

From Context Dependence to Scientific Understanding: Modeling Students' Learning Progressions of Force Concepts via Latent Transition Analysis

<https://doi.org/10.63174/xdi.WGBQ7700>

Received: 08 Apr 2025

Accepted: 17 May 2025

Published: 18 May 2025

Open Access



Yue Ming¹, Ying Nie^{1,2}, Lan Yang¹, Huixuan Yang^{1,3}, Lianxin Lian^{1,4}, Yang Xiao^{1,*}, Lei Bao^{5,*}

Abstract: In response to ongoing calls in physics and science education to support conceptual change and promote deeper scientific understanding, research on learning progressions (LPs) has increasingly focused on modeling how students' conceptual reasoning evolves over time. However, substantial evidence indicates that novice learners' understanding remains fragmented and highly sensitive to contextual features. This study investigates the relationship between students' conceptual development and their dependence on context by employing a person-centered analytical method—Latent Transition Analysis (LTA). Pre- and post-test data were collected from 474 students enrolled in a calculus-based introductory physics course, using eight items selected from the 1995 version of the Force Concept Inventory (FCI). These items targeted two fundamental conceptual domains: Force and Motion (F&M) and Newton's Third Law (NTL). The analysis identified four latent statuses for F&M and five for NTL, representing qualitatively distinct levels of understanding ranging from naïve to near-scientific. Results indicate a clear pattern: as students' conceptual understanding progressed, their reliance on surface-level contextual features decreased. These findings suggest a dynamic and interdependent relationship between conceptual development and context sensitivity. This study demonstrates the potential of LTA to reveal developmental trajectories in students' conceptual understanding and underscores the importance of incorporating contextual features in both instructional design and diagnostic assessment strategies.

1. Introduction

A central goal of physics and science education is to foster students' scientific understanding of fundamental concepts.^[1,2] However, a robust body of research has shown that students often enter science classrooms with entrenched misconceptions that are resistant to change through traditional instruction.^[3-8] Consequently, facilitating conceptual change—a shift from intuitive or incorrect ideas to scientifically accurate understandings—has become a central concern in science education.^[9]

In recent years, the learning progressions (LPs) framework has emerged as a powerful tool to examine how students' conceptual understanding develops over time. LPs are defined as “descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time”.^[10] The LP model posits that students progress through a coherent and empirically verifiable continuum—from a *lower anchor* reflecting naïve conceptions, through intermediate levels, to an *upper anchor* representing scientific understanding.^[11,12] Implicit in this framework is the assumption that student reasoning at a given level is relatively consistent across varying problem contexts.^[11,13]

However, an increasing body of research challenges this assumption by highlighting the context-dependence of students' conceptual reasoning. From the knowledge-in-pieces (KiP) perspective, diSessa argued that intuitive physics frameworks consist of fragmented knowledge elements. These elements, which are called *phenomenological primitives* or *p-prims*, are selectively activated depending on contextual cues.^[14,15] Bao and his colleagues demonstrated,

from a conceptual framework perspective, that contextual features such as surface details or representational formats can trigger different conceptual resources, leading to inconsistent reasoning even within the same conceptual domain.^[16-18]

These findings reveal a critical gap in Learning Progression (LP) research. While conceptual change is typically modeled as a smooth trajectory, learners' reasoning often remains fragmented and context-sensitive, particularly in early learning stages. The degree to which context sensitivity diminishes as students' understanding becomes more sophisticated remains underexplored. To address this gap, the present study adopts a person-centered analytical approach—latent transition analysis (LTA)—to investigate how students' understanding of force concepts evolves in relation to their context dependence. LTA is particularly suitable for modeling conceptual change that unfolds in a discrete, stage-like fashion, rather than along a continuous scale, and is thus aligned with theories of stage-sequential development.^[19] This method also enables researchers to track individuals' transitions between latent states over time, making it possible to explore developmental trajectories in ways that cross-sectional approaches or continuous latent variable models such as Rasch cannot.^[19]

While most prior applications of LTA have focused on affective or behavioral variables, such as students' engagement.^[20,21] The present study extends the application of LTA to the domain of physics conceptual understanding. In doing so, we model students' reasoning across two core domains, i.e., Force & Motion and Newton's Third Law, and further incorporate item-level contextual

