Model Analysis of Fine Structures of Student Models: An Example with Newton's Third Law

Lei Bao
Department of Physics, The Ohio State University, 174 W 18th Ave., Columbus, OH 43210
Email: lbao@pacific.mps.ohio-state.edu

Dean Zollman, Kirsten Hogg
Department of Physics, Kansas State University, Manhattan, KS 66506

Edward F. Redish
Department of Physics, University of Maryland, College Park, MD 20742

(Submitted to the Physics Education Supplement of American Journal of Physics)

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

I. Introduction

Over the past two decades, a significant amount of research has been conducted to investigate students' use of incorrect reasoning in solving physics problems.[1] These alternative ideas (often incorrect) were described as preconceptions, misconceptions, alternative concepts, etc. Studies have shown that, at both the college and pre-college levels, the formation and application of these conceptual knowledge are strongly context dependent.[2, 3, 4, 5] As indicated by research, in problem-solving situations, the contextual features of the problems can have significant influence on students' reasoning.[6,7] Many physics concepts, such as the Newton's Third Law (Newton III), can involve a variety of different contextual features. These features, when not treated properly, can often increase the difficulty in probing student learning and delivering effective instruction. Therefore, the details of how different contextual features can affect students' reasoning are of great importance to researchers and instructors in understanding students' conceptual development.

Based on the context dependence of learning process, we developed Model Analysis where the different types of students' reasoning are described with *student models*.[8] In Model Analysis, both the models which students use and the contexts in which they use the models are part 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 the Model Analysis 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. Further, one can begin to understand which contexts are difficult for the student to apply the model. Thus, the researcher can start to build a picture of the students' *model states* with respect to the concept.

In Model Analysis, we have developed numerical methods to extract the information of students' use of models with model-based multiple-choice instruments. These also provide convenient tools to study and assess the involvement of different contextual features in students' models with properly designed instruments. Popular instruments for probing students' broad understanding of concept such as Newton II exist.[9] However, for complicated concepts such as Newton III, existing instruments were often designed with multiple contextual features entangled in a single question. Thus, probing the involvement of a particular contextual feature in the formation and application of students' models is difficult.[10] In this research, we aim to develop a new type of instrument where a single question only measures the effects of one contextual feature involved with a particular concept.

In section 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 section IV, we discuss the measurement and introduce the research on the development of a new instrument. Section V gives an application example using this new instrument with detailed data analysis. In section VI, we discuss the implications of this research and conclude with a summary.

II. Student Models of Newton III and the Involvement of Contextual Features

Before going into more details of the analysis, we first briefly review what we mean as a model.[11] A model is a functional mental construct that is associated with a specific concept (or a topic) and can be applied directly in context instances relevant to the concept to obtain explanatory results. Models have direct causal relations with the responses students generate in various problem-solving situations. Other researchers have also studied this issue where they defined *facet*[12], *mental model*[13], *student views*[14], etc. A comparison with the literature reveals that facets, mental models, and student views represent mental constructs that are fundamentally similar to what we call models. However, in our definition, models have explicit attributes with respect to contexts and are considered to have direct casual relations with the responses produced by the students. In rational 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*.[15]

With a particular physics concept, through systematic research, we can identify a finite set of commonly recognized models.[16] These models usually consist one correct expert model and a few incorrect or partially correct student models. We define these as the *Physical Models* of this particular concept – these models are common to a group of students with similar background and the existence of these models can be verified through research. The identification of physics models includes all the possible forms of students' models: the ones that students bring with them to instruction, the ones that students are likely to create on the spot when confronting new contexts relevant to the concept, and the ones that students can develop during instruction.

For a single student to solve a set of questions related to a particular physics concept, there are usually two different situations in student's use of models: 1. the student can use one of the physical models and be consistent in using it in solving all questions; or 2. the student can hold different physical models at the same time and be inconsistent in using them, i.e., the student can use one of the physical models on some questions and use another model on other questions, even though all the questions are related to a single concept and the questions are seen as equivalent by experts. The different situations of the students' use of models are described with student *Model States*.[17] 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 calculated mathematically with a multidimensional probability vector to represent the probabilities for a single student to apply different physical 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 over decades. Based on the existing knowledge, we can obtain a rather clear picture for the possible forms of the student models used by students with most topics in mechanics. For example, with the concept of the relation between force and motion, we were able to identify three physical models based on the results from our research and the literature.[18] 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*,[19] where students think that during an interaction, the dominating party exerts a larger force. The dominance can come from several sources: (a) the one with greater mass, (b) the one that actively exerts a force (in contrast to a reaction force). Apparently, these two popular issues are often embedded in the contextual settings of physics questions on Newton III. Students' responses obtained with these questions reflect students' understandings on the related concept, which are in part built on the ways that students consider how the different contextual features are involved.

