into an algebraic solution that they discovered online. In a sense, students who worked with shared code were using a "closed" computational environment.

Nevertheless, we wanted to measure how effective students were at individually solving computational problems. We delivered a proctored laboratory assignment during the last lab of three different semesters to evaluate students' computational skills on an individual basis. Students received a partially completed program that created two objects (one low mass and one high mass), initialized some constants, and defined the numerical integration loop structure. We aimed to evaluate students' engagement of the modeling process by contextualizing a physics problem into a programming task. Furthermore, we intended to assess certain programming skills, namely, students' abilities to identify and assign variables and implement the numerical integration algorithm. The assignment was delivered using WebAssign in a timed mode (30 minutes) in each of the 16 lab sections over eight meeting periods. TAs were not permitted to help students debug their programs, nor were students allowed any additional aids (Internet, book, notes, lab partners, etc.). Furthermore, the assignment was IP restricted and password protected; students could only access the assignment during their assigned lab time. Even though the assignment was proctored over 4 days, we believe it is unlikely that a large number of students shared useful information about the assignment; scores for each lab section were statistically indistinguishable. The format of the assignment was

identical to students' final two homework problems; students were given a test case to check their solution before solving the grading case.

For this assignment, students modeled the motion of the low-mass object as it interacted with the high-mass object through a central force. The nature of the force (attractive or repulsive) and its distance dependence ( $r^n$ ) were randomized on a per-student basis. We also randomized some of the variable names in the partially completed program to hinder copying. After adding and modifying the necessary program statements, students ran their program and reported the final location and velocity of the low-mass object. During the assignment, students did not receive feedback from the WebAssign system about the correctness of their answer, but they were given three attempts to enter it. Similar to students' online homework, only the final numerical answer was graded. These additional attempts afforded students who experienced browser issues or made typing mistakes the ability to resubmit their answers without penalty.

Performance varied from semester to semester (Table I) because the assignment was modified slightly between each semester in order to streamline delivery (version 1 to version 2), reduce transcription errors, and improve presentation (version 2 to version 3). In the first semester, students were permitted to attempt version 1 of the assignment twice due to a logistical issue with the initial administration of the assignment. The majority of students (64.3%) were able to model the grading case successfully on the second administration of the assignment. Students' performance on version 1 was likely inflated because some students were able to work the problem twice [27]. Students solved version 2 only once, and student performance dropped. A number of students were confused by the randomized exponent on the units of one of their initial

TABLE I. As part of a final proctored lab assignment, students completed a partially constructed program which modeled the motion of an object under the influence of a central force. The partially written program defined the objects, some constants, and the numerical integration loop structure. The initial conditions, the sign ( $\pm$ ) and distance dependence ( $r^n$ ) of the force, and object names were randomized on a per-student basis. Slightly modified versions (Ver.) of this assignment were given at the end of three different semesters. Modifications were made to streamline delivery (version 1 to version 2), minimize transcription errors, and improve presentation (version 2 to version 3). Students' performance on version 1 was likely inflated because some students were allowed to work the problem on two separate occasions.

Ver.	Correct	Incorrect	% Correct
1	303	168	64.3
2	201	193	51.0
3	316	176	64.2
Overall	820	537	60.4

conditions (Sec. VI). About half of the students (51.0%) were able to model the grading case successfully. Students were more successful on version 3 of the assignment; 64.2% modeled the grading case correctly.

Overall, roughly 40% of the students were unable to model the grading case. Students with some previous programming experience were no more successful than those without any experience as indicated by a contingency table analysis [28]. To determine exactly what challenges they faced while completing this assignment, we reviewed the program of each student who failed to model the grading case. Students who uploaded their programs to the WebAssign system received a small extra credit bonus on the proctored assignment. Only a very small number of students ( $< 1\%$ ) did not upload any code. We limited our review to the programs submitted for versions 2 and 3 of the assignment.

## V. SYSTEMATICALLY UNFOLDING STUDENTS' ERRORS

Students must perform several tasks to successfully write and execute the program for the proctored assignment. Students must interpret the problem statement; that is, they must contextualize a word problem into a programming task. They must review the partially completed program and identify the variables to update. Students need to apply their knowledge of predicting motion to the problem using VPython. They must identify that the force is nonconstant and then write the appropriate programming statements to calculate the vector force. Students need to then complete the motion prediction routine by writing a statement to update the momentum of the low-mass object.

Using an iterative-design approach, we developed a set of binary (affirmative or negative) codes to check which tasks students performed correctly and which errors they made. An initial review of students' uploaded programs yielded the mistakes that were made most often. These common mistakes formed the basis for the codes. The codes were developed empirically, and several iterations were made before they were finalized. Two raters tested the codes by coding a single section of student-submitted programs ( $N = 45$ ). The raters resolved their differences (which further explicated the codes) and then recoded the section. The final codes (Table II) were used by both raters independently to code the remaining sections ( $N = 324$ ). The final codes had high inter-rater reliability; both raters agreed on 91% of the codes.

We classified each code into one of three procedural areas: *using the correct given values* (IC), *implementing the force calculation* (FC), and *updating with the Newton's second law* (SL). These areas were congruent with the broad range of difficulties and misconceptions that students exhibited with their erroneous programs. Each code is explained in greater detail in Appendix A.

TABLE II. Incorrectly written programs were subjected to an analysis using a set of codes developed from common student mistakes. The codes focused on three procedural areas: *using the correct given values* (IC), *implementing the force calculation* (FC), and *updating with the Newton’s second law* (SL). We reviewed each of the incorrectly written student programs for each of the features listed below. These codes are explained in detail in the appendix.

<b>Using the correct given values (IC)</b>	
IC1	Used all correct given values from grading case
IC2	Used all correct given values from test case
IC3	Used the correct integration time from either the grading case or test case
IC4	Used mixed initial conditions
IC5	Exponent confusion with $k$ (interaction constant)
<b>Implementing the force calculation (FC)</b>	
FC1	Force calculation was correct
FC2	Force calculation was incorrect but the calculation procedure was evident
FC3	Attempted to raise separation vector to a power
FC4	Direction of the force was reversed
FC5	Other force direction confusion
<b>Updating with Newton’s second law (SL)</b>	
SL1	Newton’s second law (N2) was correct
SL2	Incorrect N2 but in an update form
SL3	Incorrect N2 attempted update with scalar force
SL4	Created new variable for $\vec{p}_f$
<b>Other (O)</b>	
O1	Attempted to update (force, momentum, position) for the massive particle
O2	Did not attempt the problem

Determining where students encountered difficulties with these tasks might help explain how students learn this algorithmic approach to use Newton’s second law to predict motion. Because we reviewed students’ programs after they were written, we are unable to comment directly on students’ challenges with contextualizing the problem. Our work was limited to analyzing students’ procedural efforts (i.e., identifying variables and implementing the numerical integration algorithm). However, each of us has several years of experience with teaching computational modeling to introductory students from a wide variety of backgrounds. Hence, some information about students’ thoughts and actions could be inferred from this analysis.

### VI. FREQUENCY OF ERRORS IN STUDENTS’ PROGRAMS

We measured the frequency of students’ errors within each category (IC, FC, and SL) by mapping binary patterns extracted from our coding scheme to common student mistakes. The number of possible binary patterns we observed in our data ranged from nine for SL to 17 for FC with 13 possible for IC. Not all the codes within a given

category were independent; hence, the number of possible binary patterns is much less than  $2^n$ . Within a given category, we found that a large percentage of students could be characterized by just a few error patterns (between four and seven).

The errors we observed were not necessarily unique to computational problems. The most notable errors involved calculating forces or updating the momentum. Most of these errors appeared to be physics errors reminiscent of those made on pencil-and-paper problems. Many of them could have been mitigated by qualitative analysis. Some errors were unique to computational models and the iterative description of motion because they did not prevent the program from running, but nevertheless caused it to model the system inappropriately. Still others (e.g., replacing initial conditions) might have appeared to be simple careless mistakes; when investigated closely, however, these errors highlighted the fragility of students’ knowledge.

#### A. Initial condition errors

Students had to identify and update a total of eight given values: the interaction constant ( $k$ ), the “interaction strength” ( $n$ ), the mass of the less massive particle, the position and velocity of both particles, and the integration time. Most students with incorrect programs (88.8%) fell into one of seven IC patterns (Table III). Students in ICa (27.6%) did not use correct values from the grading case. Instead, these students identified and replaced all the initial conditions with those from the test case (IC2), including the integration time (IC3). Those in ICb (16.0%) used a mixed set of initial conditions, that is, conditions from both the grading and test cases (IC4), and selected one of the two given integration times (IC3). Students in ICc (14.4%) had a full set of correct initial conditions. These students identified and correctly replaced all the initial conditions with those from the grading case (IC1), including the

TABLE III. Only seven of the 14 distinct code patterns for the IC category (Table II) were populated by more than 3% of the students. The patterns (ICx) are given by affirmatives (Y) and negatives (blank) in the code columns (IC#). The percentage of students with each pattern is indicated by the last column (%). These seven patterns accounted for 88.8% of students with erroneous programs.

Pattern	Initial condition codes					%
	IC1	IC2	IC3	IC4	IC5	
ICa		Y	Y			27.6
ICb			Y	Y		16.0
ICc	Y		Y			14.4
ICd			Y		Y	13.8
ICe				Y		7.9
ICf			Y			5.2
ICg	Y					3.8

integration time (IC3). Students who appeared in ICd (13.8%) used the correct integration time (IC3) but had a single initial condition mistake; they were confused by the exponent on the units of the interaction constant (IC5). Students in ICe (7.9%) used a variety of initial conditions and given values (IC4). Those students in ICf (5.2%) only identified and employed a given integration time (IC3); their initial conditions were completely incorrect. Finally, students in ICg (3.8%) used appropriate initial conditions from the grading case (IC1) but did not employ the grading case's integration time. Most students might have simply forgotten to update one or more of the initial conditions from either the default case or the test case (ICb, ICe, ICf, and ICg). A small fraction of students with mixed initial conditions had values from all three cases.

Students in ICa were most likely stuck on the test case because they had trouble with another aspect of the problem. These students were unable to obtain the solutions provided in the test case, and so they kept working on it. It is possible a number of these students ran out of time while trying to debug their programs.

It is difficult to say definitively if students with mixed initial conditions (ICb and ICe) were unable to identify the appropriate values, as we reviewed students' programs only after they were submitted. It is possible that these students were just careless when making changes, but they might have been unable to identify and update these quantities. Some students could have been in the process of updating these quantities when they ran out of time and uploaded their programs.

Identifying and updating variables in a program is not a trivial task for students. In fact, students' difficulties with updating variables highlights the fragility of their computational knowledge. As an example, consider the students who confused the exponent on the length unit of the interaction constant ( $k$ ) for the exponent in the scientific notation of the numerical value of  $k$  when they defined it in their programs (ICd). The distance dependence of the central force was randomized, and hence the units of the interaction constant ( $k$ ) were dependent on a student's realization. In version 2 of the assignment, the exponent on the length unit of  $k$  was colored red (WebAssign's default behavior for random values). A student in ICd would read  $k = 0.1 \text{ Nm}^3$  to mean  $k = 0.1\text{E}3 = 100$  rather than  $k = 0.1 \text{ Newton times meters cubed}$ . In version 3 of the assignment, we changed the exponent's text color to black like the rest of the nonrandom text. The overall frequency of this mistake dropped from 30.5% to 9.1%.

## B. Force calculation errors

Students were given the magnitude of the force as an equation ( $F = kr^n$ ) and told that their (attractive or repulsive) force acted along the line that connected the two objects. In solving this problem, students had to correctly calculate the magnitude of the central force and identify

TABLE IV. Only six of the nine distinct code patterns for the FC category (Table II) were populated by more than 3% of the students. The patterns (FCx) are given by affirmatives (Y) and negatives (blank) in the code columns (FC#). The percentage of students with each pattern is indicated by the last column (%). These six patterns accounted for 98.8% of students with erroneous programs.

Pattern	Force calculation codes					%
	FC1	FC2	FC3	FC4	FC5	
FCa		Y		Y		23.9
FCb	Y					22.2
FCc		Y				15.7
FCd		Y	Y			14.6
FCE						14.0
FCf		Y			Y	8.4

the unit vector ( $\hat{r}$ ) and sign ( $\pm$ ) for their own realization. Almost all students (98.8%) appeared with one of five FC patterns (Table IV). Students in FCa (23.9%) implemented a reasonable force calculation procedure in the calculation loop (FC2) but reversed the direction of the net force (FC4). Those in FCb (22.2%) performed the force calculation correctly (FC1). Students in FCc (15.7%) implemented the procedure correctly (FC2) but were likely to include a force irrelevant to the problem (i.e., gravitational or electric interactions) or compute only the magnitude of the net force. Students who appeared in FCd (14.6%) implemented the force calculation procedure correctly (FC2) but attempted to raise the separation vector to a power (FC3). Students in FCE (14.0%) showed no evidence of an appropriate force calculation procedure (all codes negative); the procedure either was completely incorrect (e.g., used the differential form of the impulse-momentum theorem) or was calculated outside the numerical integration loop (i.e., a constant force). Those students in FCf (8.4%) had an appropriate force calculation procedure (FC2) but invented a unit vector for the net force (FC5).

The difficulties that students faced when numerically computing the net force likely stems from a weak grasp of the concept of vectors. Students in FCa made directional mistakes (e.g., changing the sign of one of lines 21–23 in Fig. 1) that could have been easily identified and rectified by drawing a sketch of the situation, a problem-solving strategy that is practiced in the laboratory. Those who raised the separation vector to a power (FCd) likely transcribed the central force equation (replacing  $r$  by  $\vec{r}$ ) without thinking that this operation was mathematically impossible [ $\vec{F} \sim k(\vec{r})^n$  versus  $\vec{F} \sim k|\vec{r}|^n\hat{r}$ ]. We have found that students attempt a similar operation on pencil-and-paper problems: raising components of a vector to a power [e.g.,  $(\vec{r})^n = \langle r_x^n, r_y^n, r_z^n \rangle$ ]. Such mistakes are quite common among introductory physics students [29]. However, in the pencil-and-paper case, students are not immediately directed to their mistake as they are in a programming



environment. VPython raises an exception error when this operation is attempted. These students appeared to be unable to parse this error into any useful information. Students who make this type of error might be helped by additional exposure to translating force equations to precise programmatic representations [10]. Some students invented a unit vector (FCf) for the net force. This was most likely because they had computed a scalar force and tried to add a scalar impulse to the vector momentum. VPython raises a different exception error upon any attempt to add a scalar to a vector. These students were able to parse this error but resolved it incorrectly.

Other students (FCc) might have incorrectly contextualized the problem by including an irrelevant force (i.e., gravitational or electrical interactions). The problem clearly stated that the two objects were far from all other objects. It did not explicitly state to neglect the gravitational interaction between the objects. However, the gravitational interaction could be safely neglected for the range of masses and distances we had chosen. Furthermore, nothing about the charge of the objects was mentioned in the problem statement. It is surprising that students included these interactions in their models. One possible explanation for the inclusion of these interactions is that students had memorized how to solve the gravitational and Coulomb problems because these problems had appeared on their homework several times and on an exam. They might have panicked and simply wrote all possible forces they could remember. We have observed similar contextualization issues in laboratory instruction and on handwritten exams.

A number of students (FCe) did not employ the force calculation algorithm at all. Some of these students computed the net force (e.g., lines 21–23 in Fig. 1) outside the numerical integration loop (e.g., before line 19 in Fig. 1). In this case, the net force was effectively constant and therefore only correct at  $t = 0$ . A program with correct syntax will run regardless of its physical implications. This error is unique to computational problems in which motion is predicted iteratively. Students in introductory physics rarely use Newton’s second law to predict the motion of objects that experience nonconstant forces. Other students who fell into FCe wrote “creative” program statements. Students in this group manipulated some quantities in the loop but did not perform any physically relevant calculations. The number of students with “creative” program statements was relatively small.

### C. Newton’s second law errors

Students had to write a program statement similar to line 25 in Fig. 1 to properly update the momentum using Newton’s second law. Most students demonstrated no difficulty in remembering the formula for the momentum update but some met challenges with making that description precise [10]. Nearly all students (95.7%) fell into one

TABLE V. Only four of the nine distinct code patterns for the SL category (Table II) were populated by more than 3% of the students. The patterns (SLx) are given by affirmatives (Y) and negatives (blank) in the code columns (SL#). The percentage of students with each pattern is indicated by the last column (%). These 4 patterns accounted for 95.7% of students with erroneous programs.

Pattern	Second law codes				%
	SL1	SL2	SL3	SL4	
SLa	Y				69.7
SLb		Y	Y		13.2
SLc					7.9
SLd		Y			4.9

of four SL patterns (Table V). Most students appeared in SLa (69.7%) because they wrote Newton’s second law in the correct iterative format,  $p = p + F * dt$  (SL1). A much smaller number of students fell into SLb (13.2%). These students wrote Newton’s second law in an iterative format (SL2), but they attempted to update the vector momentum with a scalar force (SL3). Students in SLc (7.9%) were unable to write Newton’s second law in any form that updated (all codes negative). A small fraction (SLd, 4.9%) wrote Newton’s second law in an iterative form (SL2) but did so incorrectly (e.g., by creating a new variable).

Students who attempted to update the momentum with a scalar force (SLb) might have been facing difficulties with understanding vectors. The momentum update was presented as a vector equation ( $\vec{p}_f = \vec{p}_i + \vec{F}\Delta t$ ). These students might have been unable to unpack that representation into a precise programmatic description, but it was more likely that they calculated a scalar force (FCc) and then simply wrote the correct (vector) second law syntax. VPython raises an exception error if an attempt to add a vector to a scalar is made. The students appeared unable to parse this error into any useful information.

Students who were unable to write Newton’s second law in any form that updated (SLc) might have experienced difficulties with converting the second law formula into a precise and useful programmatic representation. Students in this category either wrote Newton’s second law in a nonupdate form (e.g., writing `deltap = Fnet * deltat` or `pf - pi = Fnet * deltat` as line 25 in Fig. 1) or wrote a number of program statements which manipulated quantities but performed no useful calculations. In either case, these students could benefit from the precision required by a programming language [10]. By requiring them to accurately represent Newton’s second law in their programs, they might begin to distinguish between the utility and applicability of its various algebraic forms.

Students who wrote Newton’s second law in a form that updated incorrectly (SLd) either remembered the formula

for the second law incorrectly or made a typographical error. These students would generally leave off the time step in the momentum update (e.g.,  $p = p + F$ ) or divide by it ( $p = p + F/\text{deltat}$ ). Dividing by the time step is a particularly egregious error because the numerical value of the time step was quite small. Hence, the impulse added in this case would have been huge. Students who made this error were unable to assess the state of the visualization (the particle flew off to “infinity”) to debug this error.

### VII. COMMON ERROR PATTERNS IN STUDENTS’ PROGRAMS

The patterns within individual categories (IC, FC, and SL) indicated the frequency of common mistakes students made when solving the proctored assignment, but a single student could make one or more of these mistakes. Evaluating a student’s complete solution requires an analysis using all the codes (Table II). In principle, the codes we developed could have had up to  $\sim 4300$  possible error patterns using all 16 codes. In fact, the intersections of code categories indicated that the number of distinct errors made by students across all categories was relatively small; we found only 111 distinct binary patterns. It is possible to relate these unique patterns in a manner that suggests dominant common errors.

Cluster analysis, a technique borrowed from data mining, is particularly well suited for this application because it characterizes patterns in complex data sets [30,31]. This technique has been used previously to classify students’ responses to questions about acceleration and velocity in two dimensions [32]. It was used here to determine the major features in students’ incorrect programs which were responsible for their failure.

We applied the cluster analysis technique to the data generated from our set of binary codes. We used the Jaccard metric [33] to measure intercluster distances and linked clusters using their average separation [34]. We tested several other metrics (e.g., Hamming, city block,

etc.). The Jaccard metric was chosen because it neglects negative code pairs. Both the Hamming and city block metrics produced similar pairings at low levels, but higher order clusters were difficult to interpret. We used average linkage to avoid the effects of “chaining” that appeared when nearest neighbor linkage [35] was used and because useful clusters were more difficult to distinguish when farthest neighbor linkage [36] was used.

Thirty clusters with intercluster distances below 0.5 were reviewed in detail. This cutoff was selected to minimize the number of unique clusters while still rendering clusters with useful interpretations. Most students (86.5%) appeared in seven of the 30 clusters (Table VI). These clusters had very few students ( $< 1\%$ ) with affirmatives in the “other” category. Codes O1 and O2 were dropped from Table VI for this reason. Each of the other 23 clusters were populated by less than 3% ( $N \approx 10$ ) of the students, and the bottom 18 clusters had less than 1% ( $N \approx 3$ ) each. Each of the dominant clusters demonstrated a unique challenge that students faced while solving the proctored assignment (Table VI).

Students in cluster A (23.8%) tended to remain stuck on the test case (ICa) due to an error in their force calculation. Reversing the direction of the force (FCa) was the most common mistake, followed by raising the separation vector to a power (FCd). Most students in this cluster had no trouble expressing Newton’s second law (SLa). These students worked diligently to solve the test case but were unable to do so. As a result, they did not proceed to the grading case.

Cluster B (19.8%) contained students who made mistakes while replacing the given values and initial conditions (any IC code except ICa). Some of these students worked with the grading case (ICc and ICg). Others might have been working with either case and had mixed conditions (ICb and ICe) or simply incorrect ones (ICf). Still others might have incorrectly assigned the exponent on the length unit of  $k$  to the exponent of  $k$  in scientific notation

TABLE VI. Only seven of the 30 clusters with an intercluster distance of less than 0.5 were populated by more than 3% of the students. The bottom 18 clusters were populated by less than 1% of students each. These seven clusters accounted for 86.5% of students. The percentage of affirmatives for each code (Table II) within any given cluster (A–G) is given to the nearest whole percentage. Codes with affirmative percentages greater than 60% are in bold font. These clusters had very few students ( $< 1\%$ ) with any affirmatives in the “Other” category; hence, the results from this category are not reported. The percentage of students in each cluster is indicated in the last column (%).

Cluster	Initial conditions					Force calculation					Second law				%
	IC1	IC2	IC3	IC4	IC5	FC1	FC2	FC3	FC4	FC5	SL1	SL2	SL3	SL4	
A	00	<b>68</b>	<b>93</b>	18	15	00	<b>100</b>	22	<b>66</b>	09	<b>95</b>	00	00	01	23.8
B	21	01	<b>86</b>	37	41	<b>88</b>	00	00	00	00	<b>97</b>	00	00	00	19.8
C	04	33	<b>76</b>	31	22	00	<b>94</b>	00	08	00	00	<b>98</b>	<b>98</b>	08	13.3
D	<b>98</b>	00	<b>85</b>	00	00	00	<b>85</b>	18	50	00	<b>98</b>	00	00	00	10.8
E	00	00	57	<b>75</b>	36	00	<b>100</b>	<b>79</b>	00	04	<b>89</b>	00	00	00	7.6
F	00	<b>100</b>	<b>96</b>	00	00	<b>65</b>	00	00	00	00	<b>73</b>	19	00	04	7.1
G	27	00	<b>93</b>	53	07	00	<b>100</b>	00	07	<b>100</b>	<b>93</b>	00	00	00	4.1

(ICd). At any rate, most students in this cluster were able to construct a working albeit incorrect program. Given their unfamiliarity with general central force interactions, these students might have believed their solutions were correct. In fact, it is possible that students who were working with the grading case (ICc and ICg) had solved the test case correctly and simply made a typographical error.

Students in cluster C worked with either the grading or the test case and might have made a number of mistakes with their initial conditions (any IC code except ICa). The dominant error in cluster C was committed by students who computed the magnitude of the net force (FCc) and attempted to update the vector momentum with this scalar force (SLb). This mathematically impossible operation raises a VPython error. Students in this cluster were unable to parse this error into any useful information.

Cluster D (10.8%), like cluster A, was populated by students who tended to make errors in the force calculation (FCa and FCd), but students in cluster D worked with the grading case (ICc). The most common error in cluster D was reversing the direction of the net force (FCa), followed by raising the separation vector to a power (FCd). Again, like cluster A, most students met no challenges when updating the momentum using Newton's second law (SLb). These students might have started working with the test case, but we think it is more likely that they started working with the grading case because the dominant error appears in their force calculations. We have observed a number of students doing so in homework help sessions.

Students in cluster E (7.6%) tended to raise the separation vector to a power (FCd) and have mixed initial conditions (ICb, ICd, ICe, and ICf). These students generally had no difficulty with writing Newton's second law correctly (SLa). The dominant error for students in cluster E was raising the separation vector to a power (FCd). This mathematically impossible operation raises a VPython error. Students in this cluster were unable to parse this error into any useful information.

Cluster F (7.1%) contained students who worked solely with the test case (ICa) and either had no issue with their force calculation (FCb) or had no evident force calculation procedure (FCe). Most of these students had no difficulty updating the momentum using Newton's second law (SLa). Students in cluster F were able to construct a program that ran without raising any VPython errors. Students who had no issue with their solution likely completed the test case but simply ran out of time before turning to the grading case. Students with no evident procedure generally computed the net force outside the numerical calculation loop, essentially making this force constant in time. Given students' unfamiliarity with general central force interactions, it would not be surprising if students who treated the central force outside the loop believed their solutions were correct.

Students in cluster G (4.1%) all invented an incorrect unit vector for the force rather than using  $\hat{r}$  (FCf) regardless of the case with which they worked (ICc, ICb, and ICf). These students generally had no difficulty updating the momentum using Newton's second law (SLa). Most likely, these students computed the magnitude of the force (similar to students in cluster C) but were able to parse the resulting VPython error. Students in cluster G corrected their mistake by assigning some unit vector to the force before the momentum was updated.

## VIII. CLOSING REMARKS

Students with little to no programming experience can, in large introductory physics courses, develop the computational skills necessary to predict the motion of sundry physical systems. After solving a suite of computational homework problems, most students ( $\sim 60\%$ ) were able to model the motion of a novel problem successfully. In our work, we discovered that most students who were unsuccessful encountered challenges when calculating the net force acting on the object in the motion prediction algorithm (clusters A and C through G in Table VI). In contrast, there were fewer students whose primary challenge was identifying and assigning variables (cluster B in Table VI). We acknowledge that we have limited the development of our students' computational skill set to contextualizing a word problem into a programming task, identifying and updating input variables, and applying a motion prediction algorithm. We believe that further development of our homework problems and other novel deployments could broaden the scope of the skills students develop.

Our work provides a touchstone example of computational instruction for the STEM education community. This work provides clear evidence that newcomers to both computer programming and computational modeling can successfully acquire, in a large lecture introductory STEM course, basic skills needed to do computational modeling. However, the success of our students on the proctored assignment should not be overemphasized. Ours is but one of many possible implementations that might be successful in a large lecture setting. Moreover, the level of instruction in this course was limited to the iterative prediction of motion. What is more fascinating are the types of errors which students made when solving the proctored assignment and the commonalities among those errors.

Procedural errors such as those we have documented (Secs. VI and VII) could be corrected by developing additional materials aimed at addressing each error individually. However, the results from this work indicate that instructional efforts should be focused not only on correcting procedural mistakes but also on developing students' qualitative habits of mind. Training students to write programs to predict motion might help them to be successful in a highly structured environment, but they would be

better served by learning the practice of debugging. Here, debugging includes identifying syntax errors (of which we found few) and, more importantly, performing the type of qualitative analysis that is typically taught for solving analytic problems. Students who could synthesize their analytic and computational skills would be better prepared to solve the open-ended problems they will face in their future work. Students' use of debugging is the subject of active investigation by the computer science education community [37–39].

Developing the materials to teach these skills requires an evaluation of how students contextualize computational problems. We do not claim to understand this contextualization presently, although we have been able to glean some suggestive information based on students' errors. Investigating what students think about when solving computational problems requires structured student interviews (i.e., a think-aloud study). In the future, we plan to perform such a study to not only characterize students' abilities to contextualize but also to elucidate the mechanism for some of the errors we reported in Secs. VI and VII. Moreover, this work provides the foundation for others in the STEM community to perform similar investigations.

Research into skill development in math and science has shown strong correlations with student epistemology [40,41]. Epistemology is important because the views that students hold affect how they learn [42] and, ultimately, how successful they are in their science courses [43,44]. It is therefore crucial that we understand students' sentiments about learning a new tool such as computation. Our students expressed anxiety and demonstrated a lack of self-confidence in laboratory instruction and homework help sessions, even with their additional exposure to computational problems and improved scores on handwritten programming questions. We are developing an attitudinal survey aimed at exploring these and other beliefs in detail. Students who learn to use computational modeling and are confident in their abilities will be better prepared to solve challenging problems.

We do not claim to have assessed a transfer of computational knowledge. We designed a set of problems (Sec. III) that students solved over the course of the semester with an eye toward a final assessment of their skills using a novel problem. This problem (Sec. IV) was similar to some of the homework problems students had solved previously. It focused on key skills that we desired students to acquire: contextualizing a problem, identifying and assigning variables in a program, and carrying out the motion prediction algorithm. An evaluation of transfer would require that students apply these computational skills to a different domain (e.g., electromagnetism) or a different task (e.g., open-ended inquiry). Demonstrating transfer of computational knowledge is a necessary step in developing students into flexible problem solvers for the 21st century.

It is the goal of many reforms in physics education to develop students into flexible problem solvers while exploring the practice of science. Teaching computational modeling alongside physics provides support for that effort. Students learn the tools for doing science while developing a qualitative understanding of physical systems, exploring the generality of physics principles, and learning broadly applicable problem-solving methods in computation.

## ACKNOWLEDGMENTS

The authors appreciate the helpful comments from the two anonymous reviewers who helped focus and improve the paper. We would also like to thank Edwin Greco for help in deploying our computational problems, Balachandra Suri for analyzing some of the student-written code, and Scott Douglas for carefully reviewing and editing the manuscript. This work was supported by National Science Foundation's Division of Undergraduate Education (DUE0618519 and DUE0942076).

## APPENDIX: MORE DETAILS ON THE EVALUATION CODES

The codes shown in Table II were developed empirically. The procedure followed an iterative-design approach. We reviewed student work for common errors and devised a rough coding scheme. We then tested the scheme on a new set of student-submitted programs. The scheme was refined and retested. This iterative procedure was repeated several times until we captured the majority of the students' mistakes. Each code is explained in detail below.

### 1. Using the correct given values (IC) codes

We reviewed the variables in each student's program. The default values had to be updated with the values given in the problem statement in the partially completed program. We present the codes used to categorize each student's program with respect to identifying and updating the appropriate initial conditions for their realization.

*IC1—Student used all the correct given values from grading case.* A student must replace the values of all the variables (mass, position, and velocity, interaction constant  $k$  and the exponent in the force law  $n$  in  $F = kr^n$ ) with those given in the *grading case*. This code excluded the integration time. It was intended that the larger mass object was to remain at its location. This was made explicit in the problem statement; the initial position  $\langle 5, 4, 0 \rangle$  m and velocity  $\langle 0, 0, 0 \rangle$  m/s of the larger mass of the object with more mass were given in the problem statement, even though these same values appeared in the partially written program.

*IC2—Student used all the correct given values from test case.* A student must replace the values of all the variables (mass, position, and velocity, interaction constant  $k$  and the exponent in the force law  $n$  in  $F = kr^n$ ) with those given in the *test case* for this code to apply. This code excluded the

integration time. It was intended that the larger mass object was to remain at its location (see IC1).

*IC3—Student used the correct integration time from either the grading case or the test case.* A student must replace the default integration time (1 s) with the values given in the case with which they intended to work (grading or test). A student who mixed initial conditions was given an affirmative on this code if the majority of their initial conditions were from the same case as the integration time.

*IC4—Student used mixed initial conditions.* A student who used some but not all of the initial conditions from any of the cases (default, test, or grading) was given an affirmative on this code. This code excluded the integration time.

*IC5—Students confused the exponents on the units the exponent of  $k$  (interaction constant).* Many students incorrectly thought the exponent on the length unit of the interaction constant was the scientific notation exponent for the interaction constant itself. For example, a student thought  $k = 0.1 \text{ Nm}^3$  meant  $k = 100$  rather than  $k = 0.1 \text{ Newton times meters cubed}$ .

## 2. Implementing the force calculation (FC) codes

We reviewed how the students employed the force calculation algorithm in each of the programs written for the proctored assignment. The partially written program given to the students left out all statements related to the force calculation. Students were required to fill in this procedure using the appropriate VPython syntax. We present the codes used to categorize each student's program with respect to computing the vector force acting on the low-mass object.

*FC1—The force calculation was correct.* For this code to apply, a student must compute the separation vector, its magnitude, its unit vector, the magnitude of the force, and the vector force correctly. Each of these steps may be combined as long as the final result computes the vector force acting on the less massive particle at each instant. These steps must all appear in the numerical integration loop.

*FC2—The force calculation was incorrect, but the calculation procedure was evident.* In the numerical integration loop, the student must perform a position vector subtraction, a calculation of the force magnitude, and some attempt at combining magnitude with unit vector (any unit vector was acceptable) for this code to apply. If a student treated the problem using components and had some force which is a vector, it was coded as evident. If any part of the calculation was performed outside the loop, it was coded as *not* evident.

*FC3—The student attempted to raise the separation vector ( $\vec{r}$ ) to a power.* Students who raised the separation vector to a power generated a VPython exception error: `unsupported operand type(s) for ** or pow(): 'vector' and 'int'`. This error told them that VPython cannot raise a vector to a power, as it is a mathematically impossible operation.

*FC4—The direction of the force was reversed.* Students had to assign the correct unit vector and sign to the force depending on whether their force was attractive or repulsive for this code to be used. This code was not used if the student calculated the force as a magnitude only, raised  $\vec{r}$  to a power, or invented a unit vector (e.g.,  $\langle 1, 0, 0 \rangle$ ). Visual feedback (i.e., the lower mass particle flying off to infinity) indicated a simple sign mistake.

*FC5—Student had some other force direction confusion.* Some students used vectors other than  $\vec{r}$  or  $-\vec{r}$  to compute  $\vec{F}$ . Other students computed the force as a magnitude and then multiplied it by an “invented” unit vector (e.g.,  $\langle 1, 0, 0 \rangle, \hat{p}$ ). Both of these errors were given an affirmative for this code.

## 3. Updating with the Newton's second law (SL) codes

We reviewed how the students employed the momentum update in each of the programs written for the proctored assignment. The partially written program given to the students left out the one line of code necessary to update the momentum. Students were required to fill in this line using the appropriate VPython syntax. We present the codes used to categorize each student's program with respect to updating the momentum of the low-mass particle.

*SL1—Newton's second law was correct.* Correct Newton's second law meant that it was “correct as a physics principle” and also that it appeared “in the update form.” This meant that `pfinal = pinitial + Fnet * deltat` alone in a loop did not fall under “correct Newton's second law.” It is an incorrect update form.

*SL2—Newton's second law was incorrect but in a form that updates.* Newton's second law updates the momentum but not necessarily correctly (e.g., `p = p + Fnet`, `p = p + Fnet/dt`, `pf = p + Fnet`, etc.).

*SL3—Newton's second law was incorrect and the student attempted to update it with a scalar force.* Some students computed the magnitude of the force acting on the particle and then used this magnitude to update the momentum. Students who did this raised a VPython exception error: `unsupported operand type(s) for +: 'vector' and 'int'`. This might have led some to invent unit vectors in the momentum update [e.g., `p = p + vector(Fmag, 0, 0) * dt`].

*SL4—Student created a new variable for  $\vec{p}_f$ .* In computational modeling, the equal sign in a update line means “add and replace.” Some students used a new symbol for the final momentum (e.g., `pfinal`) and then replaced the momentum in the next step (e.g., `p = pfinal`). Others only did the former; that is, they did not replace the momentum with its updated value.

## D. Other codes

Two common errors were not included in the above codes because they do not reflect errors in the procedure of modeling the motion of the low-mass particle.

We present two miscellaneous codes that were common enough to consider relevant.

*O1—Student attempted to update force, momentum or position for the massive particle.* The massive particle was intended to remain in place.

*O2—Student did not attempt the problem.* Some students uploaded blank text files. We assumed they did not attempt solving the problem and intended only to receive bonus credit for uploading their code.

- 
- [1] K. Jordan and C. Woodward, <http://www.siam.org/students/resources/report.php>.
- [2] F. Fenton, E. Cherry, A. Karma, and W. Rappel, Modeling wave propagation in realistic heart geometries using the phase-field method, *Chaos* **15**, 013502 (2005).
- [3] J. González, U. Sperhake, and B. Brügmann, Black-hole binary simulations: The mass ratio 101, *Phys. Rev. D* **79**, 124006 (2009).
- [4] E. Morote and D. Pritchard, What course elements correlate with improvement on tests in introductory Newtonian mechanics?, *Am. J. Phys.* **77**, 746 (2009).
- [5] W. M. MacDonald, E. F. Redish, and J. M. Wilson, The M.U.P.P.E.T. manifesto, *Comput. Phys.* **2**, 23 (1988) [<http://dl.acm.org/citation.cfm?id=47537.47539>].
- [6] E. Redish and J. Wilson, Student programming in the introductory physics course: MUPPET, *Am. J. Phys.* **61**, 222 (1993).
- [7] H. Schecker, Learning physics by making models, *Phys. Educ.* **28**, 102 (1993).
- [8] H. Schecker, in *Proceedings of the 1994 International System Dynamics Conference, Stirling, Scotland* (System Dynamics Society, Albany, NY, 1994), pp. 74–84.
- [9] A. DiSessa and H. Abelson, Boxer: A reconstructible computational medium, *Commun. ACM* **29**, 859 (1986).
- [10] B. Sherin, A. diSessa, and D. Hammer, Dynaturtle revisited: Learning physics through collaborative design of a computer model, *Interact. Learn. Environ.* **3**, 91 (1993).
- [11] N. Finkelstein, W. Adams, C. Keller, K. Perkins, and C. Wieman, High-tech tools for teaching physics: The physics education technology project, *J. Online Learn. Teach.* **2**, 110 (2006).
- [12] R. Chabay and B. Sherwood, Computational physics in the introductory calculus based course, *Am. J. Phys.* **76**, 307 (2008).
- [13] R. Beichner, R. Chabay, and B. Sherwood, Labs for the Matter & Interactions curriculum, *Am. J. Phys.* **78**, 456 (2010).
- [14] B. A. Sherwood and R. Chabay, Vpython–3D programming for ordinary mortals, <http://www.vpython.org/> (2011).
- [15] W. Christian and A. Titus, Developing web-based curricula using java physlets, *Comput. Phys.* **12**, 227 (1998).
- [16] K. Perkins, W. Adams, M. Dubson, N. Finkelstein, S. Reid, C. Wieman, and R. LeMaster, PhET: Interactive simulations for teaching and learning physics, *Phys. Teach.* **44**, 18 (2006).
- [17] C. Wieman and K. Perkins, A powerful tool for teaching science, *Nature Phys.* **2**, 290 (2006).
- [18] F. Esquembre, Easy Java simulations: A software tool to create scientific simulations in Java, *Comput. Phys. Commun.* **156**, 199 (2004).
- [19] W. Christian and F. Esquembre, Modeling physics with easy Java simulations, *Phys. Teach.* **45**, 475 (2007).
- [20] M. Belloni and W. Christian, Proceedings of the 2007 AAPT Topical Conference on Computational Physics (ComPADRE 2007), Davidson, NC, 2007 (unpublished).
- [21] F. Esquembre, Proceedings of the 2007 AAPT Topical Conference on Computational Physics (ComPADRE 2007), Davidson, NC, 2007 (Ref. [20]).
- [22] D. Scherer, P. Dubois, and B. Sherwood, Vpython: 3D interactive scientific graphics for students, *Comput. Sci. Eng.* **2**, 56 (2000).
- [23] R. Chabay and B. Sherwood, *Matter and Interactions I: Modern Mechanics* (John Wiley and Sons, New York, 2010), 3rd ed.
- [24] This data is self-reported. Students were polled in class using a “clicker” question.
- [25] M. Kohlmyer, Ph.D. thesis, Carnegie Mellon University, 2005.
- [26] W. Conover, *Practical Nonparametric Statistics* (John Wiley and Sons, New York, 1999).
- [27] The first administration of this assignment was during a regular hour exam. Roughly 40% of the students modeled the motion correctly. However, students used their own laptop computers which created several logistical challenges.
- [28] W. Press, W. Vetterling, S. Teukolsky, and B. Flannery, *Numerical Recipes in C, The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1995).
- [29] R. Knight, The vector knowledge of beginning physics students, *Phys. Teach.* **33**, 74 (1995).
- [30] P. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining* (Pearson, Boston, 2006).
- [31] B. Everitt, S. Landau, and L. M., *Cluster Analysis* (Oxford University Press, New York, 2001).
- [32] R. Springuel, M. Wittmann, and J. Thompson, Applying clustering to statistical analysis of student reasoning about two-dimensional kinematics, *Phys. Rev. ST Phys. Educ. Res.* **3**, 20107 (2007).
- [33] P. Jaccard, Study of comparative distribution of flower in the portion of alpes and jura, *Bull. Soc. Vaudoise Sci. Nat.* **37**, 547 (1901).

- [34] R. R. Sokal and C. D. Michener, A statistical method for evaluating systematic relationships, *Univ. Kans. Sci. Bull.* **38**, 1409 (1975).
- [35] P. Sneath, The application of computers to taxonomy, *Microbiology* **17**, 201 (1957).
- [36] T. Sorenson, A method of establishing groups of equal amplitude in plant sociology based on similarity of species content, *Kongl. D. Vid. Selskab* **5**, 1 (1948).
- [37] R. McCauley, S. Fitzgerald, G. Lewandowski, L. Murphy, B. Simon, L. Thomas, and C. Zander, Debugging: A review of the literature from an educational perspective, *Comput. Sci. Educ.* **18**, 67 (2008).
- [38] S. Fitzgerald, G. Lewandowski, R. McCauley, L. Murphy, B. Simon, L. Thomas, and C. Zander, Debugging: Finding, fixing and flailing, a multi-institutional study of novice debuggers, *Comput. Sci. Educ.* **18**, 93 (2008).
- [39] L. Murphy, G. Lewandowski, R. McCauley, B. Simon, L. Thomas, and C. Zander, Debugging: The good, the bad, and the quirky—A qualitative analysis of novices' strategies, in *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education* (ACM, New York, 2008), Vol. 40, pp. 163–167, doi [10.1145/1352135.1352191](https://doi.org/10.1145/1352135.1352191).
- [40] M. Schommer, Effects of beliefs about the nature of knowledge on comprehension., *J. Educ. Psychol.* **82**, 498 (1990).
- [41] M. Schommer, Epistemological development and academic performance among secondary students, *J. Educ. Psychol.* **85**, 406 (1993).
- [42] A. Elby, Helping physics students learn how to learn, *Am. J. Phys.* **69**, S54 (2001).
- [43] I. Halloun, *AIP Conf. Proc.* **399**, 605 (1997).
- [44] K. Perkins, W. Adams, S. Pollock, N. Finkelstein, and C. Wieman, *AIP Conf. Proc.* **790**, 61 (2005).