¹ School of Physics, South China Normal University, Guangzhou, Guangdong, China

² Liang Feng Junior Middle School, Zhangjiagang, Jiangsu, China

³ Shengkang School of Hong ling Education Group, Shenzhen, Guangdong, China

⁴ Yangmei Experimental School, Shenzhen, Guangdong, China

⁵ Department of Physics, The Ohio State University, Columbus, Ohio, USA

* Corresponding Author: bao.15@osu.edu and xiaoyang@m.scnu.edu.cn

features (e.g., mass, acceleration, or motion status). This design allows us to examine how conceptual sophistication is dynamically interwoven with sensitivity to surface-level context, revealing a pattern of context dependence as students' conceptual understanding progress.

2. Theoretical Framework and Literature Review

2.1 The Conceptual Framework Model for Knowledge Integration

The conceptual framework model has been developed through the synthesis of existing learning theories of conceptual change and knowledge integration.^[22] It aims to model students' reasoning pathways by explicitly representing the connections among conceptual components and the contextual features present in problem-solving environments.^[23] This model serves as a useful operational tool to represent students' knowledge structures and assess the degree of knowledge integration they have achieved.^[24] In recent studies, the conceptual framework model has been applied to assess students' understanding of various physics domains, such as force and motion (F&M), Newton's third law (NTL), light interference, mechanical wave propagation, electric circuits, energy, and momentum.^[16–18,22–25] This application facilitates the evaluation of how students integrate multiple conceptual elements across diverse contexts.

Three interlocking assumptions characterize the conceptual framework model. First, learners' ideas and connections are activated by contextual features.^[24] In other words, contextual cues influence which conceptual elements are triggered during reasoning. Second, the structure and integration of activated ideas vary between novices and experts. While experts organize their ideas hierarchically around a central concept, leading to cohesive reasoning pathways, novices often bypass central ideas and rely on local or memorized associations between surface-level features and procedural rules. Therefore, the central idea functions as a conceptual anchor around which other ideas are connected in expert-like reasoning. Third, students' levels of knowledge integration can be evaluated by the quality and completeness of the reasoning pathways connecting contextual features to central ideas and related conceptual components. Once a conceptual framework is constructed for a particular concept, it can inform the development of assessment instruments by mapping students' reasoning pathways across different contexts. Previous research has used such instruments to categorize student understanding based on performance on typical and atypical questions.^[23,24] Typical questions involve familiar contexts frequently encountered in instruction, while atypical questions are designed to probe students' deeper conceptual understanding of less familiar or novel scenarios.

For example, Bao and Fritchman assessed students' assessed students' understanding of NTL using both typical (e.g., force magnitude, direction, type) and atypical (e.g., interaction pair, causality).^[24] Their findings showed that novice students performed poorly across both question types and exhibited high sensitivity to surface features such as perceived causality and object characteristics. Intermediate students showed more consistent reasoning in typical contexts but partial dependence on some surface cues. Expert-like students demonstrated robust understanding across both types of questions and were largely insensitive to surface-level features.

In a parallel study, Nie et al. assessed students' understanding of F&M using questions in four across four contextual categories: Force, No Force, Motion, and No Motion.^[22] The results revealed that context significantly affected novice students' performance, consistent with fragmented knowledge structures. Intermediate students showed improvement in reasoning in the No Force and Motion contexts, while expert-like students demonstrated greater consistency across all four categories.

These findings support the idea that students' understanding of force concepts evolves from context dependence to context independence. The conceptual framework model places a strong emphasis on contextual activation due to its influence on learners' reasoning processes.^[24] In applied research, this model has also guided the design of diagnostic tools to identify specific misconceptions triggered by context. Researchers have often aggregated students' responses to items sharing similar contextual features to compare levels of

understanding across student groups.^[16–18,24]

In the current study, we adopt a person-centered analytic approach—latent transition analysis (LTA)—to model students' conceptual development. Unlike traditional group-mean methods, LTA enables the identification of latent subgroups of students with distinct reasoning patterns and tracks their transitions across these states over time.^[26] This provides a more dynamic representation of how students' conceptual understanding and context sensitivity evolve in tandem during instruction.

