

Systematic study of student understanding of the relationships between the directions of force, velocity, and acceleration in one dimension

Rebecca Rosenblatt and Andrew F. Heckler

Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

(Received 24 June 2011; published 2 November 2011)

We developed an instrument to systematically investigate student conceptual understanding of the relationships between the directions of net force, velocity, and acceleration in one dimension and report on data collected on the final version of the instrument from over 650 students. Unlike previous work, we simultaneously studied all six possible conditional relations between force, velocity, and acceleration in order to obtain a coherent picture of student understanding of the relations between all three concepts. We present a variety of evidence demonstrating the validity and reliability of the instrument. An analysis of student responses from three different course levels revealed three main findings. First, a significant fraction of students chose “partially correct” responses, and from pre- to post-test, many students moved from “misconception” to partially correct responses, or from partially correct to fully correct responses. Second, there were asymmetries in responding to conditional relations. For example, students answered questions of the form “Given the velocity, what can be inferred about the net force?” differently than converse questions “Given the net force, what can be inferred about the velocity?” Third, there was evidence of hierarchies in student responses, suggesting, for example, that understanding the relation between velocity and acceleration is necessary for understanding the relation between velocity and force, but the converse is not true. Finally, we briefly discuss how these findings might be applied to instruction.

DOI: 10.1103/PhysRevSTPER.7.020112

PACS numbers: 01.40.Fk, 01.50.Kw, 01.40.Ha

I. INTRODUCTION

One of the earliest and most studied areas in physics education research is student understanding of force, velocity, and acceleration. For example, perhaps the most widely known and documented phenomenon in this field is the (incorrect) student belief that the net force on an object and its velocity must be in the same direction [1–5]. It is also well documented that students often have difficulty distinguishing between the velocity and acceleration of an object [6,7].

Nonetheless, even though this topic is relatively well studied, there remain many unanswered questions that are critical to both advancing our knowledge of student difficulties with force, velocity, and acceleration and applying this knowledge to improve student learning of these fundamental concepts. For example, empirically speaking, to what extent does the correct understanding of the relationship between, say, force and acceleration depend on the correct understanding of another relation, say, between force and velocity? Does the path to correct understanding of these relations empirically occur in steps? If so, what are the steps?

Furthermore, it is important to point out that when assessing student understanding of the relations between

force, velocity, and acceleration, the questions posed typically involve *conditional relations*, though this has not been explicitly acknowledged or systematically studied in previous work. For example, in a landmark paper, Viennot posed questions of the form “given the velocity of an object, what is the (net) force on the object?” which is a conditional relation of the form “given x , what is y ?” There were no questions in Viennot’s study probing the converse conditional relation “given a net force on an object, what is its velocity?” Nor were there any questions regarding the relations between velocity and acceleration or acceleration and force [4]. Certainly, in other studies that followed Viennot’s paper, other conditional relationships were studied. However, as can be seen from Table I, which summarizes the relationships studied in many of the existing research papers on students’ conceptual understanding of the directional relationships of force, velocity, and acceleration, there has been no systematic study of student understanding of all six possible paired conditional relations between the concepts of force, velocity, and acceleration. Furthermore, there has been an abundance of work on some of the six relations and little, if any, on others.

A systematic study of all possible pairs of conditional relations between force, velocity, and acceleration is important for two reasons. First, a within-student study of all possible pairs of relations will allow for a more holistic picture of student understanding of all relations and the possibility of determining whether understanding one relation may effect (or predict) the understanding of another relation. Second, it is not unreasonable to expect that for a

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

TABLE I. Question types investigated in a sample of previous studies. ($\vec{x} \rightarrow \vec{y}$ notation indicates a question of the form: Given \vec{x} , what can be inferred about \vec{y} ?)

Number of questions	Question type	Citation of study
4	$\vec{v} \rightarrow \vec{F}$	Viennot [4]
2	$(\Delta \vec{v}, dt) \rightarrow \vec{a}$	Trowbridge and McDermott [6]
2	$\vec{v} \rightarrow \vec{F}$	Clement [3]
2	$\vec{F} \rightarrow \vec{v}$	Clement [3]
21	$\vec{v} \rightarrow \vec{a}^a$	Reif and Allen [7]
1	$\vec{a} \rightarrow \vec{v}$	Reif and Allen [7]
7	$\vec{v} \rightarrow \vec{F}$	FCI test: Hestenes, Wells, Swackhammer [5]
9	$\vec{F} \rightarrow \vec{v}, \Delta \vec{v}^b$	FCI test: Hestenes, Wells, Swackhammer [5]
4	$\vec{F} \rightarrow \Delta \vec{v}$	FCI test: Hestenes, Wells, Swackhammer [5]
4	$\vec{v} \rightarrow \vec{F}$	Enderstein and Spargo [8]
8	$\vec{v} \rightarrow \vec{F}$	Palmer [9]
2	$\vec{F} \rightarrow \vec{v}$	Enderstein and Spargo [10]
4	$\vec{v} \rightarrow \vec{F}$	Enderstein and Spargo [10]
12	$\vec{v} \rightarrow \vec{F}^c$	FMCE test: Thornton and Sokoloff [2]
1	$\vec{a} \rightarrow \vec{F}$	FMCE test: Thornton and Sokoloff [2]
3	$\vec{v} \rightarrow \vec{a}$	FMCE test: Thornton and Sokoloff [2]

^aIn the problem set up, many of the questions specify the speed in addition to the direction of the velocity.

^bIn the problem set up, five of the questions specify a force, two of the questions imply, but do not specify, gravity, and two of the questions imply zero force.

^cIn the problem set up, six of the questions specify the speed in addition to the direction of the velocity and six do not.

given pair of variables, a conditional relation between the pair and its converse may not be answered similarly by the student. For example, the question “An object is accelerating in a certain direction, what can you infer about the object’s velocity?” may be answered differently than the question “An object has a velocity in a certain direction, what can you infer about the object’s acceleration?” Furthermore, if there is a causal relation between the variables (real or believed), such as between force and acceleration, then making inferences about the effect of a given cause may be different than making inferences about the cause of a given effect [11].

Therefore, in this paper we will investigate student understanding of all possible pairs of relations between force, velocity, and acceleration. To more precisely focus the investigation, we will only study student understanding of the relations between the *directions* of force, velocity, and acceleration in one dimension, and leave the investigation of multiple dimensions and the relations between the magnitudes of these variables for other studies.

While this investigation included a significant amount of student interviews and open-ended written answers, the bulk of the analysis is based on a multiple-choice test that we developed for this study. The multiple-choice test allows for, in principle, the identification of reliable patterns based on a large number of students. On the other hand, such a test can lack the subtlety and depth compared to a more qualitative study; nonetheless, the validity and reliability of the results claimed here were corroborated by the interviews and written answers of students. Clearly an

in-depth study using more qualitative data would also yield interesting results, but here we focus on some of the important, replicable patterns found via the carefully constructed instrument.

Finally, we have one further introductory comment before proceeding. In a relatively recent study, Alonso and Steedle [12] have investigated middle-school student (12–14 years old) understanding of force and motion. They hypothesize increasingly expertlike levels of understanding of force and motion through which middle-school students pass in a progression towards mastery of these concepts. Specifically, they construct a formal “learning progression” of force and motion for this population. The topic of learning progressions has recently generated significant interest in the science education community (e.g., see [13]) and is somewhat relevant to the study in this paper since we examine longitudinal and cross-sectional data on student performance and we are interested in the steps and hierarchies in understanding the relation between the directions of force, velocity, and acceleration. While the topic of learning progressions is not the focus of this paper, we will briefly comment on this topic and Alonso and Steedle’s study in the final discussion section.

The paper proceeds as follows. We first briefly describe the careful construction of the short, multiple-choice assessment instrument and report on its validity and reliability. Next we present test results pre- and post-instruction, and results of students at different levels of physics knowledge. These results include an analysis of within- and between-student answering patterns for all six conditional

relations and how answering patterns change both over one course and from first- to second-year university physics students. Finally, we summarize and discuss how the findings might be applied to the design of instruction aimed at improving student understanding of the relations between the directions of force, velocity, and acceleration.

II. FORCE, VELOCITY, AND ACCELERATION DIRECTION ASSESSMENT INSTRUMENT

A. Development of assessment

We constructed a 17-item multiple-choice test, called the FVA test, designed to assess student understanding of all six conditional relationships between the directions of force, velocity, and acceleration in one dimension. Each item in the test presents a simple scenario indicating the direction of one of the vectors for an object, say, acceleration \vec{a} , and asks what this implies about the direction of one of the other vectors, say, velocity \vec{v} . We label such a question as $\vec{a} \rightarrow \vec{v}$, which briefly means “given the acceleration, what can be inferred about the velocity?” (See Table II for examples of an $\vec{a} \rightarrow \vec{v}$ and an $\vec{F} \rightarrow \vec{v}$ question.) Ten of the 17 items include two questions for each of the conditional relations $\vec{F} \rightarrow \vec{v}$, $\vec{v} \rightarrow \vec{F}$, $\vec{a} \rightarrow \vec{v}$, and $\vec{v} \rightarrow \vec{a}$ and one question each for $\vec{a} \rightarrow \vec{F}$ and $\vec{F} \rightarrow \vec{a}$. These

10 items directly probe the six conditional relations between force, velocity, and acceleration, which are of particular interest in this paper.

The specific results of the remaining seven items also provide additional interesting information. However, except for being part of the reported total score results and item statistics of the FVA test, a detailed analysis of response patterns from these seven items are not reported here as they are not the focus of this paper. Nonetheless, it is worth mentioning that these seven items were included in the FVA test for several reasons. First, they provide variety in answering, so that the correct answer is not always “a, b, and c are possible” (see Table III), which is the case for the eight $\vec{v} \rightarrow \vec{F}$, $\vec{F} \rightarrow \vec{v}$, $\vec{a} \rightarrow \vec{v}$, $\vec{v} \rightarrow \vec{a}$ items. We have found that if the answer choice is always (or often) the same, then students start thinking more about the “tricks” of the item format rather than the content of the question. Second, these items probe different aspects of understanding the directional relations of force, velocity, and acceleration and as such are part of a more valid and reliable FVA test. For example, two of the seven items provide situations in which an object is explicitly at rest (item 12) or has zero net forces acting on it (item 3). Furthermore, three of the items (2, 7, and 8) provide (or ask for) information of both the velocity and the change in

TABLE II. Explanation and examples of the $\vec{x} \rightarrow \vec{y}$ notation. A $\vec{x} \rightarrow \vec{y}$ question is designed to probe a student’s understanding of how a given vector’s direction \vec{x} is related to another vector’s direction \vec{y} . For example, an $\vec{a} \rightarrow \vec{v}$ question provides a simple scenario indicating the direction of the acceleration on an object and asks the student what this implies about the direction of the object’s velocity. Two specific examples from the developed test are provided below.

Example of $\vec{a} \rightarrow \vec{v}$ question: “A car is on a hill and the direction of its acceleration is uphill. Which statement best describes the motion of the car at that time?”	Example of $\vec{F} \rightarrow \vec{v}$ question: “At a particular instant of time, there are several forces acting on an object in both the positive and negative direction, but the forces in the negative direction (to the left) are greater. Which statement best describes the motion of the object at this instant?”
a) it is moving uphill	a) it is moving to the right
b) it is moving downhill	b) it is moving to the left
c) it is not moving	c) it is not moving
d) both a and b are possible	d) both a and b are possible
e) both a and c are possible	e) both b and c are possible
f) a, b, and c are possible	f) a, b, and c are possible

TABLE III. Available student response choices for each question. Almost all possible choices for the relationship between two vectors are available responses for a student to choose from when answering a question. The other possible choices were almost never chosen, thus excluded. Consider the $\vec{a} \rightarrow \vec{v}$ question as an example: “A car is on a hill and the direction of its acceleration is uphill. Which statement best describes the motion of the car at this time?”

Response choices	Symbolic representation of choices	Description of most common choices
a) it is moving uphill	$\vec{a} \uparrow \uparrow \vec{v}$	Common “misconception”
b) it is moving downhill	$\vec{a} \uparrow \downarrow \vec{v}$	
c) it is not moving	$\vec{a} \uparrow 0\vec{v}$	(v) “cannot-be-zero” (partially correct)
d) both a and b are possible	$\vec{a}(\uparrow, \downarrow)\vec{v}$	
e) both a and c are possible	$\vec{a}(\uparrow, \uparrow 0)\vec{v}$	(v) “cannot-be-opposite” (to a) (partially correct)
f) a, b, and c are possible	$\vec{a}(\uparrow, \downarrow, \uparrow 0)\vec{v}$	
		Correct

speed. Finally, two of the items (2 and 15) are very familiar and easy, and help to establish a baseline of student understanding. Detailed analysis of response patterns on these items is a topic of further study. The complete instrument is reproduced in the appendix.

Item construction occurred over a period of several years, beginning with open-ended pencil and paper questions and over 40 individual student interviews in a think-aloud format. This was followed up by over 60 individual debriefings for students completing the final versions of the FVA test. The test development involved two major iterations, as explained in more detail in Ref. [14]. This process revealed that there were several possible student response choices for the questions posed, and it was important to include all of these possibilities as response choices in the multiple-choice format.

Specifically, for one-dimensional motion there are three ways in which any one of the three relevant vectors (net force, velocity, or acceleration) can be related to any other: the vectors can be aligned (symbolized as $\uparrow\uparrow$), opposite to one another ($\uparrow\downarrow$), or one of them can be zero ($\uparrow 0$). Using acceleration and velocity to illustrate the seven possible combinations of these relationships yields: $\vec{a} \uparrow\uparrow \vec{v}$ (\vec{v} must be parallel to \vec{a}); $\vec{a} \uparrow\downarrow \vec{v}$ (\vec{v} must be antiparallel to \vec{a}); $\vec{a} \uparrow 0\vec{v}$ (\vec{v} is zero to a nonzero \vec{a}); $\vec{a}(\uparrow\uparrow, \uparrow 0)\vec{v}$ (\vec{v} can be zero or parallel to \vec{a}); $\vec{a}(\uparrow\uparrow, \uparrow\downarrow)\vec{v}$ (\vec{v} can be parallel or antiparallel to \vec{a}); $\vec{a}(\uparrow\downarrow, \uparrow 0)\vec{v}$ (\vec{v} can be antiparallel to \vec{a} or zero); $\vec{a}(\uparrow\downarrow, \uparrow\downarrow)\vec{v}$ (\vec{v} can be parallel or antiparallel to \vec{a} or zero).

Of these seven possible combinations, we found that students rarely if ever considered the physically unnatural possibility of “can only be opposite or zero”; thus, usually only six response choices were provided. Table III provides an example of an item and the six possible response choice “models.”

B. Reliability and validity of the FVA test

The construct validity of the items, including the question and answer choice format, was supported through several stages of interview and testing-based modifications, as reported above. We report here on other measures of validity and reliability, including correlations of the test with other measures such as course level, course grade, and the Force Concept Inventory (FCI) (all measures of student knowledge) as well as various reliability measures of the instrument. Finally, we determine the extent to which the story context of the questions might affect student response choices.

1. Increases in FVA score with increasing course level and instruction

We administered the FVA test post-instruction to four different class levels—standard calculus mechanics, honors calculus mechanics for engineering majors, honors calculus mechanics for first-year physics majors, and mechanics for second-year physics majors. See Table IV for a description of each course and the enrolled students. Figure 1 reveals that the average score on the FVA test tended to increase with course level such that the average score of the second-year physics majors was 0.9 standard deviations above that of the standard mechanics course. One exception was the higher score for the first-year honors physics majors compared to the second-year majors course. This difference may be due to slightly different population, since the second-year course enrolls some student who are not physics majors and not honors students. Note also that the increase in average score with course level was not an artifact of an increase in score on a small number of questions; rather the increase was spread

TABLE IV. Outline of test administration and description of courses and populations.

Course Name (Calculus Mechanics)	Quarter/ Year	Number of students	Type of testing done	Description of course and population
Standard	Winter/2008	117	Post-FVA	Traditional lecture and recitation format.
	Winter/2009	111	Post-FVA and FCI	Majority are first- or second-year science or engineering students.
	2005–2008	>500	Test development	Students' activities include show work, free response, interviews, and multiple choice data.
Honors engineers	August/2008	86	Post-FVA	Lecture and recitation format with some Physics Education Research reforms included, e.g., clickers in lecture and group work problem solving in recitation. All first-year honors engineering students.
	August/2009	230	Pre- and Post-FVA	
Honors physics majors	August/2009	49	Pre- and Post-FVA	Traditional lecture. Majority are first-year honors physics majors.
Second-year physics majors	Winter/2009	65	Post-FVA	Traditional lecture. This is the second quarter of the second-year physics majors' sequence. Majority are second-year physics majors, with some engineering majors as well.

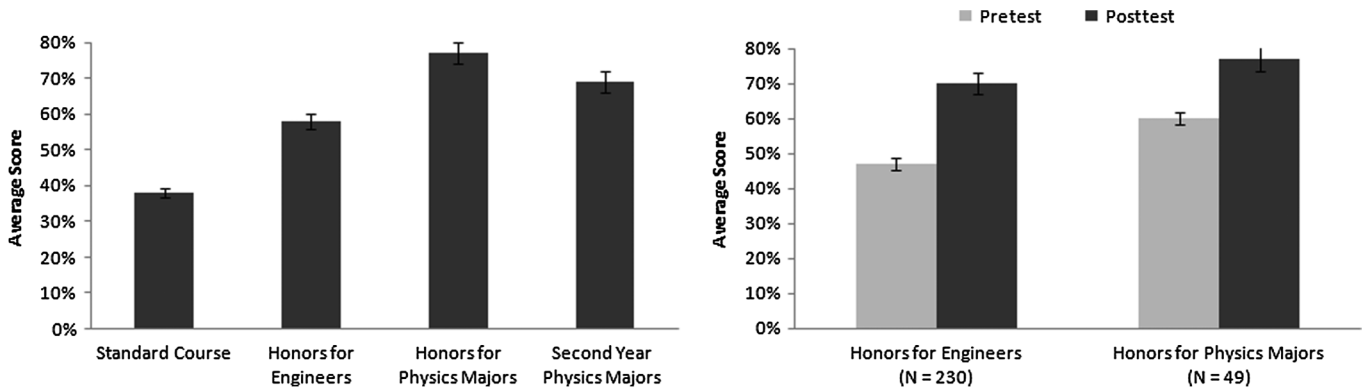


FIG. 1. Summary of FVA scores. The graph on the left shows a trend of increasing post-test score with increasing class level for four different calculus mechanics courses ranging from first-year introductory to second-year physics majors. The graph on the right shows pre- and post-test scores for the two different honors introductory mechanics courses.

among all the question types. Similarly, for all questions, the percentage of “misconception” responses decreased as class level increased. Thus, the average student in the higher level class did better on the FVA test, by both decreasing his or her misconception responses and increasing his or her correct responses, than the average student in the lower level course.

In addition to measuring post-test score differences between different course levels, we also administered pre- and post-tests to measure any changes in scores within a given course. For two courses, Fig. 1 reveals a within-student pre- to post-test increase in correct responses. In particular, a paired *t* test reveals significant gains from pre- to post-test for the honors engineering introductory calculus mechanics class [$t(229) = 16.50, p < 0.001, \text{effect size } d = 0.95$] and for the honors physics majors introductory calculus mechanics class [$t(48) = 7.13, p < 0.001, \text{effect size } d = 0.71$]. Furthermore, it is interesting to note that we pre- and post-tested students in the standard calculus mechanics course with a very similar, earlier version of the FVA test and found no significant difference between pre- and post-test averages. This suggests that for the standard calculus course there may be little gain or evolution in the concepts the FVA test assesses. This lack of significant gain in the traditional course is consistent with previous research in student conceptual understanding of force and motion [1–5].

2. Psychometric properties and correlations with course grade, course level, and FCI

Table V reports overall test statistics for several class levels and Table VI reports individual item statistics for several class levels as well. The FVA test has a reasonably high Kuder-Richardson reliability coefficient, KR-20 = 0.7–0.85, indicating that the correct responses to all of the FVA items are fairly well correlated. Furthermore, there were moderate (0.3–0.4) correlations between FVA score and final course grade. Likewise, the FVA misconception responses were negatively correlated with grade in the class and were on average about -0.4 . These correlations tended to be larger for the higher level classes. These data are consistent with the FVA test assessing a portion of the skills necessary to do well in the class. The correlations between FVA score with course level and the gains, pretest to post-test, suggest that the FVA final-grade correlation is not simply caused by something more general such as intelligence but rather by gained knowledge of force, velocity, and acceleration.

Furthermore, we administered the Force Concept Inventory to the Winter 2009 calculus mechanics class in order to compare the FVA test to a standard benchmark and further assess the validity of the FVA. The FCI is a multiple-choice concept inventory developed to assess understanding of basic concepts in force and motion. It has been widely used and generally accepted as a standard

TABLE V. Summary of FVA test statistics.

Course	<i>N</i>	Post-test FVA avg ± SD	KR-20	Correlation Grade-score	Correlation Grade-misconception score
Standard Mechanics	228	37.6 ± 17.7%	0.721	0.311	-0.286
Honors engineers	86	57.9 ± 21.4%	0.801	0.428	-0.508
Second-year majors	65	69.2 ± 22.8%	0.849
Honors engineers	230	70 ± 24.0%	0.872	0.464	-0.436

TABLE VI. Summary of individual FVA test item statistics for three course levels. Reported are the response percentages for each available response on all 17 items. Correct responses choices are in bold. (Note that question 9 required students to “circle all that apply.” Thus, response percentages reflect the percentage of students circling that response. The correct response was circling a, b, and d.) Pt-Bis. is point biserial coefficient.

Item	Question type	Standard Introductory Mechanics							Honors Mechanics for engineers							Second-year Mechanics for physics majors									
		Pt-Bis.	a	b	c	d	e	f	g	Pt-Bis.	a	b	c	d	e	f	g	Pt-Bis.	a	b	c	d	e	f	g
1	$\vec{v} \rightarrow \vec{F}$	0.377	56	4	4	4	12	21	...	0.427	23	1	1	2	30	42	...	0.486	22	3	3	0	12	60	...
2	$\vec{v}, \Delta\vec{v} \rightarrow \vec{a}$	0.151	88	2	0	7	3	0	...	-0.033	93	1	0	2	1	2	...	0.075	92	0	0	8	0	0	...
3	$\vec{F} = 0 \rightarrow \vec{v}$	0.317	1	1	82	1	15	0.628	0	0	47	2	51	0.475	0	0	29	8	63
4	$\vec{a} \rightarrow \vec{v}$	0.299	41	2	0	23	10	24	...	0.291	27	0	0	35	8	30	...	0.523	17	0	0	19	5	60	...
5	$\vec{f} \rightarrow \vec{a}$	0.162	61	3	3	5	17	11	...	0.196	88	1	1	1	6	2	...	0.562	82	0	0	5	6	7	...
6	$\vec{v} \rightarrow \vec{a}$	0.410	11	1	1	6	19	62	...	0.546	1	1	2	1	16	78	...	0.421	2	0	2	0	8	90	...
7	$\vec{v}, \Delta\vec{v} \rightarrow \vec{F}$	0.312	42	39	4	6	4	2	4	0.284	16	78	0	1	0	2	2	0.529	9	72	3	0	3	3	9
8	$\vec{a} \rightarrow \vec{v}, \Delta\vec{v}$	0.427	19	2	1	23	43	12	...	0.544	20	0	1	9	61	9	...	0.507	5	0	0	6	69	20	...
9	$\vec{v} \rightarrow \vec{F}$ -list	0.279	95	93	57	93	0	0.275	100	100	12	99	0	0.522	100	100	16	100	0
10	$\vec{F} \rightarrow \vec{v}$	0.402	62	2	0	11	10	14	...	0.562	37	4	1	26	0	32	...	0.634	23	2	2	17	2	55	...
11	$\vec{a} \rightarrow \vec{F}$	0.163	84	4	0	5	4	3	...	0.238	79	1	1	9	4	5	...	0.427	86	0	2	5	5	3	...
12	$\vec{v} = 0 \rightarrow \vec{F}$	0.437	2	30	60	1	1	5	1	0.539	1	60	30	0	0	4	4	0.516	3	68	19	0	0	8	3
13	$\vec{a} \rightarrow \vec{v}$	0.405	15	20	0	21	2	5	36	0.527	11	2	2	30	6	1	48	0.673	6	0	2	17	3	0	72
14	$\vec{F} \rightarrow \vec{v}$	0.632	4	52	1	16	2	25	...	0.625	4	16	0	28	4	48	...	0.704	0	11	3	8	2	76	...
15	$\vec{a} \rightarrow \Delta\vec{v}$	0.199	8	1	91	0.169	5	1	92	0.337	2	0	98
16	$\vec{v} \rightarrow \vec{F}$	0.387	73	4	0	2	11	10	...	0.526	34	4	1	6	21	34	...	0.427	24	0	3	2	16	56	...
17	$\vec{v} \rightarrow \vec{a}$	0.380	8	2	4	9	23	54	...	0.532	2	0	1	5	20	71	...	0.585	0	0	6	2	8	84	...

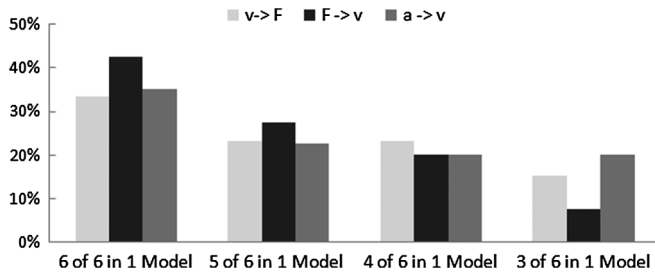


FIG. 2. Percentage of students in the Standard Mechanics course who responded using only one model—either correct, cannot-be-zero, cannot-be-opposite, or misconception—for 3, 4, 5, or 6 out of the 6 questions of the same question type— $\vec{v} \rightarrow \vec{F}$, $\vec{F} \rightarrow \vec{v}$, or $\vec{a} \rightarrow \vec{v}$ —on the “multiple context” tests. Note that a majority of students used only one model for 5 or 6 out of 6 questions.

and reasonably reliable assessment, and has also been used to evaluate instructional interventions at the high school and university level [1,15]. We found a relatively strong correlation between FVA score and FCI score ($r = 0.569$), while the correlation of FCI with final grade was 0.387, about the same as the FVA test-final grade correlation.

In summary, the positive correlations of FVA score with other measures (or expectations) of force and motion conceptual understanding such as course level, pre or post, grade, and FCI score help to support the validity of the FVA test.

3. Effect of story context on responses

One significant threat to the validity of a particular item is its potential sensitivity to construct irrelevant changes to the item. Thus far, we have only addressed issues of potential sensitivity to the item structure and format. Here we would like to address the issue of potential sensitivity to the story context of the item. For example, a force and motion question about a playground ball might be regarded differently by the student compared to an analogous question (from the perspective of the expert) about satellites in space. In order to limit the test to a reasonable length, the FVA test has at most two different story contexts for each question category, $\vec{F} \rightarrow \vec{v}$, $\vec{a} \rightarrow \vec{v}$, $\vec{v} \rightarrow \vec{F}$, etc. Therefore, if the effects of story context are significant, this could severely limit the generalizability of any

conclusions based on student response patterns in the FVA test. We constructed a series of tests and analyzed the results in two ways to investigate the possibility that our results were simply an artifact of story context.

We constructed and administered three separate “multiple context” tests to assess consistency of responses across a variety of story contexts for each of the three major question categories with which students have the most difficulty. Specifically, each multiple context test consisted of ten questions, six of which were all either $\vec{a} \rightarrow \vec{v}$, $\vec{F} \rightarrow \vec{v}$, or $\vec{v} \rightarrow \vec{F}$, and, for variety, four of which were $\vec{a} \rightarrow \vec{F}$ or $\vec{F} \rightarrow \vec{a}$ questions. Students were randomly assigned to complete one of the three multiple context tests, with 40 students in the standard mechanics course per test. We analyzed the results in two ways.

First, we analyzed the data to determine whether there were consistent within-student response patterns for a given question category. We found that on average across the three tests (see Fig. 2) 37% of students consistently (within-student) chose the same answer choice for all six of the questions, and 61% of students answered at least five of the six questions with the same answer choice (within-student).

It is also worth noting that each of the major answer choice models corresponding to correct, misconception, cannot-be-zero, and cannot-be-opposite were consistently answered on all or five out of six questions by at least some students. This suggests that these four answer choices were not just random distractors that were occasionally attractive to the student for certain question contexts, rather, they were consistently chosen. In contrast, on the regular FVA test, only 3% of students consistently (within-student) answered all six of the $\vec{a} \rightarrow \vec{v}$, $\vec{F} \rightarrow \vec{v}$, and $\vec{v} \rightarrow \vec{F}$ questions on the FVA test with the same answer choice and only 24% used the same answer choice on five of these six questions. Overall, these results suggest that for a given question type, for a variety of story contexts, within-student responses follow a specific model, such as cannot-be-opposite, but students do not necessarily use this model for other question types.

Second, we compared the answering patterns in the multiple context tests with the answering patterns in the FVA test to determine whether the response patterns for

TABLE VII. Comparison of average response choice percentages for the FVA test and the “multiple context” tests (which include six different story contexts) for each question type. Reported are averages \pm standard error between the questions.

Question type	Correct	Cannot-be-zero	Cannot-be-opposite	“Misconception”
$\vec{v} \rightarrow \vec{F}$	22 \pm 2.4%	7 \pm 1.7%	7 \pm 1.1%	60 \pm 2.9%
FVA test	15 \pm 2.7%	3 \pm 1.6%	11 \pm 2.9%	63 \pm 4.4%
$\vec{F} \rightarrow \vec{v}$	21 \pm 3.5%	12 \pm 1.4%	7 \pm 1.9%	58 \pm 4.9%
FVA test	19 \pm 3.1%	13 \pm 3.1%	6 \pm 1.8%	56 \pm 4.6%
$\vec{a} \rightarrow \vec{v}$	38 \pm 1.9%	24 \pm 2.0%	4 \pm 1.1%	34 \pm 4.7%
FVA test	30 \pm 3.4%	22 \pm 3.9%	8 \pm 2.4%	35 \pm 4.4%

corresponding question types agree. Table VII reports the average percentages and standard errors across the six different questions for each test and compares it to the averages from the two questions used for each question type on the FVA test. Both inspection of Table VII and a χ^2 test for independence reveals that there are no significant differences between answer patterns on the multiple context tests that focus on one question category, and the answer patterns on corresponding questions in the FVA test [$\chi^2(3) = 3.47$, $p = 0.325$ for $\vec{v} \rightarrow \vec{F}$ questions, $\chi^2(3) = 0.11$, $p = 0.991$ for $\vec{F} \rightarrow \vec{v}$ questions, and $\chi^2(3) = 1.05$, $p = 0.789$ for $\vec{a} \rightarrow \vec{v}$ questions].

In summary, the results of both kinds of analysis of the focused tests reveal that the *averaged* FVA test responses for each question type are relatively insensitive to story context and in that sense the results are fairly generalizable.

III. ANALYSIS OF FVA TEST RESULTS

The previous two sections have focused on the development and validation of the FVA test. The rest of the paper focuses on analyzing pre- and post-FVA test data from students enrolled in different levels of physics courses.

This analysis will allow for an investigation of possible structure and hierarchy of student understanding of the relations between the directions of force, velocity, and acceleration as well as an investigation of evolution of this understanding.

The FVA test was administered to students either during an extra session (counted as part of the total homework grade with full credit for participation) in which students came to our lab to complete the test or as an in-class activity completed during the regular laboratory or lecture for the course. In both situations, students had plenty of time to finish the quiz and appeared to take the activity seriously.

A. General response patterns for different course levels

Figure 3 presents average student response patterns for all six question types for three class levels. There are four important observations about the response patterns presented in Fig. 3, as described below.

B. Evidence of intermediate levels of understanding

There was a small but significant fraction of students (20%–30%) who displayed intermediate levels of

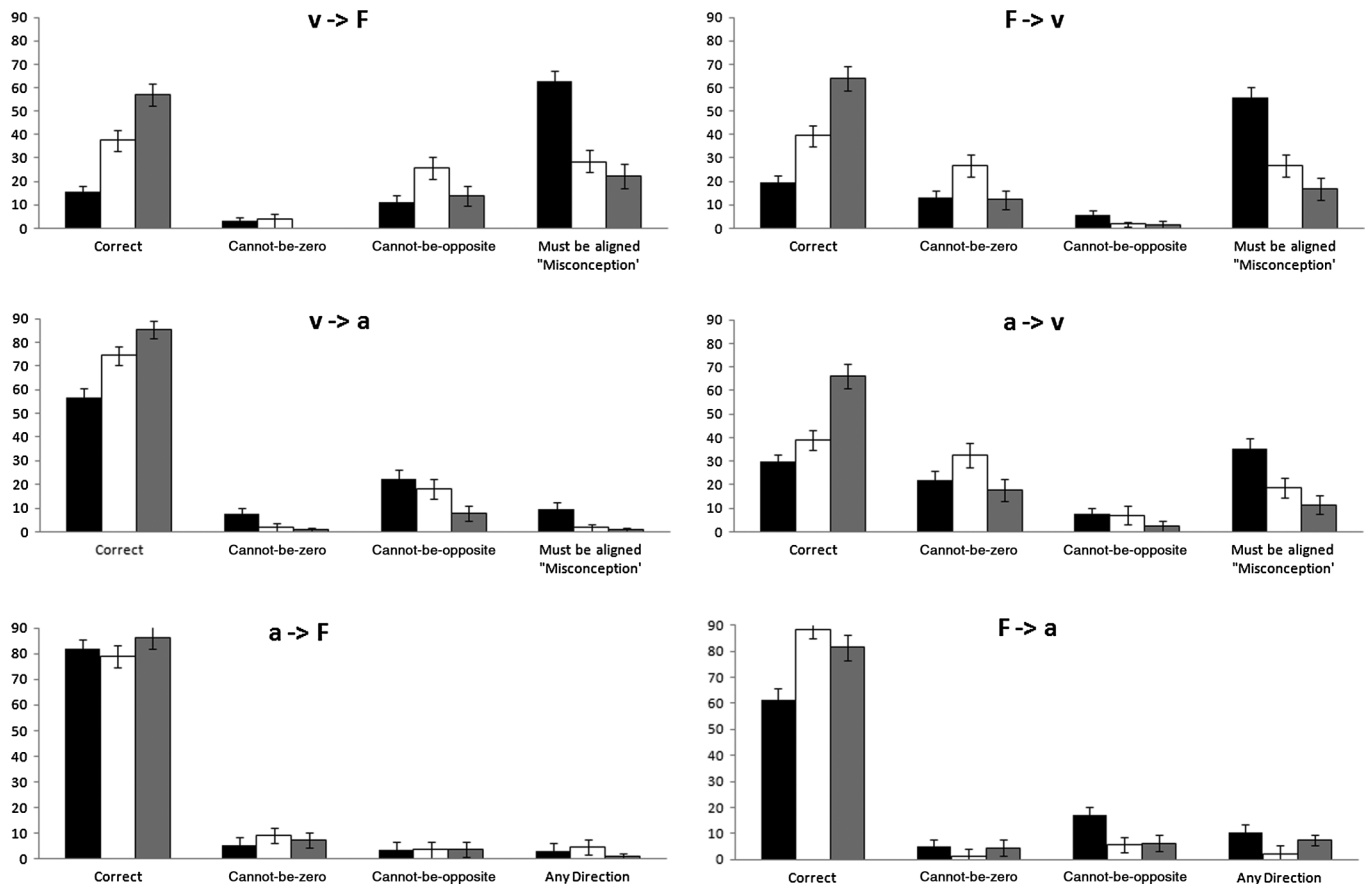


FIG. 3. Mean student response percentages for all six conditional relation question types and for three course levels. Black: standard calculus mechanics, $N = 228$; white: honors for engineers calculus mechanics, $N = 86$; gray: second-year physics majors, $N = 65$. Error bars are ± 1 standard error.

understanding of relations between the directions force, velocity, and acceleration as suggested by their choice of partially correct responses. By partially correct, we mean that the response included some physically valid possibilities not considered in the common misconception response. For example, for a $\vec{v} \rightarrow \vec{F}$ question, the common misconception response assumes that the inferred force must be nonzero and aligned with the velocity. In contrast, the somewhat common response choice that includes the possibilities that the net force is aligned *or* is zero (i.e., the “cannot-be-opposite” model) is more accurate than the common misconception response, and could be considered an intermediate, partially correct response. As seen in Fig. 3, intermediated levels of understanding occurred in all of the conditional relations between force, velocity, and acceleration.

Interviews with students further revealed that those choosing partially correct answers were often confident about their answers, for example, allowing for the possibility that a moving object can have a net force aligned with the motion or a zero net force, but certain that the net force cannot be opposite of the motion.

C. Asymmetry in response patterns between $\vec{x} \rightarrow \vec{y}$ and $\vec{y} \rightarrow \vec{x}$

Figure 3 also reveals two significant asymmetries in response patterns between a given conditional relation $\vec{x} \rightarrow \vec{y}$ and its converse $\vec{y} \rightarrow \vec{x}$. First, there were often asymmetries in scores, depending on the course level and the question types. For example, while there were only small differences between the $\vec{v} \rightarrow \vec{F}$ scores and $\vec{F} \rightarrow \vec{v}$ scores (effect sizes less than 0.12 standard deviations), there were significant differences between the $\vec{a} \rightarrow \vec{v}$ and $\vec{v} \rightarrow \vec{a}$ scores for the standard calculus-based physics course [16] [31% versus 57% correct, paired t test, $t(110) = 5.78$, $p < 0.001$, effect size $d = 0.54$] and for the honors physics majors course [39% versus 74% correct, paired t test, $t(85) = 7.49$, $p < 0.001$, effect size $d = 0.74$]. Clearly, most students correctly understand that a moving object can have an acceleration in any direction or zero acceleration, but many students also believe incorrectly that an accelerating object must be moving in the direction of its acceleration.

Another perhaps more surprising asymmetry in scores occurs for the $\vec{F} \rightarrow \vec{a}$ versus $\vec{a} \rightarrow \vec{F}$ questions. While there were no significant differences in responses for the first- and second-year physics majors courses (perhaps because they were answering at ceiling), there was a difference in responding to these two questions categories for the students in the standard calculus-based course. Specifically, the average score for the $\vec{F} \rightarrow \vec{a}$ question was 21% lower compared to the 82% score for the $\vec{a} \rightarrow \vec{F}$ question, which is a significant difference [paired t test, $t(227) = 5.50$, $p < 0.001$, effect size $d = 0.36$]. Interestingly, a similar asymmetry in scores occurs in pretest results for the first-year

physics majors course [58% correct for $\vec{F} \rightarrow \vec{a}$ and 84% correct for $\vec{a} \rightarrow \vec{F}$, paired t test, $t(227) = 6.94$, $p < 0.001$, effect size $d = 0.46$], but not for the post-test. This asymmetry in responding might be considered somewhat surprising since the relation $\vec{F} = m\vec{a}$ is a central relation in these physics courses (and readily recited by all students), but the results of the FVA test demonstrate that students in lower level courses often did not consider the conditional relationships between (net) force and acceleration to be symmetric.

A second significant kind of response asymmetry occurred in the *kinds* of intermediate, partially correct responses chosen by students. For example, for $\vec{v} \rightarrow \vec{F}$, the partially correct response chosen tended to be cannot-be-opposite, while for $\vec{F} \rightarrow \vec{v}$ the partially correct response chosen tended to be cannot-be-zero. Therefore, it appears as though if it is given that an object is moving, students more readily accepted that it may have a zero net force acting on it rather than accepting that it could have an opposing net force acting on it. On the other hand, if it is given that an object has a net force acting on it, students more readily accepted that it can move opposite the net force, rather than accepting that it is not moving at all. For $\vec{v} \rightarrow \vec{a}$ versus $\vec{a} \rightarrow \vec{v}$ questions, there are significant differences in all response choices, with the $\vec{v} \rightarrow \vec{a}$ questions tending to be answered correctly significantly more often. Similar to the $\vec{v} \rightarrow \vec{F}$ versus $\vec{F} \rightarrow \vec{v}$ questions, students tended to choose the cannot-be-opposite partially correct response for $\vec{v} \rightarrow \vec{a}$ and cannot-be-zero for $\vec{a} \rightarrow \vec{v}$ questions.

D. Other differences in scores between question types

In addition to differences in scores between a given conditional relation and its converse, there were also significant differences between other combinations of relations. For the standard calculus-based physics course the question types can be ranked as $\vec{v} \rightarrow \vec{F}$, $\vec{F} \rightarrow \vec{v}$, $\vec{a} \rightarrow \vec{v}$, $\vec{v} \rightarrow \vec{a}$, $\vec{F} \rightarrow \vec{a}$, $\vec{a} \rightarrow \vec{F}$, in order of increasing average score. The scores varied the most for the standard calculus-based physics course, but there were similar but reduced differences for the higher level courses. We will investigate the possible hierarchy of understanding of these relations in more detail in Sec. IV.

E. Difference in course levels: Evidence of evolution of understanding

While there were qualitative similarities between the patterns of the different course levels for each question type, there appears to be an “evolution” of the patterns from lower to higher course level. However, while the percentage of correct responses increased as the class level increased, the change in the misconception score between two classes was not always equal in magnitude to the change in the correct score. This appears to have been caused by a significant fraction of students choosing the

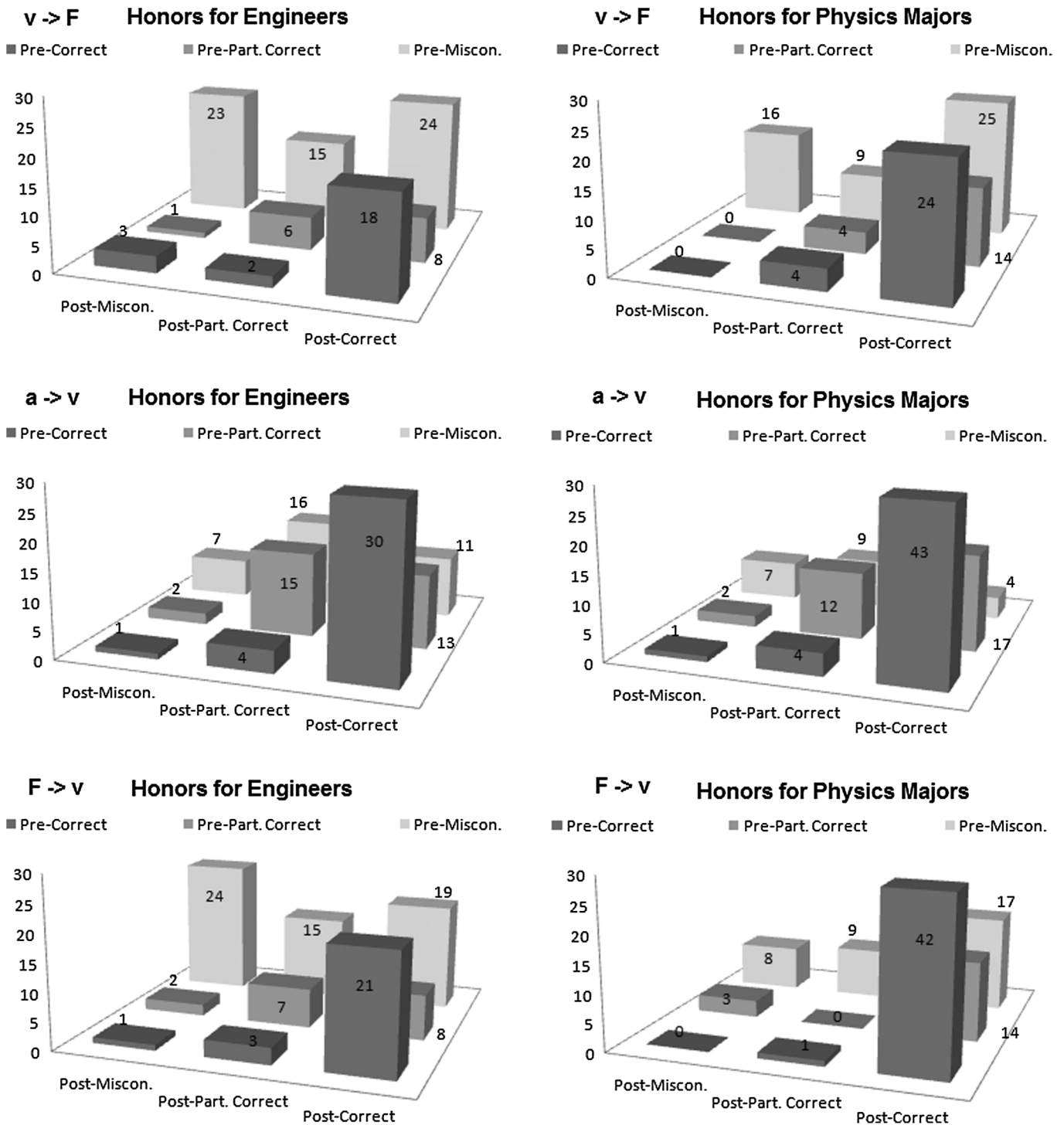


FIG. 4. Within-student, pretest versus post-test response choice percentages for the three lowest scoring question types. Responses were categorized as either Correct, Part. Correct (partially correct response), or Miscon. (misconceptionlike response). The data presented are from the honors Calculus Mechanics course for engineers, $N = 230$, and the first-year honors physics majors Calculus Mechanics course, $N = 49$. Note that roughly half of the students did not change their answer from pre- to post-test—these students are represented in the three diagonal columns from back left to front right of each graph. Also, note that a little less than half of students improved by moving into or out of a partially correct response, or directly from the misconception response to the fully correct response—these students are represented in the three columns behind and to the right of the diagonal columns—and roughly 10% of students answered less correctly from pre- to post-test—represented in the three columns to the front and left of the diagonal. (A few examples from the honors for engineers $\vec{v} \rightarrow F$ plot: 23% of students responded with a misconception on the pretest and a misconception on the post-test. 15% of students responded with a misconception on the pretest and a Part. Correct, partially correct response, on the post-test. 8% of students responded with a Part. Correct, partially correct response, on the pretest and a correct response on the post-test.)

partially correct cannot-be-zero and cannot-be-opposite responses, depending on the course level. For example, comparing the standard mechanics course to the honors for engineers mechanics course, the decrease in misconception responses was greater than the increase in correct responses, and the difference was comprised of students choosing one of the partially correct responses. Furthermore, when the difference between the honors and second-year course is considered, it is apparent that the increase in correct responses is greater than the decrease in misconception responses, and the balance is comprised of a decrease in the partially correct responses. These differences in response patterns between course levels suggest that a significant number of students evolved from an initial high level of misconceptions to the correct answer by passing through a partially correct response “state,” which indicates more knowledge than the common misconception but lacks the completeness of the correct response.

However, there is a danger in interpreting these data as evidence of evolution of understanding since it is cross sectional rather than longitudinal, and sometimes represents different kinds of students (e.g., physics majors versus engineering majors). Nonetheless, these data are consistent with the interesting possibility that students evolve through a partially correct state on the path to fully understanding the relations between the directions of force, velocity, and acceleration. We will investigate longitudinal data in the next section. (Much of the data presented here in Secs. III A–III E was shown and discussed in greater detail in Ref. [17].)

F. Pre- and post-FVA responses: Evidence of progression through intermediate levels

Pre- and post-FVA test data (i.e. longitudinal data) were analyzed in order to more closely investigate the evolution of student understanding of the relations between force, velocity, and acceleration. We were especially interested in determining whether the progression of student understanding involved passing through an intermediate, partially correct level of understanding. We administered the FVA test both pre- and post-instruction to 230 students in an honors calculus mechanics class for engineers and 49 students in a first-year honors calculus mechanics for physics majors class. As mentioned in Sec. II B and presented in Fig. 1, there were significant gains in the average scores for both classes. Perhaps more interesting, we examined within-student pre- versus post-test shifts in response choices for each item in the FVA test. Figure 4 presents a cross-tabulation of within-student pre- and post-test responses on a select set of FVA test items for the two courses. There are two important observations about the data represented in these figures.

First, it is helpful to describe the general patterns of shifts in student answering. Considering the honors

mechanics for engineers course for the $\vec{v} \rightarrow \vec{F}$, $\vec{F} \rightarrow \vec{v}$, and $\vec{a} \rightarrow \vec{v}$ questions, on average 51% of students did not change their answers, 43% answered “more correctly” on the post-test versus the pretest by either changing from the misconception response to either a partially correct response or to the correct response or changing from a partially correct response to the correct response, and conversely 6% answered less correctly. The results for the first-year physics majors course were somewhat similar, where 52% of students did not change their answers, 38% answered more correctly, and conversely 5% answered less correctly from pre- to post-test.

Second and more importantly, on average approximately 15% of students moved from the misconception response to a partially correct response and approximately 10% of students moved from a partially correct response to the correct response. These averages are representative of $\vec{v} \rightarrow \vec{F}$, $\vec{F} \rightarrow \vec{v}$, and $\vec{a} \rightarrow \vec{v}$ questions for both courses. This is to be compared with approximately 20% of students who moved directly from the misconception response to the correct response.

These results provide strong evidence that for many students the progression of student understanding involves passing through an intermediate, partially correct, level of understanding. Specifically, over half of the students who changed their answer changed either to or from an intermediate, partially correct, response.

IV. INVESTIGATING POSSIBLE HIERARCHIES IN STUDENT RESPONSES

In this section we are interested in investigating the question, “Does correctly answering a given conditional relation necessarily imply that another specific conditional relation was also answered correctly?” For example, does correctly answering $\vec{v} \rightarrow \vec{F}$ questions necessarily imply that $\vec{v} \rightarrow \vec{a}$ questions were also answered correctly? Note that, while one might make reasonable physical arguments to answer this question from an expert point of view, we are first interested in this as strictly an *empirical* question.

If such patterns in answering do exist, then one can proceed to make inferences as to the causes of these patterns. There are some standard analysis practices, such as Guttman scaling or scalogram methods, for determining the hierarchical-like structures of items for a unidimensional instrument. In fact, a full item response theory analysis of the FVA test can be used to find such hierarchies. However, here we are interested in the hierarchical relation between a number of different dimensions, such as understanding $\vec{v} \rightarrow \vec{F}$ or $\vec{v} \rightarrow \vec{a}$ probed by different items within the FVA test. Therefore, we will examine cross-tabular results between pairs of question types within the FVA test. A full Guttman scaling and/or item response theory analysis could also be informative from a more global perspective and is worth further study, but here we

TABLE VIII. A simple method for ruling out or finding supporting evidence for possible hierarchical structure in answering. Consider the generic contingency table on the top. Here, for example, a is the number of students answering both x and y incorrectly. If $c = 0$, $d \gg 1$, and $a \gg 1$, then this is consistent with the statement “answering x correctly implies answering y correctly” and the logical equivalent, “answering y incorrectly implies answering x incorrectly.” One could reasonably also use the conditions $d \gg c$ and $a \gg c$, since c will not be zero in practice due to random guessing, unusual students, etc. Furthermore, if these conditions on c are violated, then these statements are *disproved*, since c represents the number of counterexamples to these statements (and b represents the number of counterexamples to the converse of these statements). The table on the bottom presents a hypothetical example for $\vec{v} \rightarrow \vec{F}$ versus $\vec{v} \rightarrow \vec{a}$ questions. In this case one can claim this hypothetical data is consistent with the statement “correctly answering $\vec{v} \rightarrow \vec{F}$ questions implies correctly answering $\vec{v} \rightarrow \vec{a}$ questions” and “incorrectly answering $\vec{v} \rightarrow \vec{a}$ questions implies incorrectly answering $\vec{v} \rightarrow \vec{F}$ questions.” On the other hand, the relatively high counts in the “b” cell *disproves* the converse statement “correctly answering $\vec{v} \rightarrow \vec{a}$ questions implies correctly answering $\vec{v} \rightarrow \vec{F}$ questions.”

Generic contingency table			
		Question y	
		Incorrect	Correct
Question x	Incorrect	a	b
	Correct	c	d
Hypothetical example data			
		$\vec{v} \rightarrow \vec{a}$	
		Incorrect	Correct
$\vec{v} \rightarrow \vec{F}$	Incorrect	50	25
	Correct	0 ^a	25

^aThere are no counterexamples to the statement, “correctly answering $\vec{v} \rightarrow \vec{F}$ questions implies correctly answering $\vec{v} \rightarrow \vec{a}$ questions.”

will focus on hierarchies within the six conditional relations of interest.

Table VIII provides an example of a simple method to rule out or provide supporting evidence for the existence of hierarchies in response patterns for pairs of question types. In this hypothetical example, *all* of the students that answered $\vec{v} \rightarrow \vec{F}$ questions correctly also answered $\vec{v} \rightarrow \vec{a}$ correctly, but only *half* (25 out of 50) of the students that answered the $\vec{v} \rightarrow \vec{a}$ questions correctly answered the $\vec{v} \rightarrow \vec{F}$ correctly. These hypothetical data are consistent with the statement “correctly answering $\vec{v} \rightarrow \vec{F}$ questions necessarily implies correctly answering $\vec{v} \rightarrow \vec{a}$ questions” (the data are also consistent with the logically equivalent statement “incorrectly answering $\vec{v} \rightarrow \vec{a}$ questions implies incorrectly answering $\vec{v} \rightarrow \vec{F}$ questions”). Furthermore, these hypothetical data provide evidence to *disprove* the converse statement “correctly answering $\vec{v} \rightarrow \vec{a}$ questions implies correctly answering $\vec{v} \rightarrow \vec{F}$ questions,” since 25 out of 50 students are counterexamples to this statement.

TABLE IX. Within-student cross tabulations of scores between various question types. Cells represent numbers of students. Data reported are from the standard calculus mechanics course. A cell count is represented in bold face for tables which roughly satisfy conditions (discussed in Table VIII) which are consistent with significant hierarchies between the indicated question types. Some question types had two questions posed; in this case the label “Correct” indicates that at least one question of that type was answered correctly, and “Incorrect” indicates that zero questions of that type were answered correctly. Note that ϕ is the mean squared contingency coefficient between the question types, equivalent to the correlation coefficient for a 2×2 table.

	$\phi = 0.324$	$\vec{F} \rightarrow \vec{v}$	
	$\vec{v} \rightarrow \vec{F}$	Incorrect	Correct
A	Incorrect	58	17
	Correct	13	18
	$\phi = -0.026$	$\vec{a} \rightarrow \vec{v}$	
	$\vec{v} \rightarrow \vec{F}$	Incorrect	Correct
B	Incorrect	39	38
	Correct	17	13
	$\phi = 0.140$	$\vec{v} \rightarrow \vec{a}$	
	$\vec{v} \rightarrow \vec{F}$	Incorrect	Correct
C	Incorrect	25	50
	Correct	6	25
	$\phi = 0.138$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{v} \rightarrow \vec{F}$	Incorrect	Correct
D	Incorrect	16	59
	Correct	3	28
	$\phi = 0.131$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{v} \rightarrow \vec{F}$	Incorrect	Correct
E	Incorrect	27	48
	Correct	7	24
	$\phi = 0.451$	$\vec{a} \rightarrow \vec{v}$	
	$\vec{F} \rightarrow \vec{v}$	Incorrect	Correct
F	Incorrect	49	23
	Correct	7	28
	$\phi = 0.314$	$\vec{v} \rightarrow \vec{a}$	
	$\vec{F} \rightarrow \vec{v}$	Incorrect	Correct
G	Incorrect	28	44
	Correct	3	32
	$\phi = 0.130$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{F} \rightarrow \vec{v}$	Incorrect	Correct
H	Incorrect	16	57
	Correct	4	31
	$\phi = -0.066$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{F} \rightarrow \vec{v}$	Incorrect	Correct
I	Incorrect	22	50
	Correct	13	22
	$\phi = 0.238$	$\vec{v} \rightarrow \vec{a}$	
	$\vec{a} \rightarrow \vec{v}$	Incorrect	Correct
J	Incorrect	22	34
	Correct	9	42

TABLE IX. (Continued)

	$\phi = -0.022$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{a} \rightarrow \vec{v}$	Incorrect	Correct
K	Incorrect	10	46
	Correct	10	41
	$\phi = -0.013$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{a} \rightarrow \vec{v}$	Incorrect	Correct
L	Incorrect	18	38
	Correct	5	34
	$\phi = 0.064$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{v} \rightarrow \vec{a}$	Incorrect	Correct
M	Incorrect	7	24
	Correct	13	63
	$\phi = 0.038$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{v} \rightarrow \vec{a}$	Incorrect	Correct
N	Incorrect	11	20
	Correct	24	52
	$\phi = 0.126$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{F} \rightarrow \vec{a}$	Incorrect	Correct
O	Incorrect	9	27
	Correct	11	61

Therefore, one can analyze pairs of question types in this manner to either provide evidence *disproving* or supporting (but not *proving*) the existence of a particular hierarchy in answering, namely, that answering question type x correctly requires answering question type y correctly (but not the converse).

We analyzed within-student response patterns for all 15 possible pairs of question types on the FVA test (see Table IX) using the simple method shown in Table VIII, and found that there were no cases in which there were *zero* counterexamples for the statement “answering relation x correctly requires answering relation y correctly.” However, there were a number of cases in which there were a relatively small number of counterexamples.

These few counterexamples may be due to uninteresting causes such as random guessing. Thus, when there are only a relatively small number of counterexamples (rather than zero), this can still suggest the existence of a hierarchy in the answering pattern.

Using the constraint of a small number of counterexamples rather than zero counterexamples to indicate a hierarchy in answering, inspection of Table IX reveals a trend: for most pairs of relations, if the average score of a question type x was significantly less than the score for question type y , then it is also the case that correctly answering question type x implied correctly answering question type y , but *not* the converse. More specifically, if a student correctly answered the question types with the lowest average scores (the “difficult” question types), namely, $\vec{v} \rightarrow \vec{F}$ or $\vec{F} \rightarrow \vec{v}$, then most of the time this student also correctly answered question types with high

TABLE X. A summary of hierarchy trends suggested by Table IX.

Correct responses on this question type	Implies	Correct responses on this question type
$\vec{v} \rightarrow \vec{F}$		$\vec{v} \rightarrow \vec{a}$
$\vec{v} \rightarrow \vec{F}$		$\vec{a} \rightarrow \vec{F}$
$\vec{v} \rightarrow \vec{F}$		$\vec{F} \rightarrow \vec{a}$
$\vec{F} \rightarrow \vec{v}$		$\vec{a} \rightarrow \vec{v}$
$\vec{F} \rightarrow \vec{v}$		$\vec{v} \rightarrow \vec{a}$
$\vec{F} \rightarrow \vec{v}$		$\vec{a} \rightarrow \vec{F}$
$\vec{a} \rightarrow \vec{v}$		$\vec{v} \rightarrow \vec{a}$
$\vec{a} \rightarrow \vec{v}$		$\vec{F} \rightarrow \vec{a}$

average scores (i.e., “easy” question type), namely, $\vec{v} \rightarrow \vec{a}$, $\vec{F} \rightarrow \vec{a}$, or $\vec{a} \rightarrow \vec{F}$.

For example, when comparing student responses to both $\vec{F} \rightarrow \vec{v}$ and $\vec{v} \rightarrow \vec{a}$ questions for the standard calculus-based physics course, we found that the scores are 20% and 56% for $\vec{F} \rightarrow \vec{v}$ and $\vec{v} \rightarrow \vec{a}$ questions, respectively. As shown in Table IX(G), when comparing within-student responses, over 90% of students answering $\vec{F} \rightarrow \vec{v}$ questions correctly answered $\vec{v} \rightarrow \vec{a}$ questions correctly, $\frac{32}{32+3} \approx 91\%$, and over 90% of students answering $\vec{v} \rightarrow \vec{a}$ questions incorrectly answered $\vec{F} \rightarrow \vec{v}$ questions incorrectly, $\frac{28}{28+3} \approx 90\%$. For example, only about 40% of students answering $\vec{v} \rightarrow \vec{a}$ questions correctly answered $\vec{F} \rightarrow \vec{v}$ questions correctly, $\frac{32}{32+44} \approx 42\%$. These results are consistent with (but do not prove) the statement “correctly answering $\vec{F} \rightarrow \vec{v}$ questions necessarily implies correctly answering $\vec{v} \rightarrow \vec{a}$ questions, and the logical equivalent “incorrectly answering $\vec{v} \rightarrow \vec{a}$ questions implies incorrectly answering $\vec{F} \rightarrow \vec{v}$ questions.” Furthermore, this contingency table disproves the converse statement: “correctly answering $\vec{v} \rightarrow \vec{a}$ questions necessarily implies correctly answering $\vec{F} \rightarrow \vec{v}$ questions.”

Each table also includes the ϕ coefficient, which is a measure of correlation between the scores of each question type in the table. For example, in Table IX(G) discussed above, $\phi = 0.314$, denoting a medium-level correlation between the scores on $\vec{F} \rightarrow \vec{v}$ and $\vec{v} \rightarrow \vec{a}$ questions.

Finally, we use the method described in Table VIII on the data patterns in Table IX(A)–(O) to present a summary of potential hierarchies in Table X. Note that these hierarchies are only suggested by trends in the data tables. Nonetheless, these tables do provide strong evidence that statements converse to those in Table X are not true. For example, there is strong evidence (via a significant number of counterexamples) that the statement “correctly answering $\vec{v} \rightarrow \vec{a}$ questions implies correctly answering $\vec{v} \rightarrow \vec{F}$ ” is not true.

A. Comments on hierarchies in responses

There are three points we would like to address concerning the determination of hierarchies of understanding.

First, from the perspective of traditional design constraints on item statistics of a valid and reliable instrument (i.e., items have high internal consistency), it is not altogether unexpected that an answering hierarchy is aligned with increasing relative item score. Specifically, the constraint of choosing only items with a relatively high discrimination index implies that, for a given student, if items with low averages are answered correctly then items with high averages also are answered correctly. Nonetheless, this does not diminish the significance of the finding that answering patterns of some pairs of question types, such as $\vec{F} \rightarrow \vec{v}$ versus $\vec{v} \rightarrow \vec{a}$, are *not independent* and have a hierarchy (i.e., x implies y , but y does not imply x).

Second, it is worth pointing out that this analysis of hierarchies of question types can be viewed from the perspective of diagnostic assessment. Specifically, if students answer the difficult $\vec{v} \rightarrow \vec{F}$ and $\vec{F} \rightarrow \vec{v}$ questions correctly, then they are very likely to answer all other questions on the FVA test correctly. Therefore, to the extent that the FVA test measures understanding of the relations between force, velocity, and acceleration, one could view the $\vec{v} \rightarrow \vec{F}$ and $\vec{F} \rightarrow \vec{v}$ questions as the most *diagnostic* for determining understanding, at least for the level of students tested in this study.

Finally, while we have found evidence of hierarchies in student responses, our claims about hierarchies of student *understanding* of force, velocity, and acceleration are more qualified. It is important to keep in mind that “evidence of understanding” and any inferences of hierarchies that follow from such evidence depend on a careful characterization of how “understanding” is operationally defined. For example, when judging whether a student adequately understands the relation between force and velocity, it could be considered reasonable to require that a student correctly and explicitly distinguishes the differences between velocity and acceleration *as part of* their explanation of (or answers to question about) the relation between force and velocity. However, one might not expect the converse; namely, one might not require a student to explicitly distinguish the relation between force and velocity in order to demonstrate an understanding of the relation between velocity and acceleration. In this case, the evidence for a hierarchy of understanding $\vec{v} \rightarrow \vec{a}$ before $\vec{F} \rightarrow \vec{v}$ is strongly determined by the nature of the definition of understanding.

In this paper we have instead asked questions about a specific relation, say, $\vec{v} \rightarrow \vec{F}$, *without* any explicit reference to other relations, such as $\vec{v} \rightarrow \vec{a}$. To the extent that specific items in the FVA test measure understanding of each relation individually, without reference to other relations, observed hierarchical answering patterns suggest hierarchies in student understanding of relations between force, velocity, and acceleration. This finding is not simply an inevitable result of the operational definition of understanding of the relations, but appears to suggest that students at least implicitly connect different relations.

B. Hierarchies and evolution of responses

If there are hierarchies in responses to questions about the relations between force, velocity, and acceleration, then, in the course of learning, the evolution of responses should follow paths consistent with these hierarchies. Generally speaking, if correctly answering x necessarily implies correctly answering y , then gains in scores on y should precede gains in scores on x . For example, in the previous section we provided evidence that correctly answering $\vec{F} \rightarrow \vec{v}$ questions implied answering $\vec{v} \rightarrow \vec{a}$ questions, but not the converse. Therefore, for students initially performing poorly on both $\vec{v} \rightarrow \vec{a}$ and $\vec{F} \rightarrow \vec{v}$ questions, we would expect that within-student gains in answering $\vec{F} \rightarrow \vec{v}$ questions correctly would not occur without gains in answering $\vec{v} \rightarrow \vec{a}$ (though one might expect to see the converse).

The pre- and post-FVA test data described in Sec. III F allow for such a comparison of within-student gains in correct answering of the various question types. We analyzed all 15 possible pairs of question types (Table XI) using the simple hierarchy method shown in Table VIII, and found that there were no cases in which there were *zero* counterexamples for the statement “a gain in score for relations x necessarily implies a gain scores for relation y .” However, for a few pairs of relations we did find cases in which there were relatively few counterexamples to such a statement, and these cases were exactly the ones that one would expect from the evidence of hierarchies of understanding described in the previous section.

Specifically, as shown in Table XI(G), over 84% of students who improved their score on $\vec{F} \rightarrow \vec{v}$ questions also improved their score on $\vec{v} \rightarrow \vec{a}$ questions, $\frac{38}{38+7} \approx 84\%$, and over 80% of students who *did not* improve their score on $\vec{v} \rightarrow \vec{a}$ questions also did not improve their score on $\vec{F} \rightarrow \vec{v}$ questions, $\frac{29}{29+7} \approx 80\%$. Note that these patterns are not found for the converse. For example, less than 45% of students who improved their score on $\vec{v} \rightarrow \vec{a}$ questions also improved their score on $\vec{F} \rightarrow \vec{v}$ questions, $\frac{38}{38+47} \approx 45\%$. As mentioned, with a relatively small number of counterexamples notwithstanding, these findings are consistent with the evidence of hierarchy of understanding the relations $\vec{F} \rightarrow \vec{v}$ and $\vec{v} \rightarrow \vec{a}$.

In Table XII, we compiled the data even further to demonstrate the general pattern that a gain in either a $\vec{v} \rightarrow \vec{F}$ or a $\vec{F} \rightarrow \vec{v}$ question necessarily implies a gain in either a $\vec{v} \rightarrow \vec{a}$ or a $\vec{a} \rightarrow \vec{v}$ question. That is, over 92% of students who improved their score on either $\vec{F} \rightarrow \vec{v}$ or $\vec{v} \rightarrow \vec{F}$ questions also improved their score on either $\vec{v} \rightarrow \vec{a}$ or $\vec{a} \rightarrow \vec{v}$ questions, and over 77% of students who *did not* improve their score on $\vec{v} \rightarrow \vec{a}$ or $\vec{a} \rightarrow \vec{v}$ questions also did not improve their score on $\vec{F} \rightarrow \vec{v}$ or $\vec{v} \rightarrow \vec{F}$ questions. This would support the finding discussed earlier that correctly answering questions about the relations between acceleration and velocity tends to be required in order to correctly answer questions about the relations between

TABLE XI. Within-student cross tabulations of gains in scores between various question types. Cells represent numbers of students. Data reported are from the honors for engineers course. A cell count is represented in bold face for tables which roughly satisfy conditions (discussed in Table VIII) which are consistent with significant hierarchies between the indicated question types. Cases in which the score was 2 out of 2 both on pre- and post-tests on a specific question type were removed. This helps to remove the less interesting “ceiling cases” that would register as “no gain.” Some question types had two questions posed; in this case the label “Gain” indicates an increase of at least one correct response for that question type, and “No gain” indicates either no increase in correct responses or a loss in correct responses for that question type. Note that ϕ is the mean squared contingency coefficient between the question types, equivalent to the correlation coefficient for a 2×2 table.

A	$\phi = 0.352$	$\vec{F} \rightarrow \vec{v}$	
	$\vec{v} \rightarrow \vec{F}$	No gain	Gain
	No gain	66	20
	Gain	43	67
B	$\phi = 0.228$	$\vec{a} \rightarrow \vec{v}$	
	$\vec{v} \rightarrow \vec{F}$	No gain	Gain
	No gain	54	27
	Gain	44	58
C	$\phi = 0.205$	$\vec{v} \rightarrow \vec{a}$	
	$\vec{v} \rightarrow \vec{F}$	No gain	Gain
	No Gain	25	40
	Gain	11	46
D	$\phi = 0.341$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{v} \rightarrow \vec{F}$	No gain	Gain
	No gain	28	21
	Gain	14	45
E	$\phi = 0.117$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{v} \rightarrow \vec{F}$	No gain	Gain
	No gain	10	11
	Gain	12	16
F	$\phi = 0.389$	$\vec{a} \rightarrow \vec{v}$	
	$\vec{F} \rightarrow \vec{v}$	No gain	Gain
	No Gain	76	32
	Gain	20	53
G	$\phi = 0.232$	$\vec{v} \rightarrow \vec{a}$	
	$\vec{F} \rightarrow \vec{v}$	No gain	Gain
	No gain	29	47
	Gain	7	38
H	$\phi = 0.386$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{F} \rightarrow \vec{v}$	No gain	Gain
	No Gain	33	26
	Gain	9	41
I	$\phi = 0.133$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{F} \rightarrow \vec{v}$	No gain	Gain
	No gain	13	17
	Gain	10	12

TABLE XI. (Continued)

J	$\phi = 0.091$	$\vec{v} \rightarrow \vec{a}$	
	$\vec{a} \rightarrow \vec{v}$	No gain	Gain
	No gain	22	46
	Gain	12	38
K	$\phi = 0.221$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{a} \rightarrow \vec{v}$	No gain	Gain
	No gain	26	24
	Gain	15	32
L	$\phi = 0.108$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{a} \rightarrow \vec{v}$	No gain	Gain
	No gain	13	12
	Gain	8	11
M	$\phi = 0.239$	$\vec{F} \rightarrow \vec{a}$	
	$\vec{v} \rightarrow \vec{a}$	No gain	Gain
	No gain	14	11
	Gain	20	28
N	$\phi = 0.252$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{v} \rightarrow \vec{a}$	No Gain	Gain
	No gain	4	4
	Gain	11	15
O	$\phi = 0.148$	$\vec{a} \rightarrow \vec{F}$	
	$\vec{F} \rightarrow \vec{a}$	No gain	Gain
	No gain	7	8
	Gain	7	12

force and velocity. We did not find significant hierarchies for gains involving $\vec{F} \rightarrow \vec{a}$ or $\vec{a} \rightarrow \vec{F}$ questions, most likely because the score to these questions were already near ceiling, leaving little room for gain.

In summary, the contingency tables of the gains, pretest to post-test, in student scores are in agreement with the hierarchies deduced from the contingency tables of within-student answering at a single time. This provides yet more evidence that specific hierarchies exist in student understanding of the relations between force, velocity, and acceleration and these hierarchies affect the evolution of student understanding.

V. SUMMARY

We have developed a 17 item multiple choice test, the “FVA test,” designed to probe students’ understanding of the relationships between the directions of net force,

TABLE XII. Summary of hierarchy in gains for $v \rightarrow F$ and $F \rightarrow v$ with $v \rightarrow a$ and $a \rightarrow v$.

$\phi = 0.313$	$\vec{v} \rightarrow \vec{a}$ or $\vec{a} \rightarrow \vec{v}$	
$\vec{v} \rightarrow \vec{F}$ or $\vec{F} \rightarrow \vec{v}$	No gain	Gain
No gain	17	35
Gain	5	58

velocity, and acceleration of (or on) a particle in one dimension. While previous research has examined one or two of these relationships at a time, the goal here was to holistically examine answering patterns for all six possible pairs of conditional relationships in order to obtain a more coherent picture of student understanding of these relations among the concepts of force, velocity, and acceleration. The development of the instrument included multiple stages of revisions with feedback via interviews and testing with standard and honors calculus-based introductory university students as well as second-year physics majors. The test has been shown to have significant statistical reliability as well as validity for the population tested, as shown, for example, by significant correlations of FVA test score with course grade, level of the student, and the Force Concept Inventory score. The overall test scores indicate that traditional calculus-based physics students performed poorly on the test, with an average score of about 40%, and even second-year physics majors find these questions somewhat challenging, with an average score of 70%. Furthermore, detailed patterns in student responses to the FVA test were analyzed, and several interesting findings were reported, as summarized below.

First, we consistently found evidence of an intermediate, partially correct level of understanding of the relations between force, velocity, and acceleration held by up to 30% of the students pre- or post-instruction. This is in addition to finding that a significant number of students answer consistent with the well-known and common student misconception that the vector quantities should always point in the same direction. Specifically, we found two intermediate models. The first model is the belief that two vectors, such as force and velocity, need not be aligned, but they may also be pointed in opposite directions, but one cannot be zero (“cannot-be-zero” model). The second model is the belief that the two vectors need not be aligned, though one of them could be zero but not pointed in the opposite direction as the other (“cannot-be-opposite” model). Furthermore, we found that about half of the students who improved their understanding of the relations between the directions of force, velocity, and acceleration did so by evolving through these partially correct “states.” Roughly speaking, from pre- to post-test in the honors physics sections, we found that about half of the students did not change their responses, about 1/4 changed from the misconception answer to the correct answer, and about 1/4 either changed from the misconception answer to the partially correct answer or from the partially correct answer to the correct answer.

Second, we found an asymmetry in student responses to conditional relations. That is, students often treated questions that probe the concept *motion implies acceleration* differently than the concept *acceleration implies motion*. Likewise they often treated questions about *motion implies force* differently than *force implies motion*, and

perhaps surprisingly, they often treated questions about *force implies acceleration* differently than *acceleration implies force*. The differences are reflected in the response frequencies of each answer choice. For example, for $\vec{v} \rightarrow \vec{a}$ versus $\vec{a} \rightarrow \vec{v}$ questions, there were differences in the number of correct and misconception responses as well as in the kinds of partially correct responses (cannot-be-opposite versus cannot-be-zero). For the $\vec{v} \rightarrow \vec{F}$ versus $\vec{F} \rightarrow \vec{v}$ questions, there were no differences in the correct and misconception response frequencies, but there were differences in the partially correct cannot-be-opposite versus cannot-be-zero responses.

Third, we found evidence of specific hierarchies in correct responses to different question types. The evidence included both within-student scores at one point in time and within-student gains in scores from pre- to post-test. For example, we found evidence that if students correctly answered $\vec{F} \rightarrow \vec{v}$ questions, then they were very likely to also correctly answer $\vec{v} \rightarrow \vec{a}$ questions, but not the converse. Further, if $\vec{v} \rightarrow \vec{a}$ questions were answered incorrectly, then it was very likely that $\vec{F} \rightarrow \vec{v}$ questions were also answered incorrectly (but not the converse). Likewise, we found that for a given student, gains in $\vec{F} \rightarrow \vec{v}$ scores most likely occurred in the presence of gains in $\vec{v} \rightarrow \vec{a}$ scores, but not the converse. These findings are indeed interesting and suggest that it may be required to understand the relationship between the direction of velocity and acceleration in order to understand the relationship between the direction of force and velocity. However, to more firmly establish a possible causal link between understanding these relations, one must first be careful to explicitly define what is meant by *understand*, and, second, a controlled intervention (for example, manipulating the amount of velocity-acceleration instruction) is needed.

VI. COMMENT ON LEARNING PROGRESSIONS

As mentioned in Sec. I, Alonso and Steedle have hypothesized successive levels of understanding of force and motion [12]. There are several major differences between our study and theirs. For example, they studied middle-school students, they studied understanding of magnitude (including change in magnitude) of quantities of force and motion, and only to a lesser extent did they also study understanding of relative direction of force and motion. Furthermore, they did not systematically study student understanding of the concept of acceleration (including direction) and its relations to velocity. Instead they focus on “motion” and occasionally make explicit references to “acceleration.” Nonetheless, an examination of their hypothesized levels of understanding reveal that, in their model, students tend to understand issues concerning the relations between the direction of force and motion before they come to understand issues about the relations between the magnitude (and changes in magnitude) of force and

motion. This result is certainly worth confirming in further focused empirical studies.

In contrast, our study is focused solely on the understanding of the relations between the directions of force, velocity, and acceleration, and as such our results do not contradict or confirm their results. Instead, our results may add more detail and depth to a portion of a larger learning progression framework for force and motion that may also include Alonso and Steedle's work. While the term "learning progression" has not been uniquely defined (for example, see the discussion in Alonso and Steedle's work [12]), the general idea is one of successive stages of student understanding of a concept or topic, starting from incomplete or incorrect knowledge and ending with some defined level of mastery, usually described by a particular science education standard. We have not set out *a priori* to construct a learning progression, rather we found that a consistent progression emerged out of our longitudinal and cross-sectional data. This is somewhat in contrast to typical work on learning progressions (including that of Alonso and Steedle), which, rather than being primarily empirically driven, were typically constructed by an expert as some logical progression (from an expert point of view) toward mastery, with only some input on empirical data on how students are thinking or how they might progress toward mastery. As Alonso and Steedle's article states, "the learning progression represents a hypothesis about student thinking, rather than a description" [12].

Indeed our approach is more empirical. We carefully designed questions to probe student understanding of logically and scientifically relevant dimensions (from an expert's perspective), namely, the six conditional relations. Nonetheless, understanding these relations could also be seen as subgoals of understanding force and motion in general, and while an expert might logically order how these subgoals would best be learned, this does not preclude the order in which students actually learn them, which is an empirical question investigated here.

In summary, some of the results in this paper could be used to link with recent efforts to identify learning progressions of force and motion. Specifically, our results could be used to construct a more formalized, empirically based learning progression of student understanding of the directions of force, velocity, and acceleration, and this could be useful for instruction. Other implications for instruction are discussed in the next section.

VII. IMPLICATIONS FOR INSTRUCTION

We will focus on some of the most important implications of the three major findings summarized in a previous section. First, instruction may be more effective if it accounts for the existence of intermediate states of understanding, especially since these intermediate states vary, depending on the specific conditional relation. For example, if an instructor focuses on the point "an object

moving at constant velocity must have zero net force acting on it," this may help some students move from the misconception level into the somewhat common "cannot-be-opposite" intermediate state for that conditional relation, but it is unlikely to help the significant population of students who were already in that intermediate state to advance to fully correct understanding. Instead, instructors should be aware of the importance of focusing on the point that "an object which is moving may have a net force opposite to its direction of motion." Furthermore, if instructors are not careful in their assessment, they may incorrectly infer that students in the intermediate, partially correct, state have a complete understanding.

Another implication for instruction stems from asymmetry in responses to conditional relations between force, velocity, and acceleration. This implies that students may consider conditional examples differently during instruction. That is, a student who sees an example demonstrating that an object with a given instantaneous velocity can have any value of net force acting on it may perceive this differently than an example in which an object with a given net force can have any value of velocity. Furthermore, these two different examples may address different intermediate levels of understanding, as mentioned earlier. Therefore, attention must be given to both kinds of examples in order to fully address student difficulties with understanding these relations.

Finally, evidence for potential hierarchies in understanding the relations between the directions of force, velocity, and acceleration naturally has important implications for the order of instructional units and priorities for their mastery. For example, the results of this study imply that instructors must first ensure that students understand the relation between the direction of velocity and acceleration as well as force and acceleration in order to ensure that the students understand the relation between velocity and force, which is the source of common, compelling misconceptions. While from an expert point of view this order seems quite reasonable and perhaps expected, it is important to keep in mind that this study implies that teaching in the reverse order will not be as effective. Namely, teaching students *first* about the common misconceptions involving the relations between velocity and force may not be effective in preparing them to learn about the relations between velocity and acceleration or force and acceleration. While this and other implications of the order of instruction following from evidence of hierarchies of understanding is an interesting result, clearly more carefully controlled intervention studies are needed in order to better establish their validity.

In summary, we have found that the carefully designed FVA test has provided more comprehensive insight into student understanding of the relations between the directions of force, velocity, and acceleration in one dimension. Clearly the levels of understanding of these concepts has a

rich and interesting structure, and the results of this study can help to inform careful decisions about the order and priorities of instruction as well as the identification and use of critical types of example questions to improve student understanding of this fundamental topic.

ACKNOWLEDGMENTS

We would like to thank Eleanor Sayre, Thomas Scaife, and Jeffery Asch for their contributions on test development and data collection. This research was supported by a grant from the Institute of Educational Sciences of the U.S. Department of Education (R305H050125).

APPENDIX: FVA TEST

1. At exactly 2:31PM, a boat is moving to the north on a lake. Which statement best describes the forces on the boat at this time?

- (a) there may be several forces to the north and to the south acting on the boat, but the forces to the north are larger.
- (b) there may be several forces to the north and to the south acting on the boat, but the forces to the south are larger.
- (c) there may be several forces to the north and to the south acting on the boat, but the forces to the south are equal in magnitude to those to the north.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

2. A car is moving to the right and speeding up. Which statement best describes the acceleration of the car at this instant?

- (a) the car's acceleration is to the right.
- (b) the car's acceleration is to the left.
- (c) the car's acceleration is zero.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

3. A student and a dog are playing tug of war with a rubber toy. If at a particular time the student is pulling on the toy to the right and the dog is pulling to the left with an equal force, which statement best describes the motion of the toy at this time?

- (a) it is moving toward the dog.
- (b) it is moving toward the student.
- (c) it is not moving.
- (d) both a and b are possible.
- (e) a, b, and c are possible.

4. A car is on a hill and the direction of its acceleration is uphill. Which statement best describes the motion of the car at that time?

- (a) it is moving uphill.
- (b) it is to moving downhill.
- (c) it is not moving.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

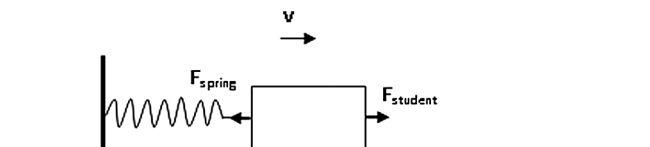
5. A group of workers is pushing on a car in a driveway. There may be several forces on the car but those toward the street are greater. Which statement best describes the acceleration of the car at this instant?

- (a) its acceleration is toward the street.
- (b) its acceleration is away from the street.
- (c) its acceleration is zero.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b and c are possible.

6. A wagon is rolling east along the sidewalk. What can you say about the acceleration of the wagon at this time?

- (a) it is accelerating east.
- (b) it is accelerating west.
- (c) it is not accelerating.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

7. At exactly $t_0 = 3.0$ sec, a student is pulling with a force F_{student} on a box which is connected to a spring, as in the diagram below. The spring is exerting a force F_{spring} on the box. At this exact time, the box is moving toward the right and slowing down. Assuming the friction is negligible, which statement best describes the forces at this time?



- (a) $F_{\text{student}} > F_{\text{spring}}$.
- (b) $F_{\text{student}} < F_{\text{spring}}$.
- (c) $F_{\text{student}} = F_{\text{spring}}$.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) both b and c are possible.
- (g) a, b and c are possible.

8. The direction of acceleration of an object is to the right. What is the most you can say about the motion of the object at this time?

- (a) it is moving to the right and its speed is increasing.
- (b) it is moving to the right and its speed is decreasing.
- (c) it is to moving to the left.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

9. As a wood block slides past a tick-mark on a smooth level surface, it has a velocity of 2 m/s to the right. There is a small amount of friction between the block and the surface, and eventually the block comes to rest. What are the forces acting on the block a few moments after it passes the mark (circle all that apply)?

- (a) weight (down).
- (b) normal force (up).
- (c) force of block (right).
- (d) friction (left).
- (e) there are no forces on the block.

10. A force sensor is attached inside a soccer ball that is used during a match. The force sensor measures the forces acting on the ball. At a randomly chosen instant during the game, the sensor detects that there is only one horizontal force on the ball, and that force is directed toward the home-team goal. Which statement best describes the motion of the ball at this instant?

- (a) the ball is moving toward the home-team goal.
- (b) the ball is moving away from the home-team goal.
- (c) the ball is not moving.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

11. At a particular instant of time during a kickball game, a ball on the playground is accelerating to the right. What can you say about the forces on the ball at this time?

- (a) there may be several forces to the right and to the left acting on the ball, but the forces to the right are larger.
- (b) there may be several forces to the right and to the left acting on the ball, but the forces to the left are larger.
- (c) there may be several forces to the right and to the left acting on the ball, but the forces to the right are equal in magnitude to those to the left.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

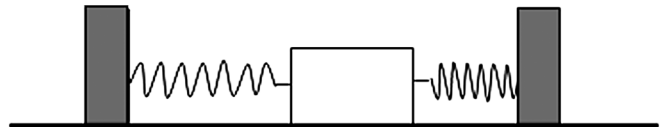
12. A boy rolls a ball toward the east on level ground into the wind. The ball rolls eastward against the wind and slows down, and after a short time the ball stops and rolls westward (with the wind) and starts to speed up. At the moment the ball turns around, the velocity is zero—which statement best describes the forces on the ball at this moment?

- (a) there may be several forces to the east and to the west acting on the ball, but the forces to the east are larger.
- (b) there may be several forces to the east and to the west acting on the ball, but the forces to the west are larger.

(c) there may be several forces to the east and to the west acting on the ball, but the forces to the west are equal in magnitude to those to the east.

- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) both b and c are possible.
- (g) a, b, and c are possible.

13. A block is attached between two springs as in the diagram below, and oscillates back and forth. At an instant of time depicted in the diagram, the acceleration of the block is to the left. Which statement best describes the motion of the block at this instant?



- (a) it is moving left.
- (b) it is moving right.
- (c) it is not moving.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) both b and c are possible.
- (g) a, b, and c are possible.

14. At a particular instant of time, there are several forces acting on an object in both the positive and negative direction, but the forces in the negative direction (to the left) are greater. Which statement best describes the motion of the object at this instant?

- (a) it is moving to the right.
- (b) it is moving to the left.
- (c) it is not moving.
- (d) both a and b are possible.
- (e) both b and c are possible.
- (f) a, b, and c are possible.

15. A child is playing with a toy car. At one instant, the acceleration of the toy car is to the north. Which statement best describes the toy car's motion?

- (a) its speed is increasing.
- (b) its speed is decreasing.
- (c) both a and b are possible.

16. At exactly 10:02 A.M., a man is pushing to the right on a box with a force, F . There is also a friction force f between the box and the floor. If at that exact moment, the box is moving to the right, which statement best describes the forces on the box at that time?

- (a) $F > f$.
- (b) $F < f$.
- (c) $F = f$.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b and c are possible.

17. A train going from Columbus to Cleveland passes a telephone pole. What can you say about the acceleration of the train when it passes the pole?

- (a) it is accelerating toward Cleveland.
- (b) it is accelerating to toward Columbus.

- (c) it is not accelerating.
- (d) both a and b are possible.
- (e) both a and c are possible.
- (f) a, b, and c are possible.

-
- [1] I. A. Halloun and D. Hestenes, The initial knowledge state of college physics students, *Am. J. Phys.* **53**, 1043 (1985).
 - [2] R. K. Thornton and D. R. Sokoloff, Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula, *Am. J. Phys.* **66**, 338 (1998).
 - [3] J. Clement, Students' preconceptions in introductory mechanics, *Am. J. Phys.* **50**, 66 (1982).
 - [4] L. Viennot, Spontaneous reasoning in elementary dynamics, *Eur. J. Sci. Educ.* **1**, 205 (1979).
 - [5] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, *Phys. Teach.* **30**, 141 (1992).
 - [6] D. E. Trowbridge and L. C. McDermott, Investigation of student understanding of the concept of acceleration in one dimension, *Am. J. Phys.* **49**, 242 (1981).
 - [7] F. Reif and S. Allen, Cognition for interpreting scientific concepts: A study of acceleration, *Cogn. Instr.* **9**, 1 (1992).
 - [8] L. G. Enderstein and P. E. Spargo, Beliefs regarding force and motion: A longitudinal and cross-cultural study of South African school pupils, *Int. J. Sci. Educ.* **18**, 479 (1996).
 - [9] D. Palmer, The effect of context on students' reasoning about forces, *Int. J. Sci. Educ.* **19**, 681 (1997).
 - [10] L. G. Enderstein and P. E. Spargo, The effect of context, culture and learning on the selection of alternative options in similar situations by South African pupils, *Int. J. Sci. Educ.* **20**, 711 (1998).
 - [11] A. Tversky and D. Kahneman, in *Progress in Social Psychology*, edited by M. Fishbein (Erlbaum, Hillsdale, NJ, 1980), pp. 49–72.
 - [12] A. C. Alonzo and J. T. Steedle, Developing and assessing a force and motion learning progression, *Sci. Educ.* **93**, 389 (2009).
 - [13] *Taking Science to School* (National Academies Press, Washington, DC, 2007).
 - [14] R. Rosenblatt, E. C. Sayre, and A. F. Heckler, Toward a comprehensive picture of student understanding of force, velocity, and acceleration, in *Proceedings of the 2008 PER Conference, Edmonton, Alberta, Canada, 2008* (AIP, Melville, NY, 2008), pp. 183–186.
 - [15] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.* **66**, 64 (1998).
 - [16] Winter2008 students received an older version of the FVA test which did not have $\vec{v} \rightarrow \vec{a}$ questions. Thus, they were excluded from this analysis.
 - [17] R. Rosenblatt, E. C. Sayre, and A. F. Heckler, Modeling students' conceptual understanding of force, velocity, and acceleration, in *Proceedings of the 2009 PER Conference, Ann Arbor, MI, 2009* (AIP, Melville, NY, 2009), pp. 245–248.