

Applying a resources framework to analysis of the Force and Motion Conceptual Evaluation

Trevor I. Smith and Michael C. Wittmann

Department of Physics and Astronomy, University of Maine, Orono, Maine 04469, USA

(Received 8 November 2007; published 10 September 2008)

We suggest one redefinition of common clusters of questions used to analyze student responses on the Force and Motion Conceptual Evaluation. Our goal is to propose a methodology that moves beyond an analysis of student learning defined by correct responses, either on the overall test or on clusters of questions defined solely by content. We use the resources framework theory of learning to define clusters within this experimental test that was designed without the resources framework in mind. We take special note of the contextual and representational dependence of questions with seemingly similar physics content. We analyze clusters in ways that allow the most common incorrect answers to give as much, or more, information as the correctness of responses in that cluster. We show that false positives can be found, especially on questions dealing with Newton's third law. We apply our clustering to a small set of data to illustrate the value of comparing students' incorrect responses which are otherwise identical on a correct or incorrect analysis. Our work provides a connection between theory and experiment in the area of survey design and the resources framework.

DOI: [10.1103/PhysRevSTPER.4.020101](https://doi.org/10.1103/PhysRevSTPER.4.020101)

PACS number(s): 01.40.Fk, 01.40.gf

I. INTRODUCTION

Students are not yet physicists. They have not had the extensive training that we, as physicists, rely on. As a result, it is often inappropriate to categorize and group student responses to physics questions solely on the basis of agreeing with correct Newtonian principles, as is commonly done with standardized tests such as the Force Concept Inventory (FCI) and the Force and Motion Conceptual Evaluation (FMCE).¹⁻³ Unfortunately student assessment using the FMCE, including work previously done by one of the authors (M.C.W.), regularly does just that.⁴ The test questions are grouped into clusters according to a physicist's view of equivalent content areas, and students' responses are evaluated based on their agreement with a physicist's viewpoint without regard for why students might choose incorrect answers. This may be a valid form of assessment for determining how well students think like physicists, but it is often an insufficient method for determining how students reason about scenarios in a physics context.

In order to effect greater conceptual development in our students, we must understand not only where we wish them to end up but also where they are beginning in terms of their understanding of the world around them. Only by having a clear picture of correct and incorrect responses in appropriately defined clusters, while accounting for context and representational dependence of questions, may we devise a manner by which to help our students truly gain a physicist's understanding of their surroundings. This paper describes more detailed analysis tools for instructors and researchers and is therefore primarily methodological. We illustrate the value of our results by comparing student data but do not highlight such an analysis in this paper.

We describe ways in which the FMCE can be organized and used to get more detailed information about students. Presently, many researchers and educators using the FMCE follow a particular procedure that includes three steps: (1) administering the FMCE pre-instruction and post-instruction, (2) dividing the questions into content-based clusters, and (3)

evaluating the correctness of each student's responses within each cluster as well as over the entire test. Several years ago a template was developed by one of the authors (M.C.W.) to analyze students' responses to the FMCE within five clusters [Velocity, Acceleration, Force (Newton's first and second laws), Newton's third law,⁵ and Energy].⁴ The template automatically scores each response as correct or incorrect, groups questions into the aforementioned clusters, and calculates a class's normalized gain for each cluster as well as over the entire test. This template has become widely used due to its availability and its succinct analysis of students' responses. Recent research using the FMCE and modeling using a resources framework⁶⁻⁸ has convinced us that analysis based on the use of this template lacks the depth and rigor we have come to expect from studies on students' understanding of Newtonian mechanics. We propose several modifications to the described analysis including a redefinition of clusters and a deeper analysis of students' incorrect responses.

The clusters mentioned previously divide the FMCE very nicely into groups of questions, each of which examines a different content aspect of physics. However many studies, including those by Beichner⁹ and Kohl and Finkelstein,^{10,11} have shown that the manner in which material is presented may significantly affect students' abilities to demonstrate their understanding. For example, Beichner⁹ showed that students' understanding (or lack of understanding) of graphs can have a profound impact on their responses to physics questions involving graphs.⁹ Furthermore, results by Dykstra¹² and others show that elements of the physical situation greatly affect reasoning. Dykstra¹² reported on students' troubles in reasoning about motion in which an object has a turning point; that is, when an object under the influence of a constant force moves in a particular direction while slowing down and then reverses direction and speeds up.¹²

Our goals in this paper are not new. Several researchers (notably Thornton^{13,14} and Dykstra¹²) have used results from the FMCE to give fine-grained analyses of students' responses to the FMCE. Unfortunately, their methods are not widely used among physics educators and education re-

TABLE I. Clusters of questions on the FMCE as previously defined by Wittmann.⁴

Cluster	Questions
Velocity	40–43
Acceleration	22–29
Force (Newton I and II)	1–4, 7–14, 16–21
Newton III	30–32, 34, 36, 38
Energy	44–47

searchers. Also, we wish to anchor our analysis in a resources framework, making explicit connections to the representational and contextual dependence of responses. In addition, Bao¹⁶ proposed new clusters of questions to investigate reasoning. We respond to Bao’s work¹⁶ in more detail in Sec. II.

In Sec. II we both respond to existing clustering methods (including our previous one) and propose a different set of clusters. In Sec. III we discuss various incorrect mental models that correspond with the content areas described by each of our clusters. In Sec. IV we examine how these mental models are aligned with particular responses to questions in the FMCE. Using a definition of clusters consistent with a resources framework allows us to go into greater detail about students’ incorrect mental models and identify responses that correspond to these models. In this section, we include a discussion of false positives, as measured by looking at responses on several questions within a representationally and contextually consistent set of questions. Finally, in Sec. V, we show two examples of how our analysis gives insight into student responses in a way that our original clustering was unable to show. We note again that our purpose in this paper is methodological, in that we use models of student reasoning (the resources framework) to revise our implementation of an experimental tool.

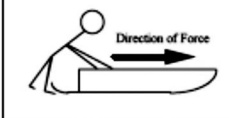
II. REVISED QUESTION CLUSTERS FOR THE FMCE

The five analysis clusters, shown in Table I, were chosen by Wittmann⁴ as a quick and dirty analysis of classroom performance on the FMCE. These clusters were defined based on the content of each question on the FMCE—the Velocity cluster asks students about the velocity of an object undergoing a series of described motions, the Force (Newton I and II) cluster asks students about the force(s) exerted on an object during a described motion, and so on. Several questions are not included in any of the clusters; Thornton and Sokoloff³ omitted these from regular analysis of the FMCE because they are intended for diagnostic purposes (such as reading ability) or do not give a definitive indication as to whether or not a student is properly using Newtonian reasoning. More details of these omissions can be found in Refs. 14,15, while issues of question weighting are discussed below.


There are obvious flaws to the old clustering of questions, the largest being that different representations and contexts are asked about in many clusters. If students are inconsistent

A sled on ice moves in the ways described in questions 1-7 below. Friction is so small that it can be ignored. A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the one force (A through G) which would keep the sled moving as described in each statement below.


You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice J.



A. The force is toward the right and is increasing in strength (magnitude).
 B. The force is toward the right and is of constant strength (magnitude).
 C. The force is toward the right and is decreasing in strength (magnitude).



D. No applied force is needed



E. The force is toward the left and is decreasing in strength (magnitude).
 F. The force is toward the left and is of constant strength (magnitude).
 G. The force is toward the left and is increasing in strength (magnitude).

- ___1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?
- ___2. Which force would keep the sled moving toward the right at a steady (constant) velocity?
- ___3. The sled is moving toward the right. Which force would slow it down at a steady rate (constant acceleration)?
- ___4. Which force would keep the sled moving toward the left and speeding up at a steady rate (constant acceleration)?
- ___5. The sled was started from rest and pushed until it reached a steady (constant) velocity toward the right. Which force would keep the sled moving at this velocity?
- ___6. The sled is slowing down at a steady rate and has an acceleration to the right. Which force would account for this motion?
- ___7. The sled is moving toward the left. Which force would slow it down at a steady rate (constant acceleration)?

FIG. 1. Questions 1–7 of the FMCE 3: contained within original Force (Newton I and II) cluster and revised Force Sled cluster.

in their thinking about the physics (as assumed in a resources framework), then results in each cluster should be noisy and inconsistent as well.

A. Defining revised clusters

To understand student reasoning better, we use a resources framework to help us group and analyze questions. We use the questions of the Force (Newton I and II) cluster, shown in Figs. 1–3, to illustrate.

In answering the questions in Figs. 1–3, students must determine the force on an object (sled, toy car, coin) undergoing a described motion. In terms of representations, Figs. 1 and 3 use pictorial representations, while Fig. 2 uses graphical representations. In terms of context, the questions in Figs. 1 and 2 are about motion in a single direction, while the questions in Fig. 3 involve an object that reverses the direction.

To measure dependence of student reasoning on the representational and contextual cues in Figs. 1–3, we have defined clusters which replace the old Force (Newton I and II) cluster: the Force Sled cluster (containing the questions in Fig. 1), the Force Graphs cluster (containing the questions in Fig. 2), and the Reversing Direction cluster (containing the questions in Fig. 3 as well as others). The Reversing Direction cluster has been expanded to include questions about acceleration as well as force; questions 27–29 on the FMCE inquire about the acceleration of a coin tossed in the air as it moves up and back down, isomorphic to questions 11–13 but in the context of acceleration.

Table II shows the full definitions of the revised clusters, ordered by question number on the test rather than difficulty

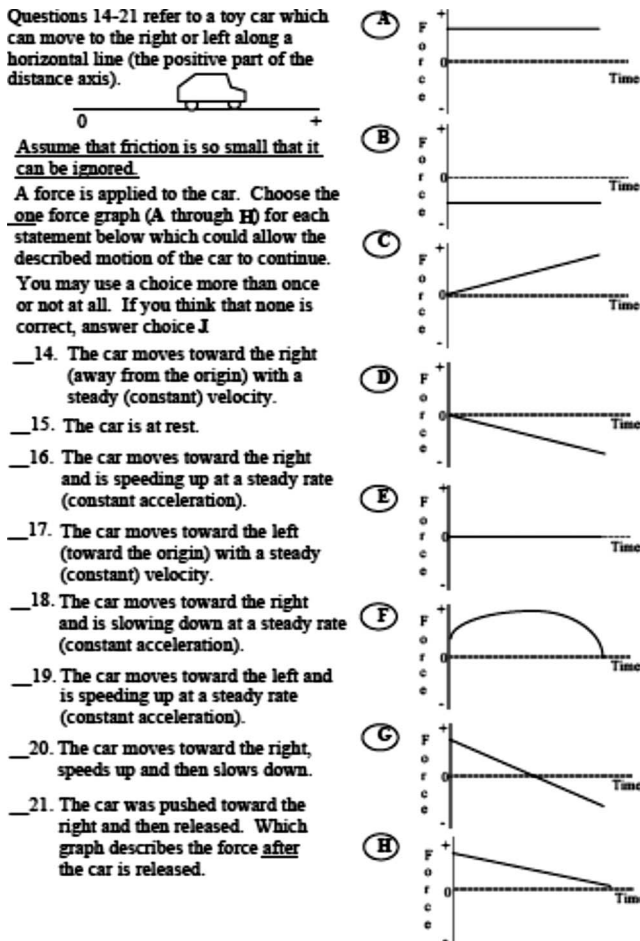


FIG. 2. Questions 14–21 of the FMCE 3: contained within original Force (Newton I and II) cluster and revised Force Graphs cluster.

of the physics material. These clusters are consistent with those used by Thornton.¹³ Note that the definitions of the original Newton III, Velocity, and Energy clusters have been directly transferred to the revised clusters. To conform with the specificity of the Force Graphs and Acceleration Graphs clusters, the Velocity cluster has been renamed the Velocity Graphs cluster. Most of the original Acceleration cluster has been transferred to the Acceleration Graphs cluster (the other Acceleration questions are in the Reversing Direction cluster). These revised clusters increase the information that can be taken from an analysis of the FMCE by highlighting and

TABLE II. Revised clusters on the FMCE.

Cluster	Questions
Force Sled	1–4, 7
Reversing Direction	8–13, 27–29
Force Graphs	14, 16–21
Acceleration Graphs	22–26
Newton III	30–32, 34, 36, 38
Velocity Graphs	40–43
Energy	44–47

Questions 8-10 refer to a toy car which is given a quick push so that it rolls up an inclined ramp. After it is released, it rolls up, reaches its highest point and rolls back down again. Friction is so small it can be ignored.



Use one of the following choices (A through G) to indicate the net force acting on the car for each of the cases described below. Answer choice J if you think that none is correct.

- (A) Net constant force down ramp
- (B) Net increasing force down ramp
- (C) Net decreasing force down ramp
- (D) Net force zero
- (E) Net constant force up ramp
- (F) Net increasing force up ramp
- (G) Net decreasing force up ramp

8. The car is moving up the ramp after it is released.
9. The car is at its highest point.
10. The car is moving down the ramp.

Questions 11-13 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (A through G) to indicate the force acting on the coin for each of the cases described below. Answer choice J if you think that none is correct. Ignore any effects of air resistance.

- A. The force is down and constant.
- B. The force is down and increasing.
- C. The force is down and decreasing.
- D. The force is zero.
- E. The force is up and constant.
- F. The force is up and increasing.
- G. The force is up and decreasing.

11. The coin is moving upward after it is released.
12. The coin is at its highest point.
13. The coin is moving downward.

FIG. 3. Questions 8–13 of the FMCE 3: contained within original Force (Newton I and II) cluster and revised Reversing Direction cluster.

isolating content areas, types of representations, and specific situations with which students may struggle. We give examples in Sec. IV.

B. Comparing to other clusters

Other ways of clustering questions on the FMCE exist. In his Ph.D. dissertation, Bao¹⁶ claimed that mixed model states are more easily detected using samples of questions that span several physical contexts. As such he defined the cluster of questions in Table III to compare students' use of two particular models (force proportional to acceleration and force proportional to velocity).³¹ His cluster contains questions from both the Force Sled (question 2) and Reversing Direction (questions 11 and 12) clusters, as well as question 5, which Thornton and Sokoloff¹⁵ suggested should only be used as a measure of reading ability rather than Newtonian thinking.

TABLE III. Question cluster defined by Bao to use model analysis with the FMCE (Ref. 16).

Question No.	$F \propto \Delta v / \Delta t$	$F \propto v$	Other
2	D	B	Others
5	D	B	Others
11	A	G	Others
12	A	D	Others

Using a resources framework, we can explain the presence of mixed model states as being the result of contextual cues or representational cues. Thus, by creating a cluster that mixes cues, he has primed his data to show mixed model states while not being able to explain the source of model mixing, be it contextual, representational, or due to deeper issues with the physics. Of course we expect our students to understand concepts in all contexts, but mixing cues makes our analysis of student reasoning much more difficult and fails to give the resolution that we, as researchers and instructors, desire.

A further weakness of the mixed context cluster is the use of question 5, which describes two kinds of motion (acceleration followed by constant velocity), and should be checked against question 2 on the FMCE for consistency. Typically, students score very well on this question (unless they have reading problems in understanding the question). Thus, adding this question necessarily skews the data toward correct model use.

We believe that it is far more beneficial for instructors to first group questions that deal with a single physical context (e.g., constant force applied to move an object horizontally) before combining questions across diverse contexts. Our approach is more consistent with the assumptions of a resources framework and gives additional insight (as described below in Sec. IV B 2). We do still observe students' use of mixed models but within a single physical context. Such a mixed model indicates a kind of inconsistency in thinking about the physics that a mixed-context cluster cannot.

III. FACETS, RESOURCES, AND MENTAL MODELS OF STUDENT REASONING ON THE FMCE

As we have stated from the outset, we use a resources framework to cluster responses on the FMCE. We are not using a conceptions approach^{14,17,18} because we are seeking a higher resolution to understand what students have mastered in their physics learning (and how best to help those who have not yet mastered the material). The resources framework can be thought of as a schema theory that emphasizes knowledge in pieces, such as phenomenological primitives (p-prims),¹⁹ facets,⁷ and resources.^{8,20} The differences and connections between a concept-based and resource-based analysis of student reasoning and teaching are discussed in more detail by Scherr.²¹

Student thinking is rarely described in terms of an individual small idea such as “closer means stronger,” for which appropriate applications are sitting by a fire to be warmer or moving from speakers at a concert to save one's hearing, and false applications include attributing the change in seasons to a difference in the distance from the earth to the sun. Instead, succinct descriptions of student thinking often require us to recognize the use of multiple resources in connection. We represent this as a type of nodal mental space network, a resource graph.⁸ The connections, or links, between these resources vary greatly in both strength and duration. Assuming that students are using a set of resources related to mechanics and kinematics, not all activated in every question, we may examine how various resources could combine to

create robust (and often incorrect) mental models (or concepts) that students use when reasoning about physics. We give several examples below.

In this section, we describe the resources that students often use when answering questions on the FMCE. In some cases, we draw resource graphs. In Sec. VI, we connect these descriptions of resources to the questions on the FMCE.

A. Dynamics

1. Newton's first and second laws

The notion that the force exerted on an object is proportional to its velocity has been reported by many researchers and is very prevalent among physics students.^{14,22,23} The $F \propto v$ model has been described as being similar to the Impetus view of physics²³ but can be described in more detail by several of Hammer's resources⁶ including “activating agency” and (in particular) a “maintaining agency” that is “dying away.”³² The activating agency resource is the notion that every event must have a cause, i.e., every object that is in motion must have had something to get it started. The maintaining agency resource embodies the idea that objects in motion must have something (some “agent”) to keep them in motion. While neither of these resources is incorrect in and of itself, the $F \propto v$ model is evident when the agent required for each of them is seen as the force applied to an object.³³ The maintaining agency when used in this context contradicts Newton's first law, but students' intuitive ideas are not unreasonable. They are, instead, consistent with years of experience in our friction-filled world where a continuous application of force is almost always needed to maintain an object's motion.

2. Newton's third law

Studies have shown that students use a variety of strategies when reasoning about the forces exerted between two interacting bodies.^{24–26} Bao *et al.*²⁴ identified four “contextual features” that students use when responding to questions regarding Newton III: velocity, mass, pushing, and acceleration. For example, an object with a larger initial speed will exert more force than an object with a smaller initial speed (during a collision), and more massive objects exert more force than less massive ones. The velocity and mass features work well together to illuminate students' implicit confusion between momentum and force. Based on their ideas about kinematics, students often have a desire to represent force as $F = mv$ (Ref. 23); furthermore, students may use the terms momentum and force interchangeably.²⁴ The pushing feature is contained in the notion that, when one object pushes another, the object that is pushing must exert more force than the object that is being pushed. This idea is typically accompanied by the reasoning that if both objects exerted the same force on each other, neither would move.

We have previously reported on students' use of three facets when considering Newton's third law: the mass dependence facet, the action dependence facet, and the velocity dependence facet.²⁵ These mental models correspond with Minstrell's facets of knowledge,⁷ in particular facets 62 (the

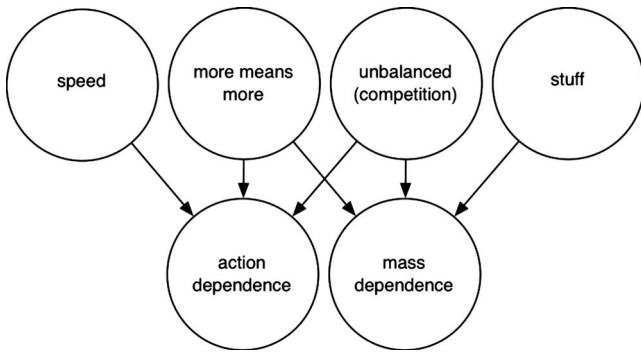


FIG. 4. The mass dependence and action dependence resources may be derived from universal primitives and observations of a scenario.

moving object or a faster moving object exerts a greater force), 63 (the more active or energetic object exerts more force), and 64 (the bigger or heavier object exerts more force).³⁴ The mass dependence facet has a direct correlation with the mass contextual feature described by Bao *et al.*²⁴ The action dependence facet combines the velocity and pushing contextual features described above to create a mental model that is applicable to both pushing and collision situations. The velocity dependence facet describes students’ use of force as an intrinsic property of an object, similar to momentum, agreeing with the velocity contextual feature described by Bao *et al.*^{24,25} Maloney²⁶ used many similar ideas to describe students’ use of a “dominance principle” to reason about the interaction between two bodies. Maloney’s dominance principle²⁶ is also very closely related to diSessa’s “Ohm’s p-prim”¹⁹ as well as Hammer’s “more is more” resource.⁶

The more is more resource might manifest itself within students’ thinking of Newton’s third law as a series of connections: the more mass or speed an object has,³⁵ the more damage it can do; the more damage an object can do, the more force it must exert on any other object. Another connection that can be made along these lines is the idea that the more an object reacts after a collision, the more force must have been applied to it.²⁷ In this case, the more is more resource is used indirectly to describe the object that is exerting the force on the object in question.

We draw a resource graph of student reasoning about Newton’s third law in order to summarize these comments. In Fig. 4 we show how four resources can combine in groups of three to create the two observed mass and action dependencies. Note that the combination of “more means more” and “unbalanced (competition)” can be interpreted as diSessa’s Ohm’s p-prim,¹⁹ in which more resistance (say, a mass in the way) requires more force for equal effect.

B. Energy

The more is more resource discussed above may also be applied when discussing students’ views of the transfer of energy. For this particular discussion we will use the scenario of a sled starting from rest at the top of an icy hill and sliding all the way down (as is seen in the FMCE). In this case, the

TABLE IV. The “Force Sled” cluster on the FMCE.

Question No.	$F \propto \Delta v / \Delta t$	$F \propto v$	Most common student response ^a
1	<i>b</i>	<i>a</i>	<i>a</i>
2	<i>d</i>	<i>b</i>	<i>b</i>
3	<i>f</i>	<i>c, g</i>	<i>c</i>
4	<i>f</i>	<i>g</i>	<i>g</i>
7	<i>b</i>	<i>e</i>	<i>e</i>

^aReference 13.

more is more resource may be used quite nicely (and correctly) in stating that the more height a hill has, the more kinetic energy (and thereby speed) the sled will have once it has reached the bottom. Students may connect more is more to other elements of the problem, instead, including more steepness or more length. Students might, for example, take the approach that the steeper a hill is, the faster (or more energetic) the sled will be when it gets to the bottom. Students most likely take this idea from their own experiences sliding down hills; the steeper hills are always more fun and get them to high speeds sooner. They attach the more is more resource to the acceleration of the sled, i.e., the rate of change of the velocity varies with the slope of the hill but not the total change in speed.

IV. MENTAL MODELS EVIDENT IN EACH CLUSTER

We return to the revised clusters defined in Sec. II, applying the resources presented in Sec. III. For each cluster of questions, we examine the possible responses to individual questions and determine which correspond with the use of correct Newtonian reasoning and which indicate the use of one of the mental models discussed in Sec. III. We will also compare these responses with the most common student responses reported by Thornton,¹³ showing that the FMCE can be interpreted in ways consistent with a resources framework. Furthermore, we will show that the use of a resources framework lets us conclude that some student responses are actually false positives (i.e., correct responses given for incorrect reasons).

A. Force Sled cluster

The Force Sled cluster (see Fig. 1) asks questions in plain language (i.e., not graphically), has no Reversing Direction questions, and deals with a single applied force that is therefore also the net force on the sled. Offered responses on this cluster include the correct idea that the net (and applied) force is proportional to its acceleration (or rate of change of velocity), as well as the notion that the net force on the sled is proportional to its velocity. Table IV shows the responses that correspond with each of these models as well as the most common student responses. The most common student responses found by Thornton¹³ are the same as those indicating a student’s use of the $F \propto v$ model. This similarity provides strong evidence that many students believe that the net force on an object is proportional to its velocity rather than

TABLE V. The “Reversing Direction” cluster on the FMCE.

Question No.	Constant downward force or acceleration	Force or acceleration in the direction of motion	$F, a \propto v$	Most common student response ^a
8–9–10	<i>a-a-a</i>	(<i>e, f, or g</i>)- <i>d</i> -(<i>a, b, or c</i>)	<i>g-d-b</i>	(<i>e, f, or g</i>)- <i>d</i> -(<i>a, b, or c</i>)
11–12–13	<i>a-a-a</i>	(<i>e, f, or g</i>)- <i>d</i> -(<i>a, b, or c</i>)	<i>g-d-b</i>	(<i>e, f, or g</i>)- <i>d</i> -(<i>a, b, or c</i>)
27–28–29	<i>a-a-a</i>	(<i>e, f, or g</i>)- <i>d</i> -(<i>a, b, or c</i>)	<i>g-d-b</i>	<i>g-d-b</i> . ^b

^aReference 13.

^bThe most common responses for questions 27–29 can be found in Ref. 28.

its acceleration. We interpret these results in terms of resource activation, though this interpretation is, in this case, not necessary.

B. Reversing Direction cluster

The Reversing Direction cluster (see Fig. 3) asks questions in which an object has been tossed in the air (or rolled up a hill). Students must think about the net force and the acceleration throughout its up-and-down free-fall motion.³⁶ Within the Reversing Direction cluster, the questions are broken into three subclusters (8–10, 11–13, and 27–29), each of which involves a single object undergoing an up-and-down motion. In each of these subclusters the students are asked about the force exerted on the object (questions 8–13) or the acceleration of the object (questions 27–29) as it goes up (questions 8, 11, and 27), at the highest point in its journey (questions 9, 12, and 28), and as it comes back down (questions 10, 13, and 29). According to Thornton and Sokoloff,³ student responses are only considered correct when all three questions within a given subcluster are answered correctly. We expand on this point below.

1. Generalized force-in-direction-of-motion model

As with the Force Sled cluster, the Reversing Direction cluster provides possible responses that correspond directly with the $F \propto v$ model (or $a \propto v$ model for questions 27–29). For the questions in the Reversing Direction cluster, however, it is beneficial to consider a generalization of the $F \propto v$ model: the force-/acceleration-in-the-direction-of-motion model. This more general model ignores the magnitude of the force or acceleration throughout each part of the motion and only describes the direction. Consider questions 11–13 (coin toss force questions). A student using the $F \propto v$ model would indicate that the force on the coin is upward and decreasing as the coin goes up, zero at the top of its motion, and downward and increasing as the coin comes back down. However, what if a second student thinks that the force is upward and constant while the coin travels up but agrees with the first student on the other two questions? This student cannot be considered as using the $F \propto v$ model but may be classified within the direction-of-motion model. In fact, our $F \propto v$ student may also be categorized as using the direction-of-motion model. In this way the direction-of-motion model allows a broader classification of students who have similar but not necessarily identical ideas.

Our decision to use the generalized direction-of-motion model is data driven. Research conducted by one of the au-

thors (T.I.S.) suggests that considering the direction-of-motion model for the Reversing Direction cluster allows many more students to be classified into a common model than the $F \propto v$ model.²⁸ Incorporating the direction-of-motion model into the results from the Force Sled or Force Graphs clusters, however, did not add any significant information. We suspect that this is due to the fact that only the Reversing Direction cluster includes scenarios in which an object moves in more than one direction during a single described motion. Moreover, the Reversing Direction scenarios do not provide information as to how the speed of the object changes throughout its motion.

Table V shows the responses that correspond with each of the models described above as well as the most common student responses. The most common student responses correspond directly with the responses indicating the use of one of our described models. We note that Thornton¹³ only provided answers that indicated the direction of the force, not its magnitude. As such, there is no way to tell from his data the likelihood that a student used the $F \propto v$ model. Also, Thornton’s work¹³ only looked at the questions pertaining to the force on the object. His results, however, were replicated by one of the authors (T.I.S.) for questions 27–29 asking about the object’s acceleration.²⁸

2. False positives in vertical toss situations

We return to the point of requiring that students answer all three questions correctly (responses *a-a-a* within each question triplet). Consider the *a-d-a* pattern of responses on a given question triplet.²⁸ This student might believe that a constant downward force is exerted on the object while it is moving both upward and downward but that no force is exerted while the object is “stopped” at the apex of its motion. This line of reasoning may come from difficulties distinguishing between instantaneous velocity and change in velocity (as it relates to acceleration).²⁹ The student may use the reasoning that since the ball has zero velocity, it is not moving; therefore, the acceleration is zero and the force exerted on the object must also be zero by Newton’s second law. For all of these reasons it is widely accepted that responses to the questions in the Reversing Direction cluster must be examined in conjunction with one another rather than individually, otherwise a student with serious problems understanding direction reversal will get 2/3 correct on this question triplet.

Furthermore, consider two students who give very similar incorrect responses: student 1 answers *g-d-b* (consistent with

TABLE VI. The “Force Graphs” cluster on the FMCE.

Question No.	$F \propto \Delta v / \Delta t$	$F \propto v$	Most common student response ^a
14	<i>e</i>	<i>a</i>	<i>a</i>
16	<i>a</i>	<i>c</i>	<i>c</i>
17	<i>e</i>	<i>a, b</i>	<i>b</i>
18	<i>b</i>	<i>h</i>	<i>h</i>
19	<i>b</i>	<i>d</i>	<i>d</i>
20	<i>g</i>	<i>f</i>	<i>f</i>
21	<i>e</i>	<i>a, h</i>	<i>h, f, a</i>

^aReference 13.

$F \propto v$), while student 2 answers *g-d-a*. The *a* given by student 2 in response to the third question indicates a constant downward force or acceleration.³⁷ The more general direction-of-motion model accounts for both sets of responses, though. Again, it would be inappropriate to give student 2 a 1/3 correct score, even though answer *a* is correct. The correct response also distinguishes the $F \propto v$ model from the direction-of-motion model. We do not believe, though, that the direction-of-motion model is 1/3 more correct than the $F \propto v$ model. For this reason, we agree with Thornton and Sokoloff³ that a student should not be considered correct for any part of a subcluster unless that student answers correctly on all three questions. We thereby avoid measuring false positives in the Reversing Direction cluster.

C. Force Graphs cluster

The Force Graphs cluster (see Fig. 2) asks questions about motions sometimes identical in physics content to those found in the Force Sled cluster, but differing in presentation. Students are provided with a description of the motion of a toy car and asked to select a graph that depicts the force exerted on the car. All of the correct responses indicate that the applied force is either zero or nonzero and constant. Table VI shows how responses to the questions in the Force Graphs cluster correspond with the various mental models as well as the most common student responses.

As with the Force Sled cluster, the most common student responses for the Force Graphs cluster correspond almost exactly with the responses indicating a student’s use of the $F \propto v$ model. We separate the clusters to observe if students master the content of one cluster before the other. Research suggests⁹⁻¹¹ that students do not display as much knowledge when working with graphs as when using descriptive language, and data from the FMCE support the separation of questions into two clusters.¹⁴ The answers to the physics depend on the context and format of the question. Philosophically, this supports the use of a resources framework, which can account for differences in reasoning based on contextual and representational differences in resource activation.

Note that in Table VI we designate multiple responses for a single model on question 21. Students are asked about the force on a car once it has been released (after being pushed). On one hand, response *a* seems to fit perfectly with the

TABLE VII. The “Acceleration Graphs” cluster on the FMCE.

Question No.	$a \propto \Delta v / \Delta t$	$a \propto v$	Most common student response ^a
22	<i>a</i>	<i>e</i>	<i>e</i>
23	<i>b</i>	<i>g</i>	<i>f</i>
24	<i>c</i>	<i>b</i>	<i>b</i>
25	<i>b</i>	<i>f</i>	<i>f</i>
26	<i>c</i>	<i>a</i>	<i>a</i>

^aReference 13.

$F \propto v$ model (see Fig. 2); the car moves at a constant velocity so a constant force must be applied. On the other hand, what if the students do not ignore friction (though they are explicitly told to do) or use an “impetus/force dies away” model? In each of these cases, the car would slow down at a (perhaps) steady rate, indicating a positive yet decreasing force (response *h*). Both of these responses can be considered a $F \propto v$ model in the sense of the need for a maintaining agency, with response *h* including the use of the dying away resource.

Also consider question 17, where research by one of the authors (T.I.S.) has shown that some students who primarily use the $F \propto v$ model will choose response *a* instead of *b* (the $F \propto v$ response). Response *a* is not entirely different from *b* as it is congruent with the $F \propto v$ model in magnitude but not direction. This may correspond to a confusion between left and right as negative and positive, indicating a difficulty with coordinate systems rather than with forces. As such we have decided to categorize response *a* to question 17 as indicative of the $F \propto v$ model if the student displays use of that model in other responses to this cluster. Strong empirical evidence and the fact that question 17 is one of only two questions in the Force Graphs cluster to describe a motion with constant velocity heavily influenced our decision to consider response *a* as corresponding to the $F \propto v$ model.

D. Acceleration Graphs cluster

The Acceleration Graphs cluster is similar to the Force Graphs cluster. Students are asked about the acceleration of a toy car undergoing various types of motion. Again, students must choose the graph they believe best represents the acceleration of the car for each scenario. It should be noted that the parenthetical reminders of “constant acceleration” that are found in the questions of the Force Graphs and Force Sled clusters are omitted from these questions. We examine the Acceleration Graphs cluster from the perspective of students’ difficulties distinguishing between the concepts of acceleration and velocity reported by Trowbridge and McDermott²⁹ to form a type of acceleration-proportional-to-velocity ($a \propto v$) model. The most common student responses shown in Table VII correspond closely with the $a \propto v$ model.

We note a discrepancy between the $a \propto v$ model and the most common response for question 23. This question asks students to choose the appropriate graph of acceleration vs time for a car that “moves toward the right (positive direc-

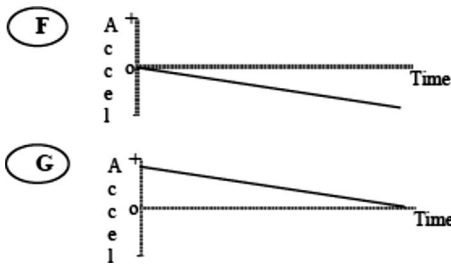


FIG. 5. Responses *f* and *g* for the Acceleration Graphs cluster.

tion), slowing down at a steady rate.”³ Figure 5 shows responses *f* and *g* that correspond to the most common student response and the $a \propto v$ response, respectively. Visually, responses *f* and *g* are incredibly similar, with identical magnitude slopes. They are also presented above one another (see Fig. 5). Even though response *f* would only accurately fit the $a \propto v$ model for a car moving to the left and speeding up (as in question 25), it is not surprising that students would choose *f* for question 23. Again, as with question 17, the problem may indicate issues with coordinate systems more than the relationship between acceleration and velocity.

E. Newton III cluster

The Newton’s third law cluster is the only cluster that commonly elicits two different incorrect student models: the mass dependence model and the action dependence model described above. Table VIII shows how the responses to the questions in the Newton III cluster separate into these models.

The most common student responses shown in Table VIII incorporate aspects of both the mass dependence and the action dependence models.³⁸ We see that the most common set of responses shows more agreement with the action dependence model (questions 32, 34, 36, and 38) than the mass dependence model (question 30). Response *b* as the most common response for question 38 might seem a bit ambiguous (as it can be classified as either mass or action dependence), but one can use response *c* for question 36 to then categorize both as uses of the action dependence model. As with much of our analysis, this requires the assumption that

TABLE VIII. The “Newton III” cluster on the FMCE.

Question No.	Forces equal	Mass dependence	Action dependence	Most common student response ^a
30	<i>e</i>	<i>a</i>	XX	<i>a</i>
31	<i>e</i>	<i>a</i>	<i>b</i>	<i>e, f</i>
32	<i>e</i>	<i>a</i>	<i>b</i>	<i>b</i>
34	<i>e</i>	XX	<i>b</i>	<i>b</i>
36	<i>a</i>	<i>b</i>	<i>c</i>	<i>c</i>
38	<i>a</i>	<i>b</i> ^b	<i>b, b, c</i>	<i>b</i>

^aReference 28.

^bCategorization for this response depends on the student’s choice for question 36.

Questions 30-34 refer to collisions between a car and trucks. For each description of a collision (30-34) below, choose the one answer from the possibilities A through J that best describes the forces between the car and the truck.

- A. The truck exerts a greater amount of force on the car than the car exerts on the truck.
- B. The car exerts a greater amount of force on the truck than the truck exerts on the car.
- C. Neither exerts a force on the other, the car gets smashed simply because it’s in the way of the truck.
- D. The truck exerts a force on the car, but the car doesn’t exert a force on the truck.
- E. The truck exerts the same amount of force on the car as the car exerts on the truck.
- F. Not enough information is given to pick one of the answers above.
- J. None of the answers above describes the situation correctly.

In questions 30 through 32 the truck is much heavier than the car.



- ___ 30. They are both moving at the same speed when they collide. Which choice describes the forces?
- ___ 31. The car is moving much faster than the heavier truck when they collide. Which choice describes the forces?
- ___ 32. The heavier truck is standing still when the car hits it. Which choice describes the forces?

In questions 33 and 34 the truck is a small pickup and is the same weight as the car.



- ___ 33. Both the truck and the car are moving at the same speed when they collide. Which choice describes the forces?
- ___ 34. The truck is standing still when the car hits it. Which choice describes the forces?

FIG. 6. Questions involving collisions within the Newton III cluster of the FMCE. Please note that question 33 is not included within analysis of the FMCE. 14

students are reasoning relatively consistently from question to question.

1. False positives in collision situations

The assumption of consistent reasoning has serious consequences when one considers question 31. Many students who answer *e* on question 31 (the correct answer) answer incorrectly on questions 30, 32, and 34 (Ref. 39) (see Fig. 6).

We can infer with a fair degree of certainty why most students respond the way they do to questions 30 and 32. In question 30, both vehicles are moving at the same speed before the collision (making action dependence a moot point), but the truck is much heavier than the car, causing students to lean heavily toward mass dependence reasoning. In question 32, the truck is still much heavier than the car, but it is not initially moving. As such, the greater “activeness” of the car wins out and students use action dependence. However, what happens between these situations? If the same two objects can interact in two different ways to get opposite results, there must be a situation in which the effects of mass dependence and action dependence will compromise or cancel out. In question 31 the smaller, lighter car is initially moving much faster than the bigger, heavier truck, but the truck is moving. In this case our “most common student” must decide how to deal with mass dependence ideas from question 30 and action dependence ideas from question 32. Response *e* is one logical conclusion. The two

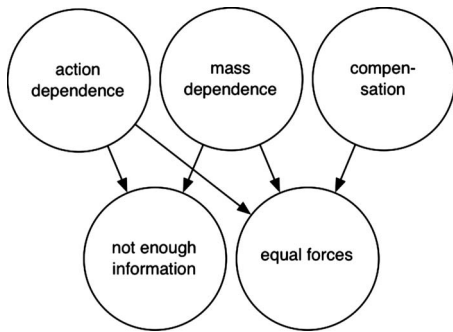


FIG. 7. The mass dependence and action dependence resources may be used together to cancel out some of their effects based on the situation (as in question 31).

effects cancel each other out to result in the car and the truck exerting forces on each other that are equal in magnitude but opposite in direction. A more discerning student, however, may feel that the effects will cancel each other to some degree but not necessarily completely, leading to response *f*, which indicates that more information is needed. Figure 7 shows how the mass dependence and action dependence resources may “balance” to produce the correct conclusion.

To avoid the measurement of false positives, we suggest clustering responses 30–32 into a “triplet” subcluster as done on the question triplet subclusters in the Reversing Directions cluster. Otherwise, one incorrectly rewards students for a situation where two different wrongs do, in fact, make a right.

2. False positives in pushing situations

A situation exists where two identical wrongs make a right as well. As shown in Table VIII the most common responses for question 36 and 38 are *c* and *b*, respectively (the questions are given in Fig. 8). Many students, however, choose responses *c* and *a* for these two questions.⁴⁰ Response *a* for question 38 indicates a correct answer of equal and opposite forces exerted by the two vehicles on each other. Response *c* for question 36, on the other hand, indicates the student’s use of the action dependence line of reasoning.

We again assume some consistency of student reasoning within a cluster of questions (that are contextually and representationally similar). In question 36 the smaller car is pushing the heavier truck and the two are speeding up. Use of the action dependence resource suggests that the car is exerting a greater force than the truck (response *c*). This result agrees with the pushing contextual feature reported by Bao *et al.*²⁴ In question 38, however, the two vehicles are at a constant cruising speed, and the truck begins to apply its brakes, causing both vehicles to slow down. Response *b* for question 38 (the truck exerts more force) might be indicative of action dependence reasoning.

When the truck begins applying its brakes, it may become the more active object in the student’s mind, causing the vehicles to slow down. Once again, we have the possibility of two effects of incorrect reasoning acting against one another. The car is the active agent, pushing the truck forward.

Questions 35-38 refer to a large truck which breaks down on the road and receives a push back to town by a small compact car.



Pick one of the choices A through J below which correctly describes the forces between the car and the truck for each of the descriptions (35-38).

- A. The force of the car pushing against the truck is equal to that of the truck pushing back against the car.
- B. The force of the car pushing against the truck is less than that of the truck pushing back against the car.
- C. The force of the car pushing against the truck is greater than that of the truck pushing back against the car.
- D. The car’s engine is running so it applies a force as it pushes against the truck, but the truck’s engine isn’t running so it can’t push back with a force against the car.
- E. Neither the car nor the truck exerts any force on each other. The truck is pushed forward simply because it is in the way of the car.
- J. None of these descriptions is correct.

- ___ 35. The car is pushing on the truck but not hard enough to make it move.
- ___ 36. The car, still pushing on the truck, is **speeding up** to get to cruising speed.
- ___ 37. The car, still pushing the truck, is at cruising speed and continues to travel at the **same speed**.
- ___ 38. The car, still pushing the truck, is at cruising speed when the truck puts on its brakes and causes the car to **slow down**.

FIG. 8. Questions involving one vehicle pushing another within the Newton III cluster of the FMCE. Please note that questions 35 and 37 are not included within analysis of the FMCE (Ref. 14).

The truck is a second active agent, pushing back against the car. A student might believe that the two effects will perfectly balance each other and might arrive at the correct response (*a*) in which the two vehicles exert equal amounts of force on one another. This result is quite similar to that found for response *e* on question 31 with the exception that question 38 presents the opportunity for students to use two conflicting versions of action dependence and eliminates the need for mass dependence.

F. Velocity Graphs cluster

The Velocity Graphs cluster is very similar to the previously described “graphing” clusters. Students are presented with various descriptions of a car’s motion, and they must choose the correct velocity vs time graphical representation of the motion. As with the Acceleration Graphs questions, the incorrect model we examine is derived from the Trowbridge-McDermott studies³⁰ of students’ understanding of kinematics and their difficulty distinguishing position from velocity. This velocity/position confusion model is also closely related to Beichner’s proposition⁹ that students view graphs as a picture of the situation no matter what the axes

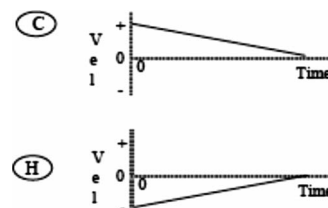


FIG. 9. Responses *c* and *h* for the Velocity Graphs cluster.

TABLE IX. The “Velocity Graphs” cluster on the FMCE.

Question No.	Correct model for velocity	Velocity/position confusion	Most common student response ^a
40	<i>a</i>	<i>d</i>	<i>d</i>
41	<i>f</i>	<i>g</i>	<i>g</i>
42	<i>b</i>	<i>c, h</i>	<i>c</i>
43	<i>d</i>	XX	XX ^b

^aReference 13.

^bThornton reported the most common student response for question 43 as “not significant” (Ref. 13).

indicate.⁹ Table IX shows how the responses in this cluster correspond with the various student models.

Once again, more than one response to a single question may indicate the use of our incorrect model. On question 42, the toy car is said to be “moving toward the left (toward the origin) at a steady (constant) velocity.” Responses *c* and *h* (shown in Fig. 9) both indicate a graph that gets steadily closer to the horizontal axis as time progresses. For response *c* a student could be picturing the car starting at the right and moving toward “0,” and students choosing *h* could be triggered by the word “left” to choose a graph that depicts negative velocity.

As an aside, we note that questions 17, 23, and 42 form a cluster, which lets one see if students have problems understanding coordinate systems, allowing for a finer grained analysis of students’ understanding of force and motion. The FMCE only measures this topic implicitly, though, and the cluster is therefore relatively badly defined.

G. Energy cluster

The Energy cluster on the FMCE contains questions that ask students to reason about the speed and kinetic energy of a sled after sliding down a hill. The incorrect model for the questions in the Energy cluster, as described in Sec. III, corresponds with the idea that steeper hills will cause a greater change in speed and kinetic energy as the sled slides down. Table X shows how the possible responses to questions in this cluster are divided among the correct and this incorrect model. The most commonly given incorrect answers correlate with the responses indicating a student’s use of the slope dependent model.

V. APPLICATIONS OF REVISED CLUSTERING TO STUDENT RESPONSES

Having shown that clustering based on a resources framework accurately accounts for the most common incorrect an-

TABLE X. The “Energy” cluster on the FMCE.

Question No.	Energy/speed depends on height	Energy/speed depends on slope	Most common student response ^a
44	<i>b</i>	<i>a</i>	<i>a</i>
45	<i>b</i>	<i>a</i>	<i>a</i>
46	<i>a</i>	<i>c</i>	<i>c</i>
47	<i>a</i>	<i>c</i>	<i>c</i>

^aMost common student responses for the Energy cluster discovered by research reported in Ref. 28.

swers, we now demonstrate the applicability of our method by discussing several sample student responses.⁴¹ This analysis is relatively brief and looks only at individual students rather than a full set of class data. To show the added resolution of our approach, we show that results identical using the old content-based correct or incorrect clustering are no longer seen as identical when using our revised clustering.

Consider the student responses to the questions contained in the Force Sled and Force Graphs clusters presented in Table XI. Both students answered 5 out of these 12 questions correctly. Using the older method of analysis, these students would be seen as identical in their understanding of Newton’s first and second laws (both being 42% correct).⁴²

Using our revised clustering, we find valuable additional information. Student 1 answered all of the Force Sled questions correctly and all of the Force Graphs questions incorrectly. Perhaps student 1 has a strong understanding of Newton’s laws but a poor understanding of graphs. Student 2, on the other hand, answered two of the five Force Sled questions correctly and three of the seven Force Graphs questions correctly. This may indicate that student 2 has a generally weak understanding of Newton’s laws but is not hindered by graphical representations. A classroom populated primarily by students like student 1 should receive very different instruction than a classroom populated by students like student 2. One class needs help on learning graphing, the other needs help on the physics.

Another example illustrates other information that can be gathered by attending to different response patterns in our revised clusters. Consider the sample responses to the Force Sled questions presented in Table XII. By the old method, all of these students would be considered identical since none of them answered any of the questions correctly.

Looking at response patterns shows that there is far greater detail to be found. The responses given by student 3 align exactly with those corresponding to the $F \propto v$ model. Student 4 gives three responses that match the $F \propto v$ model,

TABLE XI. Sample responses to questions in the Force Sled and Force Graphs clusters.

Question	1	2	3	4	7	14	16	17	18	19	20	21
Correct response	<i>b</i>	<i>d</i>	<i>f</i>	<i>f</i>	<i>b</i>	<i>e</i>	<i>a</i>	<i>e</i>	<i>b</i>	<i>b</i>	<i>g</i>	<i>e</i>
Student 1	<i>b</i>	<i>d</i>	<i>f</i>	<i>f</i>	<i>b</i>	<i>a</i>	<i>c</i>	<i>g</i>	<i>h</i>	<i>d</i>	<i>f</i>	<i>b</i>
Student 2	<i>a</i>	<i>b</i>	<i>f</i>	<i>g</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>h</i>	<i>b</i>	<i>a</i>	<i>f</i>	<i>e</i>

TABLE XII. Sample incorrect responses to the questions in the Force Sled cluster.

Question	1	2	3	4	7
$F \propto v$	<i>a</i>	<i>b</i>	<i>c, g</i>	<i>g</i>	<i>e</i>
Student 3	<i>a</i>	<i>b</i>	<i>c</i>	<i>g</i>	<i>e</i>
Student 4	<i>a</i>	<i>b</i>	<i>d</i>	<i>g</i>	<i>d</i>
Student 5	<i>g</i>	<i>f</i>	<i>e</i>	<i>a</i>	<i>e</i>

but two do not. Research conducted by one of the authors (T.I.S.) suggests that these two outliers (response *d* for questions 3 and 7) may indicate a student's use of the dying away resource in conjunction with the $F \propto v$ model. Student 5 gives only one response that fits with the $F \propto v$ model and four responses that do not align with any consistent and well-defined model of reasoning. The incorrect student responses may indicate very different ways of thinking about forces and suggest different instructional strategies for getting each of them to understand Newton's laws. As such, these students should in no way be considered identical. For these reasons it is imperative that we consider *how* our students answer questions, not just the correctness of their responses.

VI. SUMMARY

Our goal in this paper has been to increase the observational resolution of a commonly accepted analysis tool in physics education research (PER). We have described a theory-driven method for clustering questions on the FMCE. We use a resources framework to account for the most common student responses as reported by either Thornton¹³ or in the work of one of the authors.²⁸ Our clustering allows us to categorize correct and incorrect responses using a single language of resource activation.

Our clusters (Force Sled, Reversing Direction, Force Graphs, Acceleration Graphs, Newton III, Velocity Graphs, and Energy) take into account the physics content, the contextual aspects, and the representations used to ask the questions. We show that students' incorrect responses to ques-

tions on the FMCE may be indicative of a variety of mental models that correspond with well documented research results. Using the resources framework, we can analyze sets of questions within some clusters (Reversing Direction and Newton III) to describe some correct student responses as false positives.

We have presented interpretations for the most common incorrect responses for each question, but this is in no way an exhaustive list of the possible mental models that may be used by students while answering questions on the FMCE. Additional patterns of responses should be examined for prevalence among student responses and analyzed in terms of mental models that may be indicated by each. For example, questions 17, 23, and 42 are a "coordinate system" cluster that has not yet been evaluated but may affect student responses on other clusters. Also, a second tier of mixed-context clusters of questions (such as that of Bao¹⁶) could be created that "slice" data in different ways.

To show the value of our analysis, we have shown how student response patterns on sets of questions can be analyzed in greater detail than allowed by the previous correct or incorrect analysis. Such an analysis has implications for instructors seeking to target instruction in their classroom to those areas where students have their greatest needs.

Research tools such as the FMCE are most effective to educators and researchers only when responses are examined to determine not only whether or not students have the correct ideas but also what ideas they do have (correct or otherwise) and how consistently they use these ideas across similar questions. Our clustering allows such an analysis, giving insight into both how we model student thinking and how we could better address student needs in the classroom.

ACKNOWLEDGMENTS

We thank three reviewers for their valuable comments and insights on an earlier version of this paper. This paper was supported in part by NSF under Grants No. DUE-0510614 and No. REC-0633951. Portions of this paper were originally published in Ref. 28.

¹D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, *Phys. Teach.* **30**, 141 (1992).

²R. 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).

³R. 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).

⁴M. C. Wittmann, available at <http://perlnet.umephy.maine.edu/research/Wittmann2002RTPpaper.pdf>

⁵E. F. Redish, J. M. Saul, and R. N. Steinberg, On the effectiveness of active-engagement micorcomputer-based laboratories, *Am. J. Phys.* **65**, 45 (1997).

⁶D. Hammer, Student resources for learning introductory physics, *Am. J. Phys.* **68** S52 (2000).

⁷J. Minstrell, in *Research in Physics Learning: Theoretical Issues and Empirical Studies*, Proceedings of an International Workshop, Bremen, Germany, edited by R. Duit, F. Goldberg, and H. Niedderer (IPN, Kiel, 1992), pp. 110–128.

⁸M. C. Wittmann, Using resource graphs to represent conceptual change, *Phys. Rev. ST Phys. Educ. Res.* **2**, 020105 (2006).

⁹R. J. Beichner, Testing student interpretation of kinematics graphs, *Am. J. Phys.* **62**, 750 (1994).

¹⁰P. B. Kohl and N. D. Finkelstein, Effects of representation on students solving physics problems: A fine-grained characterization, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010106 (2006).

- ¹¹P. B. Kohl and N. D. Finkelstein, Student representational competence and self-assessment when solving physics problems, *Phys. Rev. ST Phys. Educ. Res.* **1**, 010104 (2005).
- ¹²D. J. Dykstra, 2002, available at <http://www.boisestate.edu/physics/dykstra/WTK1.pdf>
- ¹³R. K. Thornton, NATO Advanced Studies Institute, Series (Springer-Verlag, Berlin, 1993).
- ¹⁴R. K. Thornton, in *Thinking Physics for Teaching*, edited by C. Tarsitani, C. Bernardini, and M. Vincentini (Plenum, New York, 1995).
- ¹⁵R. K. Thornton and D. R. Sokoloff, Learning motion concepts using real-time microcomputer-based laboratory tools, *Am. J. Phys.* **58**, 858 (1990).
- ¹⁶L. Bao, Ph.D. thesis, University of Maryland, 1999.
- ¹⁷J. H. Wandersee, J. J. Mintzes, and J. D. Novak, *Handbook of Research of Science Teaching and Learning* (Macmillan, New York, 1994), pp. 177–210.
- ¹⁸S. Vosniadou and W. F. Brewer, Mental Models of the Earth: A Study of Conceptual Change in Childhood, *Cognit Psychol.* **24**, 535 (1992).
- ¹⁹A. diSessa, Toward an Epistemology of Physics, *Cogn. Instruct.* **10**, 105–225 (1993).
- ²⁰D. Hammer, A. Elby, R. E. Scherr, and Edward F. Redish, *Transfer of Learning: Research and Perspectives* (Information Age, Greenwich, CT, 2005).
- ²¹R. E. Scherr, Modeling student thinking: An example from special relativity, *Am. J. Phys.* **75**, 272 (2007).
- ²²D. E. Brown and J. Clement, Overcoming misconceptions via analogical reasoning: abstract transfer versus explanatory model construction, *Instr. Sci.* **18**, 237–261 (1989).
- ²³I. A. Halloun and D. Hestenes, Common sense concepts about motion, *Am. J. Phys.* **53**, 1056 (1985).
- ²⁴L. Bao, K. Hogg, and D. Zollman, Model analysis of fine structures of student models: An example with Newton's third law, *Am. J. Phys.* **70**, 766 (2002).
- ²⁵T. I. Smith and M. C. Wittmann, Comparing three methods for teaching Newton's third law, *Phys. Rev. ST Phys. Educ. Res.* **3**, 020105 (2007).
- ²⁶D. P. Maloney, Rule-governed approaches to physics: Newton's third law, *Phys. Educ.* **19**, 37 (1984).
- ²⁷D. Hammer and A. Elby, Tapping epistemological resources for learning physics, *J. Learn. Sci.* **12**, 53 (2003).
- ²⁸T. I. Smith, M.S.T. thesis, University of Maine, 2003.
- ²⁹D. E. Trowbridge and L. C. McDermott, Investigation of student understanding of the concept of acceleration in one dimension, *Am. J. Phys.* **49**, 242 (1981).
- ³⁰D. E. Trowbridge and L. C. McDermott, Investigation of student understanding of the concept of velocity in one dimension, *Am. J. Phys.* **48**, 1020 (1980).
- ³¹Table III also shows the response to each question that corresponds with each of these models.
- ³²These and many other resources are lightly derived from diSessa's p-prims (Ref. 19).
- ³³One could argue that this is only a problem if the students use the net force for each of these resources; however, the fact that the FMCE presents questions that involve a single applied force on a frictionless surface makes this a moot point.
- ³⁴Definitions of these facets can be found in Refs. 7,25.
- ³⁵Possibly tacit indicators of momentum or kinetic energy.
- ³⁶Note that, while the ramp on which the toy car rolls in questions 8–10 prevents it from truly being in free fall, the up and down motion of the car travelling under its own volition is analogous to the coin toss in questions 11–13 (Ref. 13).
- ³⁷The $F \propto v$ model would require an increasing downward force for each of these questions.
- ³⁸It should be noted that Thornton's study (Ref. 13) did not include the questions from the Newton III cluster. The most common student responses are those reported by one of the authors (T.I.S.) in Ref. 28.
- ³⁹Question 33 is not included in analysis of the FMCE (Ref. 14).
- ⁴⁰These responses often occur with the most common responses listed in Table VIII for questions 30–34 (Ref. 28).
- ⁴¹All sample responses are taken from research reported in Ref. 28.
- ⁴²For simplicity we will ignore the Reversing Directions component to the original Force (Newton's first and second laws) cluster.