2.2 Learning Progressions and Students' Understanding of Force Concepts

The theory of LPs posits that students' understanding evolves through a coherent, continuous, and empirically verifiable sequence of conceptual development, typically from novice to expert-like performance, and this progression is mediated through instruction.^[11,12] LPs have been proposed and validated across various disciplinary core ideas, constructing concepts, and scientific practices.^[13,27,28] The development of an LP typically begins by identifying one or more progress variables, which capture key dimensions of students' thinking within a conceptual domain. Each progress variable is then defined by a series of qualitatively distinct levels, ranging from a lower anchor (representing naïve or fragmented understanding), through one or more intermediate levels, to an upper anchor (reflecting scientific understanding).^[13,29]

The present study focuses on students' learning progressions in two core areas of Newtonian mechanics: (1) the relationship between force and motion (F&M), and (2) the reciprocal nature of forces described by Newton's Third Law (NTL). For both domains, LPs have been previously developed and empirically explored.^[13,30,31]

In the domain of F&M, Alonzo and Steedle developed one of the most widely cited LPs in science education.^[13] This model categorizes students' reasoning across four contextual dimensions—Force, No Force, Motion, and No Motion—and classifies conceptual understanding using ordered multiple-choice (OMC) instruments. Each OMC item is designed so that individual answer choices correspond to reasoning levels on the F&M-LP. After multiple rounds of empirical validation, the authors refined both the LP and the assessment instrument. The final F&M-LP is summarized in **Table 1**. Students at the upper anchor (Level 4) demonstrate a Newtonian understanding of net force and acceleration, whereas those at lower levels exhibit persistent misconceptions, such as believing that motion requires continuous force or that stationary objects are force-free.

To validate and refine the F&M-LP, several follow-up studies used different instruments, populations, and modeling techniques.^[32–35] For example, Steedle and Shavelson assessed high school students using items from the Diagnoser Project to examine consistency with the LP framework.^[34] Fulmer and his colleagues expanded this work by collecting responses from both high school and introductory university students using the Force Concept Inventory.^[33,34,36–38] Their study included a sample of nearly 200 undergraduates enrolled in an introductory psychology course, and aimed to examine whether students' FCI responses—when recoded based on LP levels—aligned with the proposed progression. While their findings offered tentative support for the LP structure, they also noted moderate reliability and challenges in distinguishing lower-level responses, especially in light of the diverse academic backgrounds of the university sample. Nevertheless, the inclusion of undergraduate students in their analysis provides valuable initial evidence that key features of the F&M-LP—such as level distinctions and response patterns—may remain observable and analytically useful at the college level. While these studies provided strong support for the upper anchor, they revealed that the lower and intermediate levels often remain less stable and are sensitive to task design and contextual features.^[33,34]

For the domain of NTL, two LPs have been proposed.^[30,31] The most difference between these two LPs appears in the fourth level, in which Neumann et al. argued that students at Level 4 often misapply Newton's Second Law ($F = ma$) in Third Law scenarios, whereas Morgan et al. suggested that students correctly identify force pairs but mistakenly believe both forces act on the same

Table 1. Force and motion learning progression (F&M-LP)

Level	Description
4	Students understand that the net force applied to an object is proportional to its resulting acceleration (change in speed or direction) and that this force may not be in the direction of motion.
3	Students understand that an object is stationary either because no forces are acting on it or no net force. A student has a partial understanding of forces acting on moving objects.
2	Students believe that motion implies a force in the direction of motion and that nonmotion implies no force. Conversely, students believe that force implies motion in the direction of the force.
1	Students understand force as a push or pull that may or may not involve motion.
0	Way off-track

Adapted from Alonzo and Steedle^[13]

Table 2. Newton’s third law learning progression (NTL-LP)

Level	Description
5	Students understand that, for every action, there is an equal and opposite reaction. The student understands that the action and reaction forces are on separate interacting bodies.
4	Students draw upon the concept of Newton’s second law and attempt to apply this concept to third-law situations.
3	Students understand that both interacting objects apply forces to each other but believe the magnitude is related to the observable behavior of the objects, such as relative motion or being the “cause” of motion.
2	Students understand that both interacting objects apply forces to each other but believe the magnitude is related to the intrinsic properties of the objects.
1	Students believe that when objects interact, there is a force applied by one object onto another and that the magnitude of this force is related to the intrinsic properties of the objects.
0	Way off-track

