Model analysis of fine structures of student models: An example with Newton's third law

Lei Bao[a)]
Department of Physics, The Ohio State University, 174 West 18th Avenue, Columbus, Ohio 43210

Kirsten Hogg and Dean Zollman
Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, Kansas 66506

(Received 29 March 2001; accepted 1 July 2001)

In problem-solving situations, the contextual features of the problems affect student reasoning. Using Newton's third law as an example, we study the role of context in students' uses of alternative conceptual models. We have identified four contextual features that are frequently used by students in their reasoning. Using these results, a multiple-choice survey was developed to probe the effects of the specific contextual features on student reasoning. Measurements with this instrument show that different contextual features can affect students' conceptual learning in different ways. We compare student data from different populations and instructions and discuss the implications.

[DOI: 10.1119/1.1484152]

I. INTRODUCTION

Over the past two decades, a significant amount of research has been conducted on students' use of incorrect reasoning in solving physics problems.[1] This reasoning has been described as preconceptions, misconceptions, alternative concepts, etc. At both college and pre-college levels, studies have shown that the formation and application of this reasoning are strongly context dependent.[2–5] In problem-solving situations, the context setting of the problems can have a significant influence on students' reasoning.[6,7] Many physics concepts, such as Newton's third law (Newton III), can involve a variety of contextual features. When not treated properly, these features can increase the difficulty of both assessing student learning and implementing effective instruction. Therefore, the details of how different contextual features can affect students' reasoning are of great importance to researchers and instructors.

The issue of context dependence has been widely studied in the literature, and a variety of definitions have been proposed for what could be considered to be contexts. In this paper, we focus on the set of context factors that are directly embedded in the content knowledge that students are studying, that is, content-based context factors.[8]

Based on the context-dependency of the learning process, we have developed a modeling method, *model analysis*, with which different types of student reasoning are described by *student models*.[9] Both the models that students' use and the contexts in which they use the models are objects of the analysis. The assessment investigates how students apply a model of a concept as well as how this application varies as the context is changed. In this approach one does not simply say that a student can or cannot apply a correct model of a given concept. Instead, one states that the student is likely to use a particular model with a certain probability on problems related to a concept. Furthermore, we can begin to understand in which contexts it is difficult for the student to apply the model. Thus, the researcher can start to build a representation of the student's knowledge status in term of *model states* with respect to the concept. This representation provides the basis for the development of a number of numerical methods to extract information on students' uses of their models when responding to research-based multiple-choice instruments. With specially designed instruments, these numerical tools can also be used to assess the effects of specific contextual features on students' use of models.

There are a number of popular instruments for probing students' broad understanding of concepts such as Newton's laws.[10] However, for complicated concepts such as Newton III, existing instruments were often designed having multiple contextual features entangled in a single question. Thus, probing the isolated effects of a single contextual feature in the formation and application of students' models is difficult.[11] In this research, we aim to develop a new type of instrument where a single question only measures the effects of a single contextual feature involved with a particular concept.

In Sec. II we present a brief review of the literature on Newton III and the involvement of contextual features in student reasoning. Section III describes our research on identifying the important contextual features. In Sec. IV, we discuss the measurement and introduce our research on the development of a new instrument. Section V gives an example of applying this instrument with quantitative data analysis. In Sec. VI, we discuss the implications of this research.

II. STUDENT MODELS OF NEWTON'S THIRD LAW AND THE INVOLVEMENT OF CONTEXTUAL FEATURES

Before going into more detail on the analysis, we briefly review what we mean by a model.[12] A model is a functional mental construct that is associated with a specific concept (or a topic) and can be applied directly in real context settings relevant to the concept to obtain explanatory results. Models have direct causal relations to the responses students generate in various problem-solving situations. Other researchers have also studied this issue and used terms such as facets,[13] mental models,[14] and student views.[15] A comparison with the literature reveals that these terms represent mental constructs that are similar to what we call models. However, in our definition, models have explicit attributes with respect to contexts and are considered to have direct causal relations to the responses produced by the students. In mental operations, models are usually involved in an explicit manner. In our research, we study the models that students have or develop in learning physics. For convenience, we call these student models.[16]

For a particular physics concept, we can identify a finite set of commonly recognized models.[17] Such a set of models usually consists of one correct expert model and a few incorrect or partially correct student models. We will refer to these as the *common models* of this particular concept. These models are common to a group of students with similar backgrounds and the existence of these models can be verified through research. In defining the set of common models, we consider all the possible forms of students' models: the ones that students have prior to instruction, the ones that students are likely to create on the spot when exposed to new contexts relevant to the concept, and the ones that students can develop during and after instruction as a result of learning interactions.

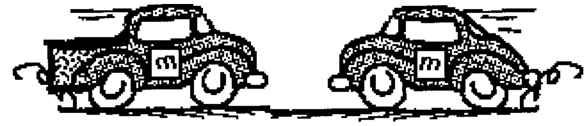
When a single student responds to a set of questions related to a particular physics concept, the student's use of models usually falls into one of the two categories: (1) the student can use one of the common models and is consistent in using it in solving all questions; or (2) the student can hold different common models at the same time and is inconsistent in using them, that is, the student can use one of the common models on some questions and use a different model on other questions, even though all the questions are related to a single concept and are seen as equivalent by experts. The different situations of the student's use of models are described in terms of student *model states*.[18] The first case corresponds to a consistent model state and the second case is a mixed model state. These model states can be measured and represented mathematically by a multidimensional probability vector in a *model space* spanned by the set of common models. We can also measure the model states of a population and study the performance of a class.

In physics education research, student understanding on topics in introductory mechanics has been thoroughly studied for several decades. Based on this existing knowledge, we can obtain a rather clear picture of the possible forms of the student models used by students in most topics of mechanics. For example, for the concept of the relationship between force and motion, we were able to identify three common models based on the results from our own research and the literature.[19] These models involve a single contextual feature—the velocity of the moving object.

On the other hand, models used by students in association with Newton III show much greater complexity, involving not one but several different contextual features. For example, in a study of students' reasoning related to this concept, Maloney found that college students use some sort of dominant principle,[20] where students think that during an interaction, the dominating object exerts a larger force. The dominance can come from a number of sources, such as (a) a greater mass and (b) the active initiator of a force (in contrast to a reaction force). Apparently, these two commonly occurring issues are often embedded in the contextual settings of physics questions on Newton III. Students' responses obtained from these questions reflect the students' understandings of the related concept, which are, in part, built on the ways that the students consider how the different contextual features are involved.

Although studies of student learning of mechanics have successfully identified the important contextual features involved in students' reasoning about Newton III, further research is needed to investigate how the different contextual features may independently (or in combination) affect students' learning. There have been studies on the effects of context on student reasoning; however, these often focus on questioning the consistency of students' use of their conceptual models in different contexts.[21] In a recent study on student understanding of forces,[22] Palmer found two types of contextual effects, primary and secondary, based on the strength of the influence that a particular contextual feature can have on student reasoning. In his research, the context is considered as an external factor that affects student reasoning.

With the model analysis and the cognitive representations we develop in this paper, the context is considered as a significant part of the student reasoning itself, and we use the contextual features as the basis for studying students' conceptual understanding.



As shown in the figure, a collision happens between a small pickup truck and a car. The small truck has the **same mass** as the car does. At the time of collision, both vehicles travel at a constant speed but the small truck is moving at a **slower speed** than the car. Describe the forces at the moment they collide.

Fig. 1. Open-ended interview question on Newton III with the contextual feature of velocity.

III. CONTEXTUAL FEATURES ASSOCIATED WITH NEWTON III

From existing research, we can identify two contextual features that are frequently used by students in their reasoning when working with problems associated with Newton III: (a) mass and (b) the initiator of the action (who is pushing). Because students often associate force with velocity,[23] our experience suggests two additional possibilities: (c) velocity and (d) acceleration.

To validate this speculation, we interviewed 9 students from an introductory physics class (*Concepts of Physics*) at Kansas State University.[24] The class has no math requirement and is for students majoring in elementary school education. The interviews were conducted in the middle of the course, about two weeks after the students had finished studying Newton III. The students volunteered to be interviewed; no attempt was made to obtain a representative sample.

In the interviews, the students were asked to think aloud about their reasoning on questions designed with the four contextual features. The protocol was designed so that each question involved only a single contextual feature. Figures 1, 2, and 3 are sample questions that are designed with isolated contextual features of velocity, mass, and pushing, respectively.

From students' responses in these interviews, we found that, in general, many students (7/9) consistently employed the dominance viewpoint in describing the forces on the

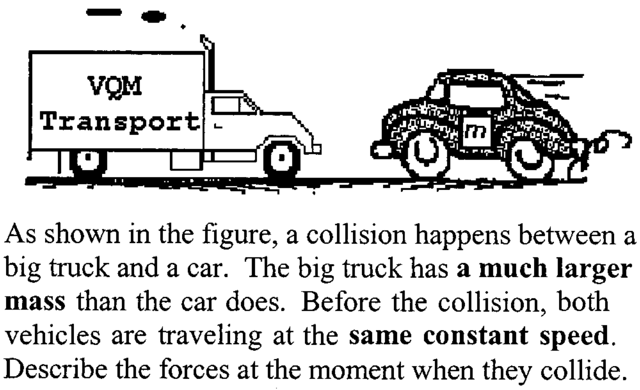


Fig. 2. Open-ended interview question on Newton III with the contextual feature of mass.

objects—the object with dominating features applies a greater force. The dominating features are selected from three context features: velocity (V), mass (M), and the initiator of the action (pushing P). Table I shows some typical incorrect student reasoning identified in the interviews.

Some students (2/9) were found to be in an explicitly confused state[25] and used mixed ideas (combining the correct model and the incorrect dominance viewpoint) on questions related to the three context features. For example, on the question shown in Fig. 1, Kathy said:

> "Same amount of force but I am not sure. Because when two things collide, they exert the same amount of force. I don't know why it is always equal and opposite. Because I think speed might have something to do with it.... It is common sense that something moving faster is going to have more force. Now I am not sure."

Later in the interviews, a modified version of the question shown in Fig. 1 was revisited, and we asked students to "consider the case that before the collision, the car is traveling at a constant speed while the truck starts slow and is speeding up. At the moment of collision, both vehicles happen to be traveling at the same speed. Describe the forces between them at the moment when they collide." When we mentioned that at the instance of the collision, both objects have the same velocity, all but one student claimed that the force should be equal and considered the acceleration (A) irrelevant.[26] Therefore, we conclude that most students we interviewed did not consider acceleration as a significant factor in their reasoning on questions related to Newton III. In our later analysis, we will keep this context feature to see if other populations might deal with this feature differently.

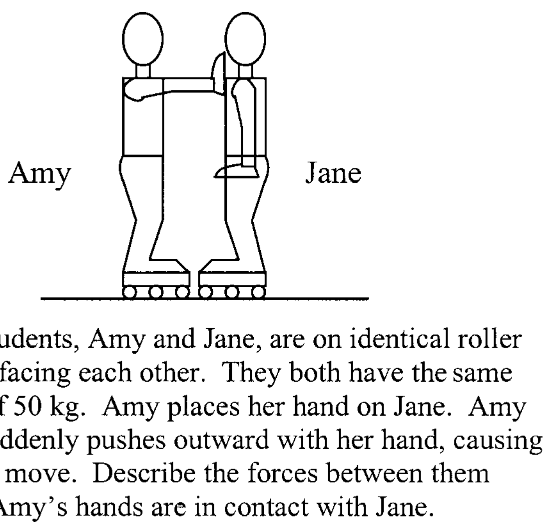


Fig. 3. Open-ended interview question on Newton III with the contextual feature of pushing (the initiator of action).

Table I. Students' incorrect reasoning involving the four contextual features. These are identified in our interviews.

Contextual features and common incorrect reasoning	Student responses in interviews
Velocity—Object with larger velocity exerts a larger force.	*"The car is going faster and it has a greater push against the truck."*
Mass—Object with larger mass exerts a larger force.	*"It has more weight so the momentum behind it is greater."*[a]
Pushing—Object that "pushes" exerts a larger force.	*"Amy actually reaches out and pushes Jane and Jane was just there. Her (Jane's) force was a non-equal but opposite force that she pushes back."*
Acceleration—Object that is speeding up exerts a larger force.	*"Because it is speeding up so it has more acceleration and more momentum behind."*[a]

[a]Most students use momentum as another word for force.

We also found that students can use combinations of different contextual features in their reasoning and may consider them with different levels of significance for specific questions. At this stage, we did not pursue these details further and focused on the study of the independent-first-order relation between student reasoning and the individual context features.[27]

Table II briefly summarizes the incorrect reasoning of the interviewed students corresponding to the four contextual features. As indicated by the results, these four contextual features represent the ones that are frequently used by students in their reasoning; therefore, we define them as the *physical features* related to Newton III. In our definition, the term "physical feature" describes a unique contextual aspect of a physical scenario that is considered relevant to the related physics concept by experts and/or students.[28]

IV. MEASURING THE EFFECTS OF CONTEXTUAL FEATURES ON STUDENTS' REASONING

The physical features can be used as the basis for studying the detailed structure of student models of a particular concept. To do so, we need a probe that can measure the effects of each physical feature on students' reasoning. However, many of the questions in existing instruments are not designed with isolated physical features. For example, the question shown in Fig. 4 mixes two physical features, mass

Table II. Interview results on student reasoning.

Contextual features	Incorrect	Mixed	Correct[a]
Velocity (V)	7	2	0
Mass (M)	7	2	0
Pushing (P)	6	1	2
Acceleration (A)	0	1	8

[a]Students consider the corresponding contextual feature irrelevant.

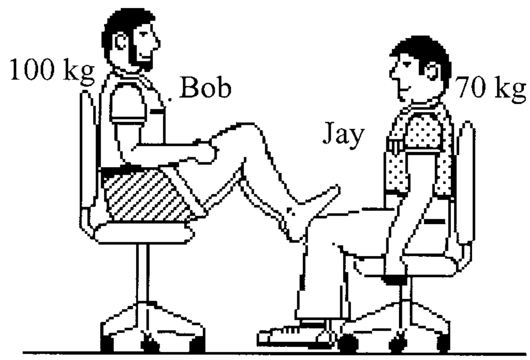


Fig. 4. Question 11 in FCI (original version). This question is on Newton III with mixed contextual features of mass and pushing.

and pushing, together. If a student answers that Bob exerts a larger force, no further evidence can indicate if the incorrect response is generated based on consideration of mass, of pushing, or both.

A. A new multiple-choice instrument on Newton III and model-based assessment

Based on published research and the results from our interviews, we developed a new multiple-choice instrument where each question only measures students' reasoning related to a single physical feature of Newton III. To measure the possible mixing of students' use of their models,[29] for each of the four physical features we designed three questions using different context settings. In Figs. 5 and 6, we show two sample questions: question 7 on velocity and ques-

In a soccer game, two players, John and Tom who happen to have the **same mass**, are running to chase a ball that is flying close to them. John runs about **twice as fast** as Tom. Unfortunately, neither of the players notices the other, and they run into each other. At the time they hit, which of the statements is true?

A. John exerts a greater force on Tom than Tom exerts on John.

B. John exerts the same amount of force on Tom as Tom exerts on John.

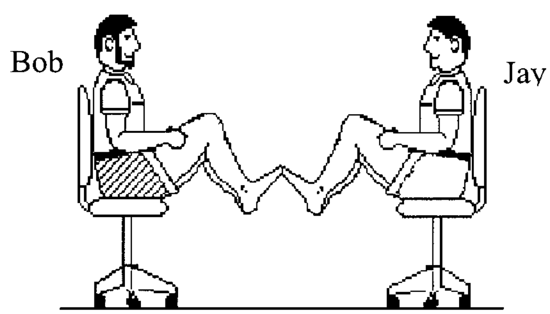
C. Tom exerts a force on John but John doesn't exert a force on Tom.

D. Tom exerts a greater force on John than John exerts on Tom.

E. John exerts a force on Tom but Tom doesn't exert a force on John.

F. None of the above answers describes the situation correctly.

Fig. 5. Question 7 in the new survey on Newton III. This question only involves the contextual feature of velocity.



Two students, Bob and Jay, sit in identical office chairs facing each other. Bob has a mass of 100 kg and Jay has a mass of 70 kg. Both Bob and Jay place their feet against the other. They then both suddenly push outward with their feet at the same time, causing both chairs to move. In this situation, while their feet are still in contact, which of the following choices describes the force?

A. Jay exerts a force on Bob, but Bob doesn't exert a force on Jay.

B. Bob exerts a force on Jay, but Jay doesn't exert a force on Bob.

C. Each student exerts a force on the other, but Jay exerts the larger force.

D. Each student exerts a force on the other, but Bob exerts the larger force.

E. Each student exerts the same amount of force on the other.

F. None of above is appropriate (please write down your own).

Fig. 6. Question 15 in the new survey on Newton III. This question only involves the contextual feature of mass.

Table III. Questions in the new survey on Newton III and the physical features they are measuring.

	Velocity	Mass	Pushing	Acceleration	Others
Questions	1, 5, 7	4, 9, 15	2, 8, 10	3, 13, 14	6, 11, 12, 16

Table IV. The common models corresponding to the four physical features and the associations between the choices of the questions and the common models.

Physics features	Common models	Questions/ choices[a]		
		1	**5**	**7**
Velocity	M_0^V: null model	x	x	x
	M_1^V: correct model	e	b	b
	M_2^V: incorrect model —larger velocity larger force	b	a	a
		4	**9**	**15**
Mass	M_0^M: null model	x	x	x
	M_1^M: correct model	e	b	e
	M_2^M: incorrect model —larger mass larger force	a	d	d
		3	**13**	**14**
Pushing	M_0^P: null model	x	x	x
	M_1^P: correct model	e	e	e
	M_2^P: incorrect model —the one that pushes exerts a larger force	e	e	e
		2	**8**	**10**
Acceleration	M_0^A: null model	x	x	x
	M_1^A: correct model	e	b	c
	M_2^A: incorrect model —the one that speeds up exerts a larger force	a	a	a

[a]"x" is used to represent all other choices.

tion 15 on mass. If a student selects choice (d) of question 15, "*Each student exerts a force on the other, but Bob exerts the larger force*," we then have strong evidence to infer that this student is using an incorrect model based on the physical feature of mass. The complete survey is included in the Appendix. In Table III, we list the questions in clusters based on the physical features that these questions are intended to measure.

This instrument is designed to be used with the model analysis, and each physical feature is treated as an independent dimension to represent students' model structures. On each dimension (corresponding to a specific physical feature), we can further construct a multidimensional model subspace spanned by the common models involved with this particular physical feature. For Newton III, the subspaces for all four physical features have three dimensions (three common models for each physical feature). In general, the dimensions of these subspaces can be different. As an example, with the physical feature of pushing (P), we define the following common models:[30]

M_0^P: Null model (incorrect student ideas that do not involve pushing).

M_1^P: The force has the same magnitude and opposite direction during the interaction regardless of which object is the initiator of the force (correct model).

M_2^P: The object exerting the force will exert a larger force during interaction (incorrect student model).

In Table IV, we list the common models corresponding to the four physical features. The associations between the choices of the questions and the common models are also listed and are used in our later analysis to analyze students' responses. Note that in this example, we have only one incorrect model for each physical feature. In general, we can imagine situations where many incorrect models exist.[31]

B. Validation of the instrument

As a partial validation of the measurement consistency of this instrument, we selected 6 questions from the survey and used them in the 9 interviews that were also used to confirm the existence of the physical features. This approach reduced the time for conducting this research by combining two tasks in a single interview: (1) identifying and confirming the existence of the physical features, and (2) validating our design of the survey instrument. The first task relies on an analysis of open-ended explanations and discussions from the students. In the task of validating the instrument, we first asked students to answer all the multiple-choice questions, and then had them explain the reasoning used to generate their answers.[32] The consistency between students' responses to the questions and their reasoning is used to evaluate if the questions can measure accurately the underlying student reasoning.

In all the interviews, students' explanations and their selections of the answers were found consistent, and we did not find any apparent communication problems in the questions—most students understood the questions and their explanations show consistency between their understanding and the intentions of the measurement. As often observed in interviews, some students may change their minds when explaining their reasoning. We also observed this situation; however, in our cases the reasoning that these students brought up initially (before their second thoughts after some extensive discussion) are all consistent with the answers they selected. Even when students did change their minds, they all came up with answers compatible with the given choices of the questions and their modified explanations were also consistent with their new answers.

Table V. Student background and course information for the five introductory physics courses at Kansas State University.

Courses	Types of courses	Majors	Math prerequisites
Physical World	Algebra, Mech.	Liberal arts	No math
General Physics 1	Algebra, Mech.	Life science	Algebra
General Physics 2	Algebra, E&M	Life science	Algebra
Engineering Physics 1	Calculus, Mech.	Eng. and Phys.	Calculus
Engineering Physics 2	Calculus, E&M	Eng. and Phys.	Calculus

As mentioned previously, students can use mixed ideas in their reasoning. When multiple questions related to a single physical feature are presented to students, we observed that some students responded with different models on different questions. This result indicates that by using these types of questions, we can obtain measurements on students' mixed model states and on the effects of different context features in triggering students' use of models.

V. ASSESSING STUDENTS' MODELS WITH MULTIPLE-CHOICE INSTRUMENT

A. The population

The new multiple-choice survey was used in five introductory physics courses at Kansas State University (Fall, 1999). These courses include: Physical World (PW), a conceptual physics course for nonscience majors with no math prerequisites; General Physics 1 (GP1), the first semester of a two-semester, algebra-based physics course; General Physics 2 (GP2), the second semester of a two-semester, algebra-based physics course; Engineering Physics 1 (EP1), the first semester of a two-semester, calculus-based course for physics and engineering majors; and Engineering Physics 2 (EP2), the second semester of a two-semester, calculus-based course for physics and engineering majors. A brief summary of these courses is listed in Table V. All the courses used traditional instruction.

In the beginning of the five courses, we surveyed a total of 280 students—about 60 students from each course. Students in PW, GP1, and EP1 hadn't had any college instruction on mechanics before they took the courses. Students in GP2 and EP2 all had instruction on mechanics from their previous courses (GP1 and EP1). Therefore, we can approximately study the change of student understanding before and after traditional instruction.

In the following two sections, we apply two numerical methods, *concentration analysis* and *model state estimation*, to analyze the data. In this paper, we provide only limited descriptions of the operations of these tools. More details are provided in Refs. 2 and 9.

B. Concentration analysis

As a way to validate the effectiveness of this multiple-choice instrument, we first used the *concentration analysis* to evaluate the design of the distracters.[33] As we have learned from qualitative research on student learning, student responses to problems in many physical contexts can be considered to be the result of their application of a small number of conceptual models. The way in which the students' responses are distributed on model-based multiple-choice questions can yield information on the students' state: for a particular question; highly concentrated responses implies that many students are applying a common model associated with the question; whereas randomly distributed responses often indicate that the population has less commonality in reasoning. (Sometimes, this situation corresponds to the case where most students have no systematic model and/or the question is not designed to extract information on students' models.)

It is convenient to construct a simple measure that gives information on how the students' responses are distributed among the choices of a particular multiple-choice question. This measure is defined as the concentration factor, **C**, which is a function of students' responses and has values in the interval [0,1]. Larger values represent more concentrated responses with 1 being a perfectly correlated response and 0 a random response. The concentration factor is defined as

$$\mathbf{C}=\frac{\sqrt{m}}{\sqrt{m}-1}\times\left(\frac{\sqrt{\sum_{i=1}^{m}n_i^2}}{N}-\frac{1}{\sqrt{m}}\right), \tag{1}$$

where m represents the number of choices for a particular question, N is the number of students, and n_i is the number of students who select choice i of the question.

Table VI. The average scores (S) and concentration factors (C) of student responses on all 16 questions for the five courses (only the average results of the question groups corresponding to the different physical features are shown).

Class		V	M	P	A	Others
PW	S	0.06	0.10	0.11	0.73	0.35
	C	0.69	0.67	0.34	0.59	0.28
GP1	S	0.12	0.11	0.17	0.85	0.34
	C	0.80	0.76	0.40	0.75	0.29
GP2	S	0.25	0.18	0.32	0.64	0.50
	C	0.57	0.67	0.49	0.50	0.38
EP1	S	0.20	0.23	0.27	0.67	0.35
	C	0.61	0.55	0.37	0.55	0.33
EP2	S	0.35	0.29	0.46	0.65	0.52
	C	0.49	0.54	0.51	0.51	0.40
Average	S	0.20	0.18	0.27	0.71	0.41
	C	0.63	0.64	0.42	0.58	0.34
Standard	ΔS	0.11	0.08	0.14	0.10	0.11
Deviation	ΔC	0.13	0.13	0.12	0.13	0.09

As seen in Eq. (1), the concentration factor has a complicated relation with n_i. Based on our results, a value of **C** greater than 0.5 represents a fairly high concentration (>60% students selected the same choice). A value between 0.2 and 0.5 is considered to be a medium concentration, in which case students' responses are often concentrated on two choices indicating a possible two-model situation. A value less than 0.2 indicates that the students' responses are somewhat evenly distributed among three or more choices. In this case, students either have no consistent reasoning at all and respond rather randomly, or they may exist in an evenly distributed population for all the possible models involved in the question. (Further clarification on this situation requires looking at the content of the question and student behavior in interviews.)

In Table VI we show the results of the concentration analysis of student responses on the newly developed multiple-choice test (see Appendix) for the five courses. For easy comparison, we first calculate the concentration factors for all 16 questions and then group the questions based on the different physical features to obtain the average results of the groups.

The concentration analysis provides a convenient tool for determining whether a question can select common incorrect student models or if students actually have a common model at all. Therefore, it can be used to assess student learning as well as to facilitate the development of multiple-choice instruments. For a concept involving two common models, as in the case of Newton III, a well-designed question often has medium to high concentration on students' pretest data. As shown by Table VI, the questions corresponding to the four physical features all have high concentration factors, whereas the questions we used to explore certain interesting possibilities, denoted by "others," have systematically lower concentration factors. This difference indicates that for the four physical features, most students have common types of reasoning (models) similar to the ones that we have identified from qualitative research. It also shows that the choices of the questions match well with these models.

A more detailed look at the data shows that the students' responses to questions with the physical features of velocity and mass have high concentration values but low scores. This indicates that most students selected the same incorrect answers on these questions (common incorrect models). In contrast, for questions that feature acceleration, students' responses show high scores and high concentrations, indicating that most students selected the correct answer. On questions that feature pushing, students' responses have a medium value for C, which indicates that students often select between two popular answers. In this case, students usually have a mixed state of understanding. To look into the details of all these possible situations of student models, we need to use our knowledge of the content of the questions and apply the methods of model analysis to extract the probability states of students' use of different models as discussed in the following.

C. Model state estimation

Using each physical feature as an independent dimension, we analyzed student model states. As discussed in Refs. 2 and 9, the model state for a single student gives the amplitude of the distributed probabilities for the student to use (due to the context settings of the questions used in the measurement) the different common models associated with the set of questions. The model state for a population gives the amplitude of the distributed probabilities for the population to use the different common models. These distributed probabilities are stored in a model state vector, and its structure can provide important information on the ways that students apply their conceptual models. In particular, it provides a numerical measure of how a single student or a population may inconsistently use different conceptual models in contexts that are regarded equivalently by experts. Such mixed states are often a crucial intermediate stage of a favorable conceptual change.[34]

To calculate the model states, we first code students' raw responses to obtain single-student model vectors by using the scheme shown in Table IV. Then a class model density matrix is obtained for each class using the single student model vectors. We then calculate the eigenvectors and eigenvalues of the class model density matrix. Figure 7 shows a schematic of the procedures for the calculation.[35]

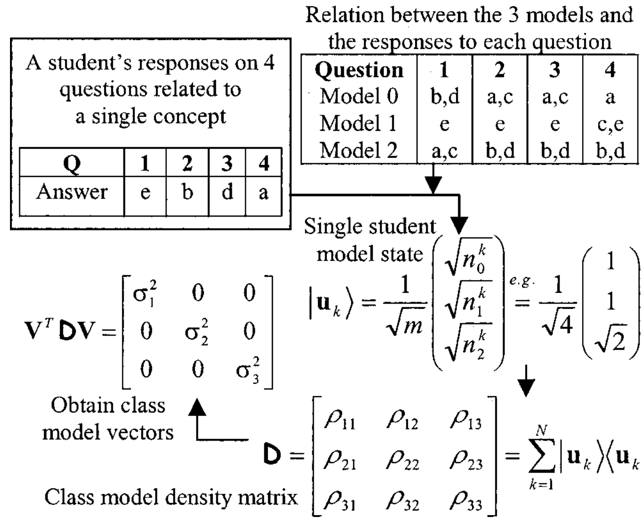


Fig. 7. Schematics of the procedures for calculating the class model states.

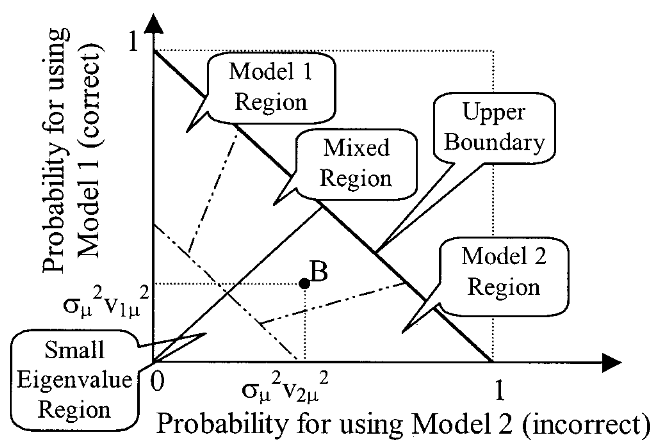


Fig. 8. Model plot used to represent the class model states. Model 1 (Model 2) region represents comparatively consistent model states with dominant model 1 (model 2) components. Mixed model region represents mixed model states. Secondary region represents model states with small eigenvalues. In the figure, σ_μ^2 is the μth eigenvalue of the class model density matrix and $v_{\eta\mu}$ is the ηth vector component of the eigenvector corresponding to σ_μ^2.

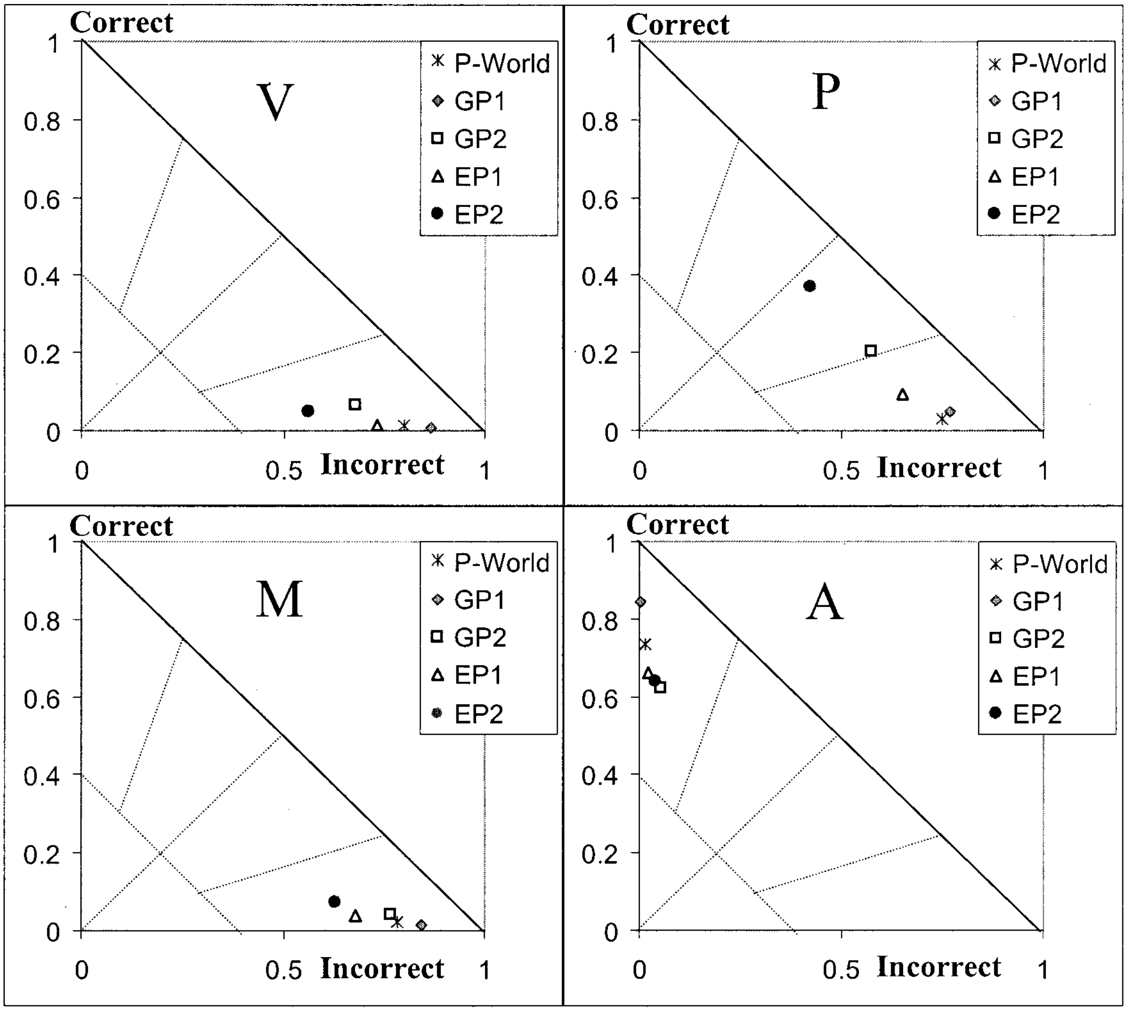


Fig. 9. Model plots of student class model states on Newton III with the four physical features: velocity (V), mass (M), pushing (P), and acceleration (A). The data is taken from five introductory physics courses at Kansas State University.

As discussed in Ref. 9, the class model state vectors are the eigenvectors of a class model density matrix and reflect the set of unique features of the model states held by the individual students in the class, whereas the eigenvalues reflect the popularity of the corresponding class model states. A useful way to investigate the shift in student thinking between two common models is to create a model plot.[36] As shown in Fig. 8, a particular model state as well as its eigenvalue can be represented by a point on a model plot (for example, B), where the horizontal and vertical components are equal to the products of the square of the class model state vector's two corresponding components and the model state vectors' eigenvalues. The calculated results give the probabilities for the class to apply the common models represented by each of the corresponding axes.

From a model plot, we can obtain information about the class population and individual students' use of their models. In general, the value of the largest eigenvalue can provide a measure of the consistency of the population. For example, a class model state with a large eigenvalue, which results in a point close to the upper boundary line, indicates that a large number of students in the class have model states similar to this class model state, that is, the class has a somewhat consistent population.

The information on the individual student's use of their models is reflected by the eigenvectors (or class model state vectors). If most students in a class use their models consistently, which corresponds to a large number of "pure" single student model states, the point representing the class model state will be in either model 1 (correct) or model 2 (incorrect) regions. When individual students use their models inconsistently, which corresponds to a large number of "mixed" single student model states, the point representing the class model state will be in the mixed model region.[37]

The student class model states with the four physical features of Newton III are calculated and plotted in Fig. 9. For each class, only the model state with the largest eigenvalue (called the primary model state) is shown.

From Fig. 9, we can see that for mass and velocity, the primary model states of all the classes are in the region representing a consistent incorrect model (model 2), which indicates that most students have a dominant consistent incorrect model. The popularity of the incorrect model decreases somewhat in higher-level courses—from 90% (GP1) to 60% (EP2), but the model states stay in the model 2 region, showing that most students in the five classes apply their models consistently, that is, no mixed use of different models. In this situation, the eigenvalue of the primary model state provides an estimate of the fraction of the class population using the incorrect model.

Student model states for acceleration appear to be the opposite of the situations with mass and velocity. In this case, most students hold a consistent "correct" model when they consider acceleration irrelevant. Although students give correct responses on the related questions, it does not mean that student models are the same as the expert one. Correct understanding of the underlying student reasoning requires further studies with detailed interviews. As a preliminary indication, the analysis of our interviews suggests that a possible reason for the students to consider acceleration irrelevant is not that they truly understand the nature of Newton III, but rather that they believe the velocity to be the major factor and acceleration is related to velocity and does not have a direct effect.[38]

For pushing, the student model states show a different structure. The low level classes still have a dominant consistent incorrect model. As the class level gets more advanced, the student model states become more mixed. The most advanced class (EP2) had almost a perfectly mixed model state with a large eigenvalue (~ 0.8), which also indicates that most students in this class have mixed model states and the structures of the individual single-student model states are also similar (students behave similarly). This is very different from the situations with the other physical features and implies a different process in conceptual development. Notice that among the five classes, three (PW, GP1, and EP1) haven't had any college-level instruction on mechanics, and two (GP2 and EP2) had instruction on mechanics in the previous semester. As clearly shown in the model plot, the three classes without instruction are in the model 2 region representing a consistent incorrect model. The two classes with instruction are in the mixed region. The shifts of the class model states are consistent with our presumption that students often start with a consistent incorrect model and go through a stage of mixed models toward building a consistent correct model.[39]

An advantage of this model analysis representation over score-based representation is that it shows more details of student conceptual development status. For example, if one uses scores to analyze the same data, the results (for example, with the EP2 class as shown in Table VI) would show scores ranging from 29% to 46% for the three question groups on mass, velocity, and pushing. As shown by the model analysis method for velocity and mass, the class population exists in two groups that each uses consistently either the correct model (about 30% of the total population) or the incorrect model (about 70% of the total population). On the other hand, for pushing, the class population predominantly holds a mixed model state where each student uses both the correct and the incorrect models in an inconsistent manner. However, if we collapse these details into scores, the two situations, (1) class population exists in two groups that each uses consistently one of the two models, and (2) class population exists in a single group that uses both models inconsistently, would produce numbers showing the percentage of correct responses, but with no further information on how such responses were produced. Therefore, using a score-based measurement would not allow us to distinguish between the different patterns of conceptual development indicated by the model-based analysis.

D. Implications on conceptual development

As recognized by many researchers, the stage of a mixed model state is often an important intermediate step for a complete favorable conceptual change. Therefore, we put more emphasis on the study of student reasoning concerning pushing. When we ask the students in our interviews to explain their reasoning related to the physical features of pushing, many (7/9) of them specifically said that "when you are pushing something, you get pushed back." A significant number (4/9) of students even explicitly said that "the force is equal and opposite" and tried to use this idea in their reasoning. Some of the students can associate these correct ideas with examples such as push against a wall from their experiences. In the following, we summarize some common behaviors of the students identified in our interviews.

(1) Students often use the two above quotes in their explanations on questions involving pushing. With questions that do not explicitly involve the issue of pushing, students instantly look for mass or velocity in their reasoning without even bothering to recall the two sentences, which many of them can memorize (especially the first one) and relate to examples from their personal experience. It appears that for the students in our interview, the two sentences are strongly associated with the issue of pushing only.

(2) When students use the two sentences, the first one is very easy for them. On the other hand, many students still have problems with the second sentence and have the tendency to think the one who pushes exerts a larger force. So students can sometimes give contradictory answers on similar questions with pushing resulting in a mixed model state.

With the results from the qualitative and quantitative methods, we can infer a possible explanation for why student model states are different with the physical feature of pushing. It appears that pushing is often the first and the most common issue in examples used to introduce Newton III. More importantly, most students have the experience of being pushed back when they are pushing an object. Integrating the sensory cues of being pushed back in instruction can directly link this particular aspect of the concept of Newton III to students' real life experience and presumably make it more meaningful for them and thus easier to understand. Therefore, students can make significant changes in their models of this physical feature even with traditional instruction. On the other hand, students' strong naive models associated with mass and velocity often receive inadequate or ineffective treatment by traditional instruction and changes of these models are rather insignificant.

VI. IMPLICATIONS FOR INSTRUCTION

This study provides evidence for the context dependence of conceptual learning. The result implies that effective instruction often requires that the instructional contexts be integrated with the students' existing knowledge system. As we can see from this example, when the context used to present the new concept is treated properly, even traditional instruction can make a significant impact on students' conceptual understanding. Therefore, instruction should be developed based on a good understanding of the possible forms of student models as well as the effects of the related contextual features. Successful instruction should also include effective assessment tools to provide accurate and context-rich information on students' state of understanding. This research suggests that the model analysis method can be a useful assessment tool in research and instruction. It has several advantages:

(1) It uses multiple-choice instruments making it appropriate and feasible to implement in large classes.
(2) The instruments and analysis methods are based on systematic research of student conceptual models and thus can provide detailed and validated information on the state of student understanding.
(3) The method of using physical features to study the structures of student models can yield explicit information for both researchers and instructors on the details of how contexts and students' conceptual models interact during the process of conceptual development for a single student and/or a population.

VII. SUMMARY

We found that student models show different structures for different physical features and that student model evolution during instruction also shows different processes with different physical features. Such information is often unavailable using assessment instruments designed with entangled physical features. As an example, the new instrument and algorithms in model analysis were found to be effective in measuring and analyzing the details of student model structures. With this new method, we can obtain detailed quantitative information on the status of student conceptual understanding and the changes of such understanding with respect to specific contextual features.

ACKNOWLEDGMENTS

We would like to thank Professor Edward F. Redish at the University of Maryland for his many constructive suggestions and Professor Alan Van Heuvelen at the Ohio State University for his insightful discussions. We want to express our sincere appreciation to the anonymous reviewers for their extremely valuable feedback. We also thank the NSF for their support on a number of grants including DUE-9652877, REC-0087788, and REC-0126070.

APPENDIX: SURVEY ON STUDENT MODELS OF NEWTON'S THIRD LAW

Instructions to students:

- *Please read each question carefully. If you have any questions, ask the instructor for help.*
- *For each question, please only select* **one** *choice that best describes your understanding.*
- *This test is for diagnostic purposes only and will not influence your grade in any way. Therefore, we would like you to give your true thoughts on the physics involved. This will greatly help us to design better instruction and help you to improve your performance on exams.*
- *Your name and ID on this survey are for administration purposes, and your result will be kept strictly confidential.*

Use the following choices to answer questions **1–4**.

A. The truck exerts a greater force on the car than the car exerts on the truck.

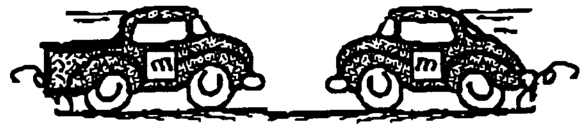
B. The car exerts a greater force on the truck than the truck exerts on the car.

C. The truck exerts a force on the car but the car doesn't exert a force on the truck.

D. The car exerts a force on the truck but the truck doesn't exert a force on the car.

E. The truck exerts the same amount of force on the car as the car exerts on the truck.

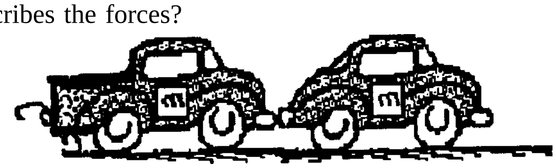
F. None of the answers above describes the situation correctly.



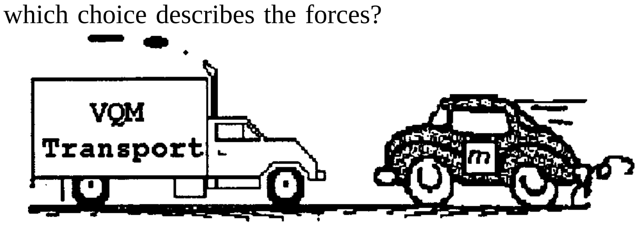
1. As shown in the figure on the right, a collision happens between a small pickup truck and a car. The small truck has the same mass as the car does. At the time of collision, both vehicles travel at a constant speed but the small truck is moving at a slower speed than the car. When they collide, which choice describes the forces?

2. Now consider the case that before the collision, the car is traveling at a constant speed while the truck starts slow and is speeding up. At the moment of collision, both vehicles happen to be traveling at the same speed. When they collide, which of the above choices describes the forces?

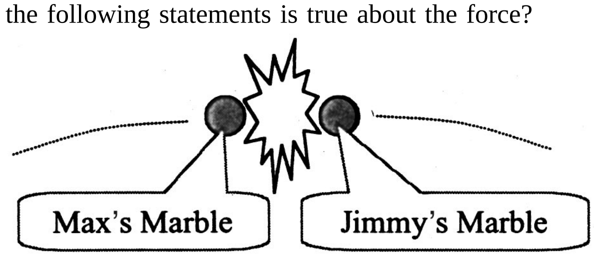
3. As shown in the figure below, a car breaks down on the road and is pushed into town by a small AAA service truck. The small truck has the same mass as the car does. While the truck, still pushing the car, is speeding up, which choice describes the forces?



4. As shown in the figure below, a collision happens between a big truck and a car. The big truck has a much larger mass than the car does. Before the collision, both vehicles are traveling at the same constant speed. When they collide, which choice describes the forces?



5. Two boys, Jimmy and Max, are tossing identical marbles at each other. Max is a little stronger than Jimmy, so the marble from him goes faster than the marble from Jimmy. During one nice shot, the two marbles, one from Max and one from Jimmy, collide in mid air (see figure below). Which of the following statements is true about the force?



A. Max's marble exerts a greater force on Jimmy's marble than Jimmy's marble exerts on Max's marble.

B. Max's marble exerts the same amount of force on Jimmy's marble as Jimmy's marble exerts on Max's marble.

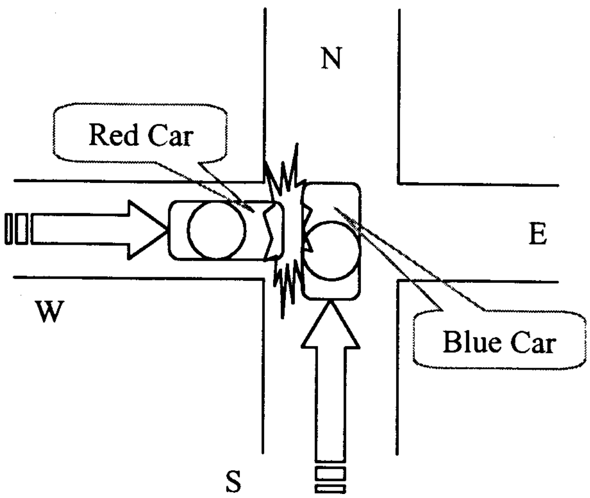
C. Max's marble exerts a smaller force on Jimmy's marble than Jimmy's marble exerts on Max's marble.

D. Jimmy's marble exerts a force on Max's marble but Max's marble doesn't exert a force on Jimmy's marble.

E. Max’s marble exerts a force on Jimmy’s marble but Jimmy’s marble doesn’t exert a force on Max’s marble.

F. None of the answers above describes the situation correctly.

6. A blue car and a red car both enter an intersection where the traffic light is just broken. Both cars are the same model except for the color. As shown in the figure below, the blue car was originally traveling from south to north and the red car from west to east. At the moment when the red car hit the blue car on the left side, both cars are traveling at the same constant speed. Which of the following statements is true?



A. Red car exerts a greater force on blue car than blue car exerts on red car.

B. Red car exerts the same amount of force on blue car as blue car exerts on red car.

C. Blue car exerts a greater force on red car than red car exerts on blue car.

D. Blue car exerts a force on red car but red car doesn’t exert a force on blue car.

E. Red car exerts a force on blue car but blue car doesn’t exert a force on red car.

F. None of the answers above describes the situation correctly.

Use the following choices to answer questions 7 and 8.

A. John exerts a greater force on Tom than Tom exerts on John.

B. John exerts the same amount of force on Tom as Tom exerts on John.

C. Tom exerts a force on John but John doesn’t exert a force on Tom.

D. Tom exerts a greater force on John than John exerts on Tom.

E. John exerts a force on Tom but Tom doesn’t exert a force on John.

F. None of the answers above describes the situation correctly.

7. In a soccer game, two players, John and Tom who happen to have same mass, are running to chase a ball that is flying close to them. John runs about twice as fast as Tom. Unfortunately, neither player notices the other, and they run into each other. At the time they hit, which of the above statements is true?

8. Still in that soccer game, John and Tom run into each other again. This time, Tom starts early and is running at a constant speed. John starts late and is speeding up. At the time they hit, they both happen to be running at the same speed. Which of the above statements is true?

9. Later in that soccer game John runs into another player, Bill, who is almost twice as heavy as John. This time, they are both running at the same speed. At the time they hit each other, which of the following statements is true?

A. John exerts a greater force on Bill than Bill exerts on John.

B. John exerts the same amount of force on Bill as Bill exerts on John.

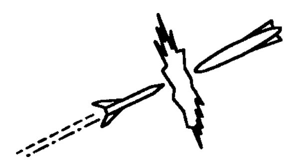
C. Bill exerts a force on John but John doesn’t exert a force on Bill.

D. Bill exerts a greater force on John than John exerts on Bill.

E. John exerts a force on Bill but Bill doesn’t exert a force on John.

F. None of the answers above describes the situation correctly.

10. As shown in the figure below, an enemy aircraft drops a bomb, which is gliding down at a constant speed toward people on the ground. A missile is launched and is accelerating toward the bomb. The missile and the bomb happen to have the same mass. At the moment the missile hits the bomb, they both are moving at the same speed and neither of them explodes after the hit. Which choice describes the forces?



A. The missile exerts a greater force on the bomb than the bomb exerts on the missile.

B. The bomb exerts a greater force on the missile than the missile exerts on the bomb.

C. The missile exerts the same amount of force on the bomb as the bomb exerts on the missile.

D. The missile exerts a force on the bomb but the bomb doesn’t exert a force on the missile.

E. The bomb exerts a force on the missile but the missile doesn’t exert a force on the bomb.

F. None of the answers above describes the situation correctly.

Use the following choices to answer questions 11 and 12.

A. The floor exerts a greater force on the feet than the feet exert on the floor.

B. The feet exert a greater force on the floor than the floor exerts on the feet.

C. The feet exert a force on the floor but the floor does not exert a force on the feet.

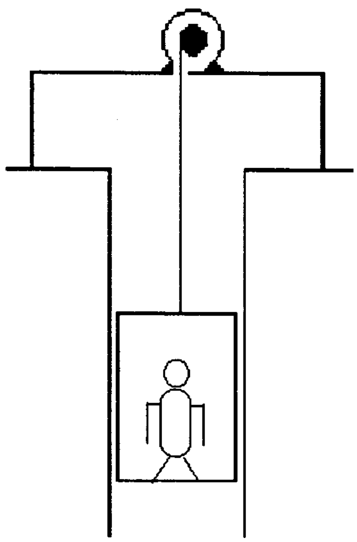
D. The floor exerts a force on the feet but the feet do not exert a force on the floor.

E. The floor exerts the same amount of force on the feet as the feet exert on the floor.

F. None of the answers above describes the situation correctly.

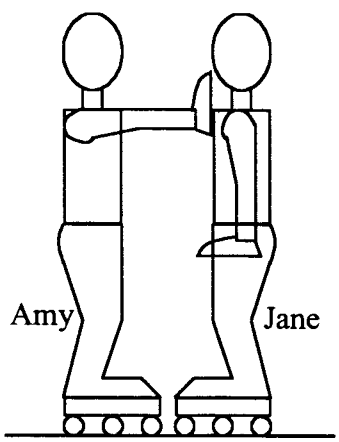
11. Anna is taking an elevator from the fourth floor of the physics building to catch the next class in the basement. In

the elevator, Anna is standing right in the middle without touching anything else when the elevator just started to go down. Which of the above choices correctly describes the forces between the floor of the elevator and Anna's feet?



12. After class, Anna is taking the same elevator from the basement back to the fourth floor to ask for help from her physics professor. In the elevator, Anna is standing right in the middle without touching anything else when the elevator just started to go up. Which of the above choices correctly describes the forces between the floor of the elevator and Anna's feet?

13. Two students, Amy and Jane, are on very good identical roller blades facing each other. They both have the same mass of 50 kg. Amy places her hand on Jane, as shown to the right. Amy then suddenly pushes outward with her hand, causing both to move. In this situation, while Amy's hands are in contact with Jane, which choice describes the forces?



A. Jane exerts a force on Amy, but Amy doesn't exert any force on Jane.

B. Amy exerts a force on Jane, but Jane doesn't exert any force on Amy.

C. Each student exerts a force on the other, but Jane exerts the larger force.

D. Each student exerts a force on the other, but Amy exerts the larger force.

E. Each student exerts the same amount of force on the other.

F. None of these answers is correct.

14. Susan, a little girl, and her Mom are traveling to Europe. They have a lot of luggage. At the airport, Susan tries to help her Mom by handling her luggage all by herself. She has a big suitcase that weighs the same as she. So she can only push the case to move it a little bit at a time. But she keeps doing it. Each time while she is pushing the case, which choice describes the forces?

A. The case exerts a force on Susan, but Susan doesn't exert any force on the case.

B. Susan exerts a force on the case, but the case doesn't exert any force on Susan.

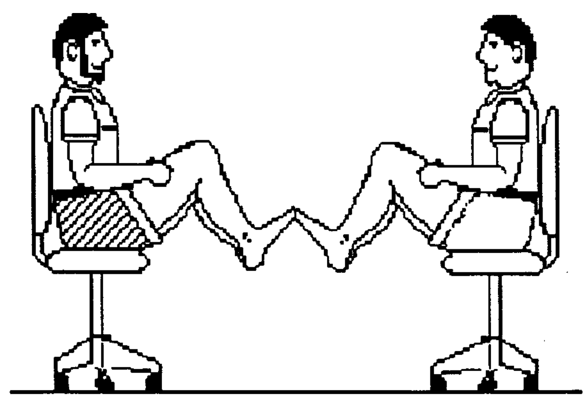
C. Both Susan and the case exert a force on the other, but the box exerts a larger force on Susan.

D. Both Susan and the case exert a force on the other, but Susan exerts a larger force on the box.

E. Both Susan and the case exert the same amount of force on the other.

F. None of these answers is correct.

15. Two students, Bob and Jay, sit in identical office chairs facing each other. Bob has a mass of 100 kg and Jay has a mass of 70 kg. Both Bob and Jay place their feet against the other, as shown to the right. They then both suddenly push outward with their feet at the same time, causing both chairs to move. In this situation, while their feet are still in contact, which of the following choices describes the force?



A. Jay exerts a force on Bob, but Bob doesn't exert a force on Jay.

B. Bob exerts a force on Jay, but Jay doesn't exert a force on Bob.

C. Each student exerts a force on the other, but Jay exerts the larger force.

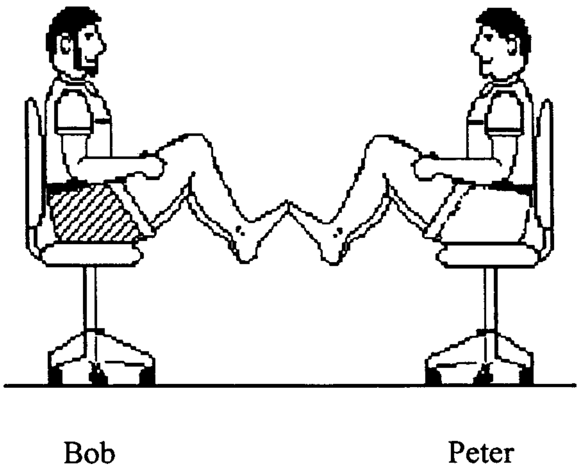
D. Each student exerts a force on the other, but Bob exerts the larger force.

E. Each student exerts the same amount of force on the other.

F. None of these answers is correct.

16. Now Bob and Peter sit in identical office chairs facing each other. They all have the same mass of 100 kg. Peter is a football player and is much stronger than Bob. Again, they then both suddenly push outward with their feet, causing both chairs to move. In this situation, while their feet are still

in contact, which of the following choices describes the force?



A. Peter exerts a force on Bob, but Bob doesn't exert a force on Peter.

B. Bob exerts a force on Peter, but Peter doesn't exert a force on Bob.

C. Each student exerts a force on the other, but Peter exerts the larger force.

D. Each student exerts a force on the other, but Bob exerts the larger force.

E. Each student exerts the same amount of force on the other.

F. None of these answers is correct.

[a)]Electronic mail: lbao@pacific.mps.ohio-state.edu

[1]L. C. McDermott and E. F. Redish, "Resource letter PER-1: Physics education research," Am. J. Phys. **67** (9), 755–767 (1999).

[2]L. Bao, "Dynamics of student modeling: A theory, algorithms, and application to quantum mechanics," Ph.D. dissertation, University of Maryland, December 1999.

[3]G. Gliner, "College students organization of math word problems in relation to success in problem solving," School Sci. Math. **89** (5), 392–404 (1989).

[4]G. Gliner, "College students organization of math word problem solving in terms of mathematical structure versus surface structure," School Sci. Math. **91** (3), 105–110 (1991).

[5]*How People Learn: Brain, Mind, Experience, and School*, edited by J. D. Bransford, A. L. Brown, and R. Cocking (National Academy Press, Washington, DC, 1999).

[6]D. Palmer, "The effect of context on students reasoning about forces," Int. J. Sci. Educ. **19** (6), 681–696 (1997).

[7]E. Engel Clough and R. Driver, "A study of consistency in the use of students conceptual frameworks across different task contexts," Sci. Educ. **70** (4), 473–496 (1986).

[8]In research, we generally consider three major domains of context factors: content-based factors, educational settings, and certain student internal status such as views and motivations on learning.

[9]L. Bao and E. F. Redish, "Concentration analysis: A quantitative assessment of student states," PERS of Am. J. Phys. **69** (7), S45–53 (2001); "Model analysis: Assessing the dynamics of student learning," Cognit. Instruction (submitted).

[10]For example: D. Hestenes and M. Wells, "Mechanics baseline test," Phys. Teach. **30**, 159–169 (1992); D. Hestenes, M. Wells, and G. Swackhammer, "Force concept inventory," *ibid.* **30**, 141–153 (1992).

[11]Reference 2, Chap. 4.

[12]See Refs. 2 and 9.

[13]J. Minstrell, "Facets of students knowledge and relevant instruction," *Research in Physics Learning: Theoretical Issues and Empirical Studies, Proceedings of an International Workshop, Bremen, Germany, 4–8 March, 1991*, edited by R. Duit, F. Goldberg, and H. Niedderer (IPN, Kiel, Germany, 1992), pp. 110–128.

[14]S. Vosniadou, "Capturing and modeling the process of conceptual change," Learning Instruction **4**, 45–69 (1994).

[15]Ronald K. Thornton, "Conceptual dynamics: Changing student views of force and motion," in *The Changing Role of Physics Departments in Modern Universities, Proceedings of the International Conference on Undergraduate Physics Education*, edited by E. F. Redish and J. S. Rigden (Wiley, New York, 1997), pp. 241–266.

[16]Detailed definitions and discussions can be found in Refs. 2 and 9.

[17]Supporting evidence and method of investigations can be found in Refs. 2 and 9.

[18]A detailed formulation is discussed in Refs. 2 and 9.

[19]See Ref. 2 and I. A. Halloun and D. Hestenes, "Common sense concepts about motion," Am. J. Phys. **53** (11), 1056–1065, 1985; "The initial knowledge state of college physics students," Am. J. Phys. **53** (11), 1043–1055 (1985).

[20]Reference 7 and D. P. Maloney, "Rule-governed approaches to physics—Newton s third law," Phys. Educ. **19**, 37 (1984).

[21]Reference 7 and D. Palmer, "How consistently do students use their alternative conceptions?," Res. Sci. Educ. **23**, 228–235 (1993).

[22]See Ref. 6.

[23]See Refs. 2, 7, and 9. In addition, our experience also suggests that it is sometimes possible for students to have different considerations for an object that is speeding up and one that is moving at a constant speed.

[24]D. A. Zollman, "Preparing future science teachers: The physics component of a new program," J. Phys. Educ. **29**, 271–275 (1994); "Learning cycles for a large enrollment class," Phys. Teach. **28**, 20–25 (1990).

[25]When dealing with the mixing of student models, we consider two different types of mixed states: explicit mixing and implicit mixing. Explicit mixing describes the situation where a single context setting activates a student to explicitly consider two different models at the same time. Implicit mixing describes the situation where a single context setting always activates a student to use one model, but different equivalent context settings will cue the same student to use different models. More details on how to study and assess the different types of mixing are discussed in Ref. 9 and in a manuscript in preparation.

[26]Here, although students consider the acceleration irrelevant, it doesn't mean that these students have the correct expert model on this issue. It only reflects that most students don't associate this issue in their reasoning.

[27]We are currently studying this issue.

[28]See Ref. 2 for more details on physical features.

[29]See Ref. 2.

[30]The numbering tradition of the common models is different from the one used in Ref. 2. The system used in this paper makes it more convenient to represent model space when the number of common models with different physical features may vary. The superscript P is used to represent the dimension related to pushing.

[31]See Ref. 2, Chap. 5.

[32]Strictly speaking, the development of the survey and the interviews were done in the same time frame. The interviews were conducted over a period of two weeks and the results from the earlier ones were used to make slight modifications to the survey questions. Our initial design of the survey was largely based on the existing research in the literature as well as our own empirical experience. However, based on the interview results, this initial guess seems to have been quite successful, which saved much time in our research.

[33]See Ref. 9.

[34]See Refs. 9, 13, 14, Ref. 2, Chap. 5, and D. P. Maloney and R. S. Siegler, "Conceptual competition in physics learning," Int. J. Sci. Educ. **15** (3), 283–295 (1993).

[35]See Ref. 2 for more details.

[36]For more details, see Ref. 9 and Ref. 2, Chap. 4.

[37]See Ref. 2 for details on model plots and class model states.

[38]Quite interestingly, the results also indicate that more students in higher level classes change their ideas on this issue. Our preliminary results indicate that this can happen when students get into confusing/transitional stages (often with mixed model states). We are looking further into the mechanisms behind this phenomenon.

[39]See Refs. 2 and 14.