Although studies on 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' using their conceptual models in different contexts.[20] In a recent study on student understanding of forces,[21] 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 the student reasoning.

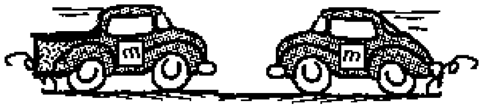
With Model Analysis and the framework we develop in this paper, the context is considered as a significant part of the student reasoning itself, and we use the contextual features as the basis to study the student conceptual understandings.

III. Contextual Features in Newton III

As indicated from existing research, we can easily identify two contextual features that are frequently used by students in their reasoning about Newton III: (a) mass, (b) source of the force (who is pushing). Since students often associate force with velocity,[22] our empirical experience suggests two additional possibilities: (c) velocity, and (d) acceleration.

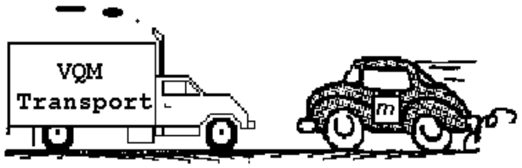
To validate this speculation, we interviewed 9 students from an introductory physics class, Concepts of Physics, at Kansas State University[23]. The class has no math requirement and is for students majoring in elementary school education. The interviews were conducted in the middle of the course, about 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 from this class.

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



As shown in the figure, a collision happens between a small pickup truck and a car. The small truck has the same weight 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.

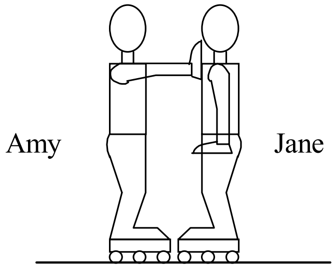
Figure 1. Open-ended interview question on Newton III with the contextual feature of velocity.



As shown in the figure, a collision happens between a big truck and a car. The big truck has a much larger mass (weight) than the car does. Before the collision, both vehicles are traveling at a same constant speed. Describe the force at the moment when they collide.

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

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 objects – they determine that the object with dominating features applies a greater force. The dominating features are selected from three contextual features: velocity (V), mass (M), and source of the force (P). Table 1 shows some typical incorrect student reasonings identified in the interviews.



Two students, Amy and Jane, are on identical roller blades facing each other. They both have a same mass of 50 kg. Amy places her hand on Jane. Amy then suddenly pushes outward with her hand, causing both to move. Describe the forces between them while Amy's hands are in contact with Jane.

Figure 3. Open-ended interview question on Newton III with the contextual feature of pushing (the source of force).

Some students (2/9) are also found to be in a confused state and use mixed ideas (between the correct model and the *dominance* incorrect viewpoint) on questions related to the three contextual features. For example, on the question shown in figure 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.*"

Using the similar setting shown in Figure 1, we ask 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 said 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.[24] Therefore, we can conclude that most students we interviewed do not consider acceleration as a factor with significant effects in their

Table 1. Students' incorrect reasoning involving the four contextual features. These are identified in our interviews.

Contextual Features – Common Incorrect Model	**Student Responses**
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.*"[*]
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 an 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.*"[*]

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

reasoning on questions related to Newton III. In our later analysis, we still kept this contextual feature to see if other populations may incorporate this feature differently.

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 of research, we did not pursue further into these details and focused on the study of the first order relation of the contextual features.[25]

For a brief summary, Table 2 shows the incorrect reasoning of the interviewed students corresponding to the four contextual features. As indicated from the results, these four contextual features represent the ones that are frequently used by students in their reasoning and are defined as the *physical features* related to Newton III – a physical feature describes an unique contextual aspect of a physical representation that is considered relevant to the related physics concept by experts and/or students.[26]

Table 2. Interview results on student reasoning.

Contextual Features	Incorrect	Mixed	Correct*
Velocity (V)	7	2	0
Mass (M)	7	2	0
Pushing (P)	6	1	2
Acceleration (A)	0	1	8

*Students consider the corresponding contextual feature irrelevant.

IV. Measuring the Effects of Contextual Features on Students' Reasoning

The physical features can be used as the basis to study the detailed structures of student models of a particular concept. To do so, we need a probing instrument that can provide measurement on the effects of the individual physical features on students' reasoning. However, many of the questions in existing instruments are not designed with isolated physical features. For example, the question shown in Figure 4 mixes two physical features, mass 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 of both.

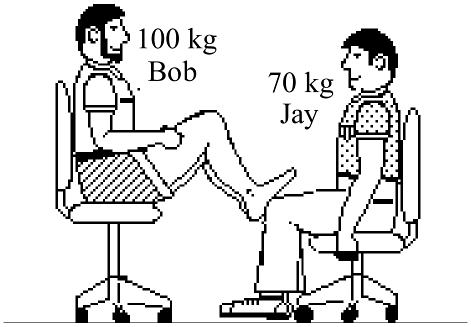


Figure 4. Question 11 in FCI (original version). This question is on Newton III involves two contextual features of mass and pushing.

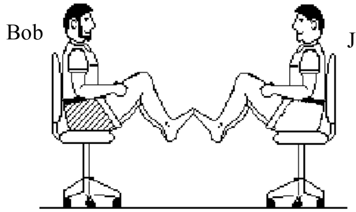
A New Multiple-Choice Instrument on Newton III and Model-Based Assessment

Based on the existing research in the literature 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. In order to measure the possible mixing of students' use of their models,[27] for each of the four physical features, we designed three questions using different context settings. In Figures 5 and 6, we show two sample questions: question 7 on velocity and question 15 on mass. As we can see, with question 15, if a student selects choice (d), "*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.

In a soccer game, two player, John and Tom who happen to have same weight, 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 a 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.*

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

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

Table 3. Questions in the new survey on Newton III and the physical features they are designed to measure.

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

The complete survey is included in the Appendix. In Table 3, we group the questions based on the physical features that these questions are intended to measure.

This instrument is designed to be used with Model Analysis, and each physical feature is used as an independent dimension to represent students' model structures. On each dimension (corresponding to a specific physical feature), we can further construct a multi-dimensional model sub-space spanned by the physical models involved with this particular physical feature. In the case of Newton III, the sub-spaces for all four physical features happen to have three dimensions (three physical models for each physical feature). In general, the dimensions of these sub-spaces can be different. As an example, with the physical feature of pushing (P), we define the following physical models based on research:[28]

M_0^P: The 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 the source of the force. (correct model)
M_2^P: The one exerting the force will exert a larger force during interaction. (incorrect model)

In Table 4, we list all the physical models corresponding to the four physical features. The associations between the choices of the questions and the physical models are also listed and are used in our later analysis to analyze students' raw responses. Note that in this example, we have only one incorrect model for each physical feature. In general, one can imagine situations where more than one incorrect model exists.[29]

Table 4. The physical models corresponding to the four physical features and the associations between the choices of the questions and the physical models.

Physics Features	Physical Models	Questions/ Choices*		
Velocity		**1**	**5**	**7**
	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
Mass		**4**	**9**	**15**
	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
Pushing		**3**	**13**	**14**
	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	a,c	b,d	b,d
Acc.		**2**	**8**	**10**
	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

* "x" is used to represent all other choices.

Validation of the Instrument

As a partial validation on measurement consistency of this instrument, we selected 6 questions from the survey and used them in the 9 interviews that are also used to confirm the "existence" of the physical features. This approach reduces the time span of 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 analysis of open-ended explanations and discussions from students. In the second task of validating the instrument, we first ask students to solve the multiple-choice questions altogether and then have them explain their reasoning used to generate their answers. [30] The consistency between students' responses to the questions and their reasoning is used to evaluate if the questions can measure accurately the underlying student models.

In all the interviews, students' reasoning and their selections of the answers are found consistent, and we didn't find any apparent communication problems with the questions – most students understand the questions well and their explanation show consistency between their understanding and the intentions of the experts. As often observed in interviews, some students may change their minds when explaining their reasoning. In this research, we have 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 choices of the multiple-choice questions and their modified explanations are also consistent with their new answers.

As discussed earlier, students can use mixed ideas in their reasoning. When multiple questions related to a single physical feature are presented to students, they may respond with different models on different questions. This indicates that using these questions we can obtain measurement on students' mixed model states and the significance of different contexts in triggering students' use of models.

V. Assessment with Multiple-Choice Instrument

The Population

The new multiple-choice survey was used in 5 introductory physics courses at Kansas State University (Fall, 1999). These courses include: Physical World (PW), an conceptual- physics course for non-science majors with no math pre-requisites; 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 5. All courses used traditional instruction.

Table 5. Student background and course information of the five intro physics courses at Kansas State University.

Courses	Format	Majors	Prerequisites
Physical World	Algebra, Mech.	Liberal arts	No math
General Phys. 1	Algebra, Mech.	Life science	Algebra
General Phys. 2	Algebra, E&M	Life science	Algebra
Eng. Phys. 1	Calculus, Mech.	Eng and Phys	Calculus
Eng. Phys. 2	Calculus, E&M	Eng and Phys	Calculus

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 haven't had any 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, using this setup, we can approximately study the change of student understandings 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 only provide limited descriptions of the operations of these tools. More details are provided in references 2 and 8.

Concentration Analysis

As a way to validate the effectiveness of this multiple-choice instrument, we first use the *Concentration Analysis* to evaluate the design of the distracters.[31] As we learn from qualitative research into student learning, student responses to problems in many physical contexts can be considered as the result of their applying a small number of conceptual models. The way in which the students' responses are distributed on research-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 students have less commonality in reasoning (sometimes, this situation corresponds to the case where most students have no systematic model).

It is then convenient to construct a simple measure that gives the information of how all students' responses are distributed among the choices of a particular multiple-choice question. This measure is defined as the concentration factor, $\boldsymbol{C}$, which is a function of students' responses and takes a value in [0,1]. Larger values represent more concentrated responses with 1 being a perfectly correlated response and 0 a random response. We want all other situations to generate values between 0 and 1. This concentration factor can be calculated with Eq. (1),

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

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.

Due to the non-linearity of the concentration factor, a value greater than 0.5 represents a fairly high concentration (>60% students select the same choice of a question). A value between 0.2 and 0.5 is considered as 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 can either have no consistent reasoning at all and respond rather randomly or they may have a evenly distributed population for all the possible models involved in the question (further clarification of the detail requires looking at the content of the question and student behavior in interviews).

In Table 6 we show the results of the concentration analysis of student responses on this multiple-choice test for the five courses. For easy comparison, we first calculated the concentration factors for all 16 questions and then grouped the questions based on the different physical features to obtain the average results which are shown in the table.

Table 6. The average scores and concentration factors 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
	$\boldsymbol{C}$	0.69	0.67	0.34	0.59	0.28
GP1	S	0.12	0.11	0.17	0.85	0.34
	$\boldsymbol{C}$	0.80	0.76	0.40	0.75	0.29
GP2	S	0.25	0.18	0.32	0.64	0.50
	$\boldsymbol{C}$	0.57	0.67	0.49	0.50	0.38
EP1	S	0.20	0.23	0.27	0.67	0.35
	$\boldsymbol{C}$	0.61	0.55	0.37	0.55	0.33
EP2	S	0.35	0.29	0.46	0.65	0.52
	$\boldsymbol{C}$	0.49	0.54	0.51	0.51	0.40
Average	S	0.20	0.18	0.27	0.71	0.41
	$\boldsymbol{C}$	0.63	0.64	0.42	0.58	0.34
Standard Deviation	ΔS	0.11	0.08	0.14	0.10	0.11
	$\Delta\boldsymbol{C}$	0.13	0.13	0.12	0.13	0.09

The concentration analysis provides a convenient tool to study if a question can pick up 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 involves two common models, as in the cases of Newton III, a well-designed question often has medium to high concentration on students' pre-test data. As shown by table 6, the questions corresponding to the 4 physical features all have quite high concentration factors, whereas the questions we used to explore certain interesting possibilities, denoted with "others", have systematically lower concentration factors. This indicates that with the 4 physical features, most students have common types of reasoning (models) similar to the ones that we have identified. 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 on questions with the physical features of velocity and mass have high concentration but low scores. This indicates that most students selected the same incorrect answers on these questions (common incorrect models). On the contrary, for questions with the physical feature of acceleration, students' responses show high score and high concentration indicating that most students selected the correct answer. On questions with the physical feature of pushing, students' responses have a medium value for concentration factor, which indicates that students often select between two popular answers. In this case, students usually have mixed state of understanding. To look for the detail of all these possible situations of student models, we need to use our knowledge from qualitative research (the content of the questions) and apply algorithms in Model Analysis to extract the probability states of students use of different models.

Model State Estimation

Using each physical feature as an independent dimension, we analyzed student model states. As detailed in references 2 and 8, the model state for a single students gives the amplitude of distributed probabilities for the student to be triggered into using the different physical models associated with set of questions used in the measurement. The model state for a population gives the amplitude of the distributed probabilities for the population to use the different physical models. These distributed probabilities are stored in a model state vector and the structure of it 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 mix to use different conceptual models in contexts that are regarded equivalent by experts. As shown by research, such mixed states are often a crucial intermediate stage of a favorable conceptual change.[32]

To calculate the model states, we first code students' raw responses to obtain single-student model vectors by using the scheme shown in table 4,. Then a class model density matrix is obtained for each course 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.[33]

As discussed in reference 8, the class model state vectors, eigenvectors of a class model density matrix, reflect the set of unique features of the models held by the individual students in the class, whereas the eigenvalues reflect the popularity of the corresponding class model states.

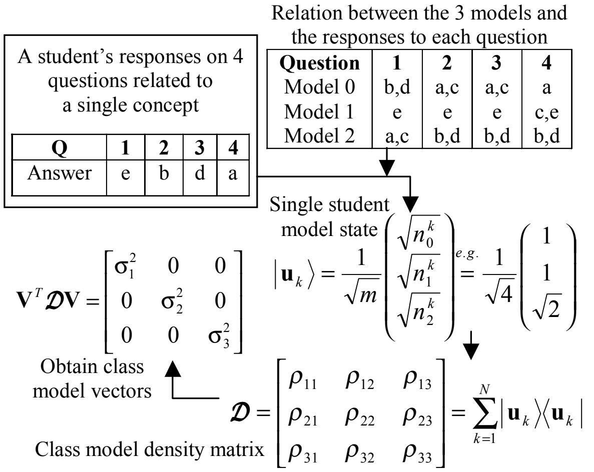


Figure 7. Schematics of the procedures for calculating the class model states.

A useful way to investigate the shift in student thinking between two common models is to create a model plot.[34] As shown in figure 8, a particular model state as well as its eigenvalue can be represented with a point (e.g. B) on a model plot where the horizontal and vertical components equal to the products of the square of the class model state vector's two elements and the corresponding eigenvalue. The values of the two components give the probabilities for the class to apply the models represented with the corresponding axis.

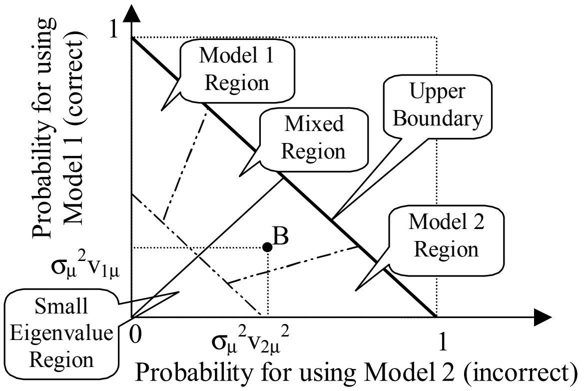


Figure 8. Model plot used to represent the class model states. Model 1 (Model 2) region covers comparatively consistent model states with dominant model 1 (model 2) components. Mixed model region represents mixed model states. Small eigenvalue region represents model states with small eigenvalues.

From a model plot, we can obtain useful information about the class population and individual students' use of their models. In general, the value of the largest eigenvalue can provide a measure on 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, i.e., the class has a somewhat consistent population.

The information on the individual students' using their models is reflected by the eigenvectors (or class model state vectors). If most students in a class are consistent in using their models, which results in "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 are inconsistent in using their models, which results in "mixed" single student model states, the point representing the class model state will be in the mixed model region.[35]

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

From figure 9, we can see that for the physical features of mass and velocity, the primary model state of all the classes stay in the region representing a consistent incorrect model (model 2). This indicates that most students have a dominant consistent incorrect model. The popularity of the incorrect model somewhat decreases 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 these five classes apply their models consistently, i.e., no mixed use of different models. In this situation, the eigenvalue of the primary model state can provide an estimation of the size of the students using the incorrect model.

Student model states with the physical feature of acceleration appear to be the opposite of the situations with mass and velocity. In this case, most students hold a consistent "correct" model where 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

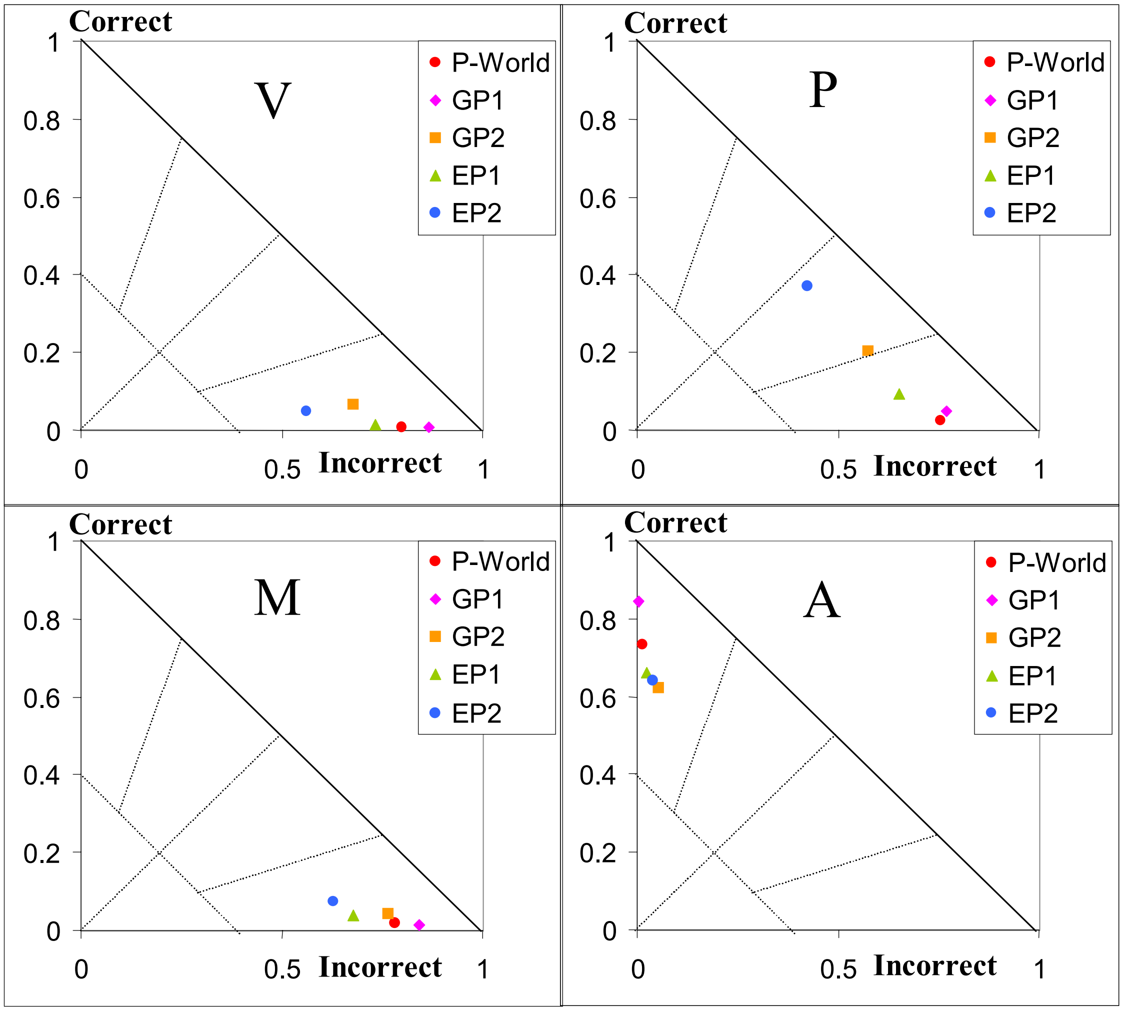


Figure 9. Model plots of student class model states on Newton III with the four physical features: Velocity (V), Mass (M), Pushing (P), and Acceleration (A) . The data is taken from 5 introductory physics course at Kansas State University.

rather that they consider the velocity is the major factor and acceleration is something related to velocity and will not make direct effect.[36]

With the physical feature of pushing, student model states show a different structure. The low level classes still have a dominant consistent incorrect model. As the level of class gets more advanced, student model states become more mixed. The most advanced class (EP2) has nearly a perfectly mixed model state with a quite large eigenvalue (~0.8), which indicates that most students in this class have mixed model states and the structures of the individual single-student model states are also similar. This is very different from the situations with the other physical features and implies a different process in conceptual development.

Implications on Conceptual Development

As recognized by many researchers, the stage of mixed model state is often an important intermediate step for a complete favorable conceptual change.[37] Therefore, we put more emphasis on the study of student reasoning with pushing to see why this physical feature makes different contribution. In our interviews, when asked for the reasoning on the physical features of pushing, many (7/9) students specifically quoted that "*when you 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 these students can even associate these correct ideas with examples such as push against a wall from their experiences. In the following, we summarize some common student behaviors identified in our interviews:

1. Students often use the two sentences discussed above in their explanations on questions involving pushing. With questions that do not explicitly involve the issue of pushing, students look for mass or velocity instantly in their reasoning without even bother to recall the two sentences, which many of them can memorize (especially the first one) and relate to examples from their personal experience. It seems that for these students, the two sentences are 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 "pusher" 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 the fact that student model states are different with the physical feature of pushing: It appears that "Pushing" is often the first and most common example used to introduce Newton III. More importantly, most students all have the experience of being pushed back. Integrating this piece of student experience as examples in instruction can make this side of the concept of Newton III directly linked to students' life experience and presumably more meaningful for students to understand. Therefore, students can have significant changes of their models on this physical feature even with traditional instruction. On the other hand, students' strong naïve models associated with mass and velocity often receive inadequate treatment through traditional instruction and students' changes on their models with these physical features are fairly insignificant.

VI. Implications on Instruction and Summary

This study can provide a piece of evidence for the context dependence of conceptual learning (conceptual changes). 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 contexts used to present the new concept are treated properly, even traditional instruction can make significant impacts on students' conceptual understandings; however, such learning process happens in a highly context-dependent manner. Therefore, instruction should be developed based on a good understanding of the possibilities of student models as well as the effects of contextual features. Successful instruction should also include effective assessment tools to provide accurate and context-rich information of students' state of understanding.

The method discussed in this research can be a useful assessment tool in research and instruction. It has several advantages over score-based methods:

1. It uses multiple-choice instruments making it appropriate and feasible to implement this method in large classes.
2. The probing 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 to both researchers and instructors on the details of how contexts and students' conceptual models interact during the process of conceptual development of a single student and or a population.

In this study, we found that student models show different structures with different physical features and the student model evolution also show 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 are found effective in measuring and analyzing the details of the structures of student models. With this new method, we can obtain detailed quantitative information on student models with a particular physical feature. In addition, the results from these methods can provide explicit information on student understandings with respect to specific contextual features for both researchers and instructors.

Acknowledgment

We would like to thank Alan Van Heuvelen and Eugenia Etkina for their constructive discussions and suggestions.

Endnotes and references:

---

[1] Lillian C. McDermott and Edward F. Redish , "Resource Letter PER-1: Physics Education Research," Am. J. Phys. **67** (9), 755, 1999.

[2] Lei Bao, "Dynamics of Student Modeling: A Theory, Algorithms, and Application to Quantum Mechanics," Ph.D. dissertation, University of Maryland, December 1999.

[3] Gliner, G. (1989). College students' organization of math word problems in relation to success in problem solving. *School Science and Mathematics*, 392-404

[4] Gliner, G. (1991). College students' organization of math word problem solving in terms of mathematical structure versus surface structure. *School Science and Mathematics*, 105-110

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

[6] David Palmer, "The effect of context on students' reasoning about forces," *International Journal of Science Education*, **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," Science Education, **70** (4), 473-496, 1986.

[8] Bao, L. and Redish, E. F. (2001) Concentration Analysis: A Quantitative Assessment of Student States, accepted for publication by *PERS of AJP*, July 2001;
Bao, L. and Redish, E. F. "Model Analysis: Modeling and Assessing the Dynamics of Student Learning," submitted to *PERS of AJP*.

[9] Hestenes, D., & Wells, M. (1992). Mechanics Baseline Test. *The Physics Teacher*, **30**, 159-169.

[10] Chapter 4 of reference 2.

[11] See reference 2 and 8.

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

[13] Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. Learning & Instruction, 4, 45-69.

[14] R. K. Thornton, "Conceptual Dynamics: Changing Student Views of Force and Motion," Proceedings of the International Conference on Thinking Science for Teaching: the Case of Physics. Rome, Sept. 1994;

[15] Detailed definitions and discussions can be found in references 2 and 8.

[16] Supporting evidence and method of investigations can be found in references 2 and 8.

[17] Detailed formulation is discussed in references 2 and 8.

[18] See reference 5;
I. A. Halloun and D. Hestenes, "Common sense concepts about motion," Am. J. Phys. **53** (11), Nov. 1985;
I. Halloun and D. Hestenes, "The initial knowledge state of college physics students," *Am. J. Phys.* **53** (11), 1043-1055 (1985);

[19] Reference 7; and
D.P. Maloney, "Rule-governed approaches to physics – Newton's Third Law," Phys. Ed. **19**, 37 (1984).

[20] E. Engel Clough and R. Driver, "A study of consistency in the use of students' conceptual frameworks across different task contexts," Science Education, **70** (4), 473-496, 1986.;
D. Palmer, "How consistently do students use their alternative conceptions?," *Research in Science Education*, 23, 228-235, 1993

[21] David Palmer, "The effect of context on students' reasoning about forces," *International Journal of Science Education*, **19** (6), 681-696, 1997

[22] See reference 5, 7, and 8.;
In addition, our experience also suggests that it is sometimes possible for students to have some different considerations between object that is "speeding up" and object that is moving at a constant speed.

[23] Zollman, D. A. (1994). Preparing future science teachers: The physics component of a new programme. *Journal of Physics Education*, 29, 271-275.; Zollman, D. A. (1990). Learning cycles for a large enrollment class. *The Physics Teacher*, 28 20-25.

[24] Here, although students consider the acceleration irrelevant, but 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.

[25] We are studying this issue and will report in the near future.

[26] See reference 2 for more details on physical features.

[27] See reference 5.

[28] The numbering tradition of the physical models is different from the one used in reference 5. The system used in this paper makes it more convenient to represent model space when the number of the physical models with different physical features may vary. The superscript "P" is used to represent the dimension related to pushing.

[29] See chapter 5 of reference 5.

[30] Strictly speaking, the development of the survey and the interviews are done in the same time frame. The interviews are conducted over a period of two weeks and the results from the earlier ones are used to make slight modifications to the survey questions. Our initial design of the survey is 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 be quite successful which saves significantly amount of time in our research.

[31] Lei Bao, Edward F. Redish, "Concentration Analysis: A Quantitative Assessment of Student States," Submitted to PERS.

[32] See reference 8, chapter 5 of reference 2; reference 14; and

D. P. Maloney and R. S. Siegler, "Conceptual competition in physics learning," *Int. J. Sci. Educ.*, **15** (3), 283-295, (1993);

S. Vosniadou, "Capturing and modeling the process of conceptual change," Learning & Instruction, (**4**), 45-69, 1994;

[33] See reference 5 for more details.

[34] For more details, see reference 5 and chapter 4 of reference 2.

[35] See reference 5 for details on model plots and class model states.

[36] 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 to the mechanisms behind this phenomenon.

[37] See references in endnote 28.