Adapted from Neumann et al.^[31]

object.^[30,31] Previous empirical studies supported both frameworks to varying degrees.^[39–41] For example, Zhou et al. found that younger students frequently confuse interaction forces with balanced forces, which aligns with Morgan et al.’s view.^[41] Meanwhile, Low and Wilson observed that post-instruction students often attempt to apply Newton’s Second Law to NTL contexts, consistent with Neumann’s model.^[30,31,39–41]

Table 2 presents Neumann et al.’s version of the NTL-LP, which is adopted in this study due to its partial empirical validation.^[30,31,37,42] Fulmer et al. partially validated Neumann’s LP using responses from Singaporean secondary students on four FCI items.^[37] Two recent works demonstrated that FCI Items 15 and 16 can jointly diagnose Level 4 misconceptions—specifically, the misapplication of Newton’s Second Law in Third Law scenarios.^[39,40] Accordingly, it seems possible to diagnose students’ progression levels based on their response patterns across these four items related to NTL measured by the FCI.

While LPs have proven valuable in identifying and modeling student reasoning patterns, they have also faced important theoretical critiques. Early LP models often assumed that conceptual development was internally coherent and linear, consistent with a “knowledge-as-theory” perspective in which students are viewed as holding stable, theory-like mental models.^[10,43] More recent research challenges this assumption by adopting the knowledge-in-pieces (KiP) framework, which posits that learners possess loosely connected knowledge elements that are selectively activated depending on context.^[44,45] Recent empirical findings further support the notion that scientific conceptions and

misconceptions can coexist and compete for activation depending on context.^[46–49]

These developments have led to a more nuanced, middle-ground perspective, in which LPs are recognized as useful tools but must be adapted to account for non-linear, context-sensitive reasoning.^[50] This middle-ground view aligns with the conceptual framework model described in the previous section, which integrates both structured progressions and context activation as dual mechanisms for conceptual change.^[51–54] Accordingly, the current study employs latent transition analysis (LTA) to operationalize this dual framework by modeling how students shift between latent reasoning states over time, and how these shifts relate to their context dependence.

2.3 The Current Study

Building upon the theoretical foundations and empirical findings reviewed above, the present study employs latent transition analysis (LTA) to examine how introductory physics students’ understanding of force concepts develops over the course of instruction. This approach enables the modeling of students’ latent conceptual states and their transitions over time, providing insights into both the structure of students’ conceptual reasoning and the dynamics of their progression.

Specifically, this study aims to extend prior work by investigating how students’ understanding of F&M and NTL evolves in relation to their context dependence—that is, the extent to which their reasoning is influenced by surface features of problems across contexts. By integrating the learning progression

Table 3. Items and their contextual features selected from the FCI 1995 version

Item	Contextual features			
	<i>Force:</i>	<i>No Force:</i>	<i>Motion:</i>	<i>No motion:</i>
F&M	To identify the resulting motion of an object having a nonzero net force.	To identify the resulting motion of an object with net force is zero.	To identify applied force(s) acting on a moving object.	To identify applied force(s) acting on a stationary object.
22	√			
27		√		
13			√	
29				√
NTL	<i>Mass:</i> the one with a larger mass exerts a larger force.	<i>Active/ Pushing:</i> only the one that pushes exerts a force (or a larger force).	<i>Acceleration:</i> the one that causes the acceleration exerts a larger force.	<i>Velocity:</i> the one with a larger velocity exerts a larger force.
4	√			
15	√	√	√	
16	√	√		
28	√	√		

Note: “√” indicates the contextual features were addressed by an item. No item in the FCI was developed for the *Velocity* context feature.

framework with the conceptual framework model of knowledge integration, this study addresses the coexistence of misconceptions and scientific conceptions as students move toward expert-like reasoning.

The following research questions guide the current investigation:

1. Are there qualitatively distinct subgroups of introductory physics students who exhibit different patterns of understanding in relation to force concepts? If so, to what extent do these subgroups align with the levels described in the proposed learning progressions for F&M and NTL?
2. To what extent can students' conceptual understanding progressions be explained by their dependence on the contextual features embedded in force-related problems?
3. How do students' latent statuses—i.e., their underlying conceptual understanding of force concepts—change from pre- to post-instruction, and what patterns characterize these transitions?

3. Methodology

3.1 Participants and Instruments

To investigate students' learning progression in force concepts and their dependence on contextual features, this study employed LTA using data collected through the 1995 version of the FCI.^[38] The dataset was drawn from a large public research university in the United States, ranked within the top 60 nationally and admitting approximately 52% of applicants. A total of 474 students ($N = 474$) enrolled in calculus-based introductory physics courses completed the FCI. The FCI was administered as a pre-test during the first week of instruction and again as a post-test during the week preceding the final exam. The average score on the FCI was 33.77% at the pre-test (Cronbach's $\alpha = .760$) and 45.56% at the post-test (Cronbach's $\alpha = .825$), providing a baseline for interpreting students' conceptual profiles and progression.

The FCI consists of multiple-choice items covering various force-related topics.^[38,55-59] For the purposes of this study, only items aligned with the **LP frameworks for F&M and NTL** were selected. Based on prior research, eight items were identified as representative of these two LP domains.^[33,37,60]

Table 3 categorizes these items by the contextual features they represent.^[13,22,51]

For the F&M domain, four distinct contextual categories were used: Force, No Force, Motion, and No Motion. Each of the four selected items—Items 13, 22, 27, and 29—was mapped to one of these categories. These categories are known to activate different student conceptions and are integral to identifying shifts in reasoning along the F&M-LP.^[24,51]

For the NTL domain, four additional items—Items 4, 15, 16, and 28—were selected. These items address one or more of the following contextual features: Mass, Acceleration, Active/Pushing, and Velocity. While all four NTL items assess the Mass context, only Item 15 specifically addresses the Acceleration context. Some items—such as Item 28—combine multiple contextual features, in this case, Mass and Active/Pushing. Notably, none of the selected FCI items target the Velocity context independently, which is a limitation previously noted in the literature.^[51]

3.2 Analysis

Each item on the FCI includes five response options: one scientifically correct answer and four distractors designed to elicit common student misconceptions. Rather than dichotomously scoring responses as correct or incorrect, the present study retained the full set of response options for each item. This approach preserves the diagnostic value of students' choices and enables classification into distinct latent reasoning patterns. Accordingly, latent transition analysis (LTA) was used to classify students into latent statuses based on their selected options and to track how their membership in these statuses changed over time.

LTA is a longitudinal extension of latent class analysis that models transitions between qualitatively distinct cognitive states over multiple time points. A recent study explored the dynamic patterns of student role transitions across asynchronous (AOD) and synchronous (SOD) online collaborative discussions.^[20] The researchers employed latent transition analysis (LTA), a longitudinal extension of latent profile analysis (LPA), to examine the dynamic changes in roles across discussion modes (from AOD to SOD), estimating item response probabilities (students' engagement in specific indicators), class prevalence (proportion of each role), and transition probabilities (likelihood of

role change). Through LTA-based dynamic modeling, the study not only identified the static distribution of student roles but also uncovered the asymmetry in role transitions across discussion settings (e.g., highly active roles were more likely to decline, while low-active roles were less likely to improve) and structural heterogeneity (distinct behavioral patterns associated with different transition categories).^[20] These findings offer quantitative evidence to inform the design of targeted intervention strategies, such as enhancing the stability of active roles.

It is particularly well-suited for modeling stage-like conceptual development, which aligns closely with the structure of learning progressions (LPs).^[61,62] The rationale for employing LTA in this study is twofold.

First, LTA captures conceptual change as a discrete and categorical process rather than as movement along a single continuous dimension.^[19] This distinction is important because LPs define conceptual understanding as a series of qualitatively different levels. In contrast, Rasch-based approaches typically impose a continuous, unidimensional scale that may overlook important structural features of student reasoning.^[63] Therefore, the LTA was used to understand how the data conform to the LPs of force concepts.

Second, LTA provides a framework for modeling latent cognitive constructs that are not reducible to a single trait.^[19] In the context of force and motion understanding, Alonzo and Steedle identified four conceptual aspects—Force, No Force, Motion, and No Motion—that collectively characterize students' reasoning.^[13] Likewise, the present study emphasizes that students' responses may be differentially influenced by contextual features embedded in FCI items (e.g., mass, acceleration, or motion status), and that such variation provides key evidence for context-sensitive knowledge integration.

To conduct the LTA, the PROC LTA procedure in SAS used.^[19] The analysis began by determining the optimal number of latent statuses that best described the data. Specifically, five models were estimated, containing 2 to 6 latent statuses, respectively. Model fit was evaluated using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the likelihood-ratio chi-square statistic (G^2).^[64,65] Following prior research, we assumed that the latent profiles maintained the same structure across pre- and post-instruction time points, rather than conducting separate LPAs for each occasion^[21]. This approach allowed for conceptual comparability of the profiles over time and enabled meaningful interpretation of longitudinal transitions. Within this framework, the number of latent profiles was determined by comparing LTA models of increasing complexity, while allowing profile membership to vary across time. The model with the lowest AIC and BIC values and a statistically significant improvement in G^2 statistic over adjacent models ($p < 0.05$) was selected as the best-fitting model.

Once the optimal model was identified, it produced a set of interpretable parameters, including: (1) Latent status membership probabilities at pre-instruction (i.e., the proportion of students in each latent class before instruction), (2) Transition probabilities between latent statuses from pre- to post-instruction, and (3) Item-response probabilities, which express the likelihood of selecting each item option given latent status membership and time point. These item-response probabilities provide the basis for interpreting each latent status in terms of conceptual coherence, consistency with the proposed learning progression levels, and context sensitivity. By analyzing shifts in latent class membership and transition pathways, this study aims to uncover not only how students progress conceptually, but also how their reasoning remains contingent on the contextual features of physical scenarios.

4. Results

4.1 LTA of students' understanding of force and motion

To identify the optimal number of latent statuses that best represented the heterogeneity in students' conceptual understanding of force and motion (F&M), latent transition models with 2 to 6 statuses were fitted to the data. **Table 4** presents the model fit statistics including log-likelihood (G^2), degrees of freedom (DF), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and likelihood-ratio tests between adjacent models.

Among all candidate models, the 4-status solution provided the best fit to the data. It yielded the lowest AIC and BIC values and showed a statistically significant improvement over the 3-status model based on the likelihood-ratio test

Table 4. Comparisons of models for the F&M data

Number of statuses	G^2	DF	AIC	BIC	ΔG^2	ΔDF	p
2	3015.85	390589	3085.85	3231.49			
3	2868.20	390568	2980.20	3213.23	147.64	21	0.000
4*	2717.75	390545	2875.75	3204.48	150.46	23	0.000
5	2671.88	390520	2879.88	3312.64	45.87	25	0.007
6	2622.25	390493	2884.25	3429.37	49.62	27	0.005

Note: Asterisk (*) indicates the selected model.

Table 5. Item-response probabilities and prevalence of latent statuses (F&M data)

	Item-response Probabilities of Latent Status				
	Naïve	Mixed	Tran. hybrid	Near-scientific	
Item 13. To determine the force(s) acting on a ball that was thrown straight up.					
A. A downward force of gravity along with a steadily decreasing upward force.	0.086	0.071	0.047	0.042	
B. A steadily decreasing upward force on the way up and a steadily increasing downward force of gravity on the way down.	0.256	0.265	0.202	0.000	
C. An almost constant downward force of gravity and a steadily decreasing upward force on the way up or a steadily increasing downward force of gravity on the way down.	0.513	0.485	0.738	0.182	
D. *An almost constant downward force of gravity only.	0.127	0.179	0.000	0.776	
E. No force.	0.018	0.000	0.013	0.000	
Item 22. To determine the rocket's speed when its engine produced a constant force on it.					
A. Constant.	0.340	0.429	0.281	0.315	
B. *Continuously increasing.	0.210	0.214	0.342	0.574	
C. Continuously decreasing.	0.022	0.044	0.052	0.023	
D. Increasing for a while and constant thereafter.	0.375	0.295	0.282	0.082	
E. Constant for a while and decreasing thereafter	0.053	0.018	0.043	0.006	
Item 27. To determine the speed of a box that was moving across a horizontal floor at a constant speed if the force acting on it was canceled.					
A. Immediately come to a stop.	0.568	0.683	0.000	0.063	
B. Continue moving at a constant speed for a while and then slow to a stop.	0.084	0.159	0.179	0.070	
C. *Immediately start slowing to a stop.	0.348	0.146	0.800	0.842	
D. Continue at a constant speed.	0.000	0.007	0.013	0.025	
E. Increase its speed for a while and then start slowing to a stop.	0.000	0.005	0.008	0.000	
Item 29. To determine the force(s) acting on an empty office chair that is at rest on a floor.					
A. A downward force of gravity only.	0.608	0.012	0.065	0.016	
B. *A downward force of gravity and an upward force exerted by the floor.	0.000	0.797	0.563	0.767	
C. A downward force of gravity and an upward force exerted by the floor.	0.000	0.054	0.019	0.012	
D. A downward force of gravity, an upward force exerted by the floor and a net downward force exerted by the air.	0.369	0.107	0.313	0.199	
E. No force.	0.023	0.030	0.040	0.006	
● Prevalence of Statuses					
	Per-test	0.317	0.183	0.358	0.142
	Post-test	0.000	0.416	0.323	0.261

Note: Item-response probabilities are constrained to be equal on pre- and post-test occasions. Bold font indicates the enormous item-response probabilities given Latent Status for an item. Starred options are the keys for corresponding items.

($\Delta G^2(23)=150.46, p < 0.001$). Additionally, qualitative interpretability of latent profiles supported the selection of the 4-status model, consistent with prior studies using latent class and transition models in previous research. [20,21,26,63]

The 4 latent statuses were labeled based on students' item-response probabilities: (1) Naïve, (2) Mixed, (3) Transitional Hybrid, and (4) Near-Scientific (see Table 5 and 6). Table 5 and 6 summarized the conditional item-response probabilities, latent status prevalence at each time point, and transition probabilities from pre- to post-test. These labels were informed by both the structure of student misconceptions across the four F&M contextual categories (Force, No Force, Motion, No Motion) and their alignment with previously proposed levels in F&M learning progressions.

The Naïve status is characterized by consistent misunderstandings across all four contextual features of force and motion. Students in this group predominantly subscribe to the intuitive belief that *motion implies force* and *force implies motion* in the Motion or Force contexts (items 13 and 22). For example, in item 22 (Force context), 37.5% of these students selected option D and 34.0% selected option A—both reflecting the misconception that a constant force causes an object to move at a constant speed. In the *No Force* context (item 27), a majority (56.8%) reported that the object would stop immediately if the applied force was canceled, further reinforcing the flawed association

between motion and force. Similarly, in the *No Motion* context (item 29), 60.8% selected option A, indicating the belief that no motion implies no force, in which the downward force of gravity is understood as a pull by the earth. These patterns reflect reasoning dominated by context-triggered intuitive ideas rather than integrated conceptual understanding.

In contrast, the Mixed status represents an intermediate status where students begin to develop context-specific understanding. Most notably, 79.7% of students in this group correctly answered item 29 (*No Motion*), suggesting that they have successfully integrated the concept of force balance in static scenarios. However, their responses to items involving the *Force*, *No Force*, and *Motion* contexts reveal lingering misconceptions similar to those seen in the Naïve status. This pattern implies that the *No Motion* context may be the most accessible entry point for conceptual integration in introductory physics learners.

Students in the Transitional Hybrid status demonstrate a broader, though still incomplete, conceptual development. A majority responded correctly to both the *No Force* (80.0% correct on item 27) and *No Motion* (56.3% correct on item 29) items, indicating growing coherence in their understanding of force equilibrium and inertia. In the *Force* context (item 22), their response distribution reveals partial progress: although 34.2% selected the scientifically ac-

Table 6. Transition Probabilities Among Latent Statuses (F&M data)

	Latent Status			
	Naïve	Mixed	Tran. hybrid	Near-scientific
<i>Naïve</i> : showing various misunderstandings of F&M activated by all contextual features.	0.000	0.796	0.125	0.079
<i>Mixed</i> : showing limited scientific understanding of F&M and misunderstandings activated by some contexts.	0.000	0.890	0.096	0.014
<i>Transitional hybrid</i> : showing partially scientific understanding of F&M and misunderstandings activated by few contexts.	0.000	0.000	0.743	0.257
<i>Near-scientific</i> : showing almost consistent understanding of the scientific statement of F&M and minor misunderstanding.	0.000	0.000	0.000	1.000

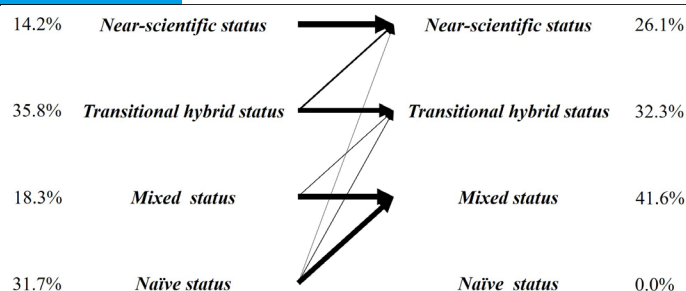


Figure 1. Latent status distributions and transition probabilities from pre- to post-test for force and motion understanding. Each node represents a latent status at either the pre-test (left) or post-test (right), labeled with the corresponding proportion of students in that status. Arrows indicate transitions taken by at least 5% of students, with arrow thickness proportional to the conditional probability of the transition. The diagram highlights the directional flow of conceptual change, showing predominant transitions from less coherent (e.g., Naïve) to more coherent (e.g., Near-Scientific) reasoning statuses.

curate option B, the combined probability of selecting incorrect options D and A (56.3%) suggests that the misconception of constant force causing constant velocity persists. Similarly, in the *Motion* context (item 13), 73.8% selected option C, implying that students still associate motion with an active, diminishing force—an interpretation inconsistent with Newtonian mechanics. Compared to those in the Mixed status, these students appear to have incorporated the *No Force* context into their mental models, yet their reasoning remains vulnerable to context-activated misconceptions.

Students in the Near-Scientific status show the most consistent application of Newtonian reasoning across all four contexts. For the *Motion*, *No Force*, and *No Motion* contexts (items 13, 27, and 29), over 75% selected the correct option, reflecting well-integrated conceptual understanding. Nevertheless, even in this advanced group, the *Force* context (item 22) proved challenging: only 57.4% selected the correct option, and 31.5% still held the misconception that a constant force results in constant speed. This suggests that even high-performing students may struggle to fully override entrenched intuitive beliefs, particularly when context features strongly cue everyday reasoning.

The prevalence of latent statuses was further examined. At pre-test, the Transitional Hybrid status was the most common (35.8%), followed by Naïve (31.7%), Mixed (18.3%), and Near-Scientific (14.2%). After instruction, substantial shifts occurred: the Naïve status became nearly extinct (<0.1%), while the Mixed and Near-Scientific statuses increased to 41.6% and 26.1%, respectively. This distribution suggests broad movement toward more sophisticated understanding, particularly in context-sensitive areas.

Transitions among latent statuses from pre- to post-test are summarized in Table 6 and visualized in Figure 1. The diagonal entries of the transition matrix represent the probability that students remained in the same latent status after instruction. These probabilities indicate a high degree of stability for the more advanced statuses. Specifically, students in the Near-Scientific status at pre-test had a 100% probability of remaining in that status post-test, reflecting well-consolidated conceptual understanding. The Mixed and Transitional Hybrid groups also demonstrated substantial stability, with 89.3% and 74.3% of students remaining in their respective statuses. In contrast, the Naïve status showed virtually no retention (<0.1%), suggesting that nearly all students in this group experienced some degree of conceptual change following instruction.

The off-diagonal entries reveal a clear and consistent trend of progression toward more scientifically accurate reasoning. Students in the Naïve status who transitioned were most likely to enter the Mixed status (79.6%) or the Transitional Hybrid status (12.5%). Similarly, among students in the Mixed group, those who advanced most commonly moved into the Transitional Hybrid status (9.6%). Notably, the highest likelihood of progression into the Near-Scientific status was observed among students starting in the Transitional Hybrid group, with 25.7% making this transition.

Overall, these findings suggest that conceptual progression of F&M after instruction tends to occur between adjacent stages along the learning progres-

sion. This stepwise pattern of change supports the notion of stage-sequential conceptual development, wherein students build coherence gradually by integrating individual conceptual features across multiple contexts.

4.2 LTA of students' understanding of Newton's third law

To examine students' conceptual development regarding Newton's Third Law (NTL), latent transition models with 2 to 6 statuses were fitted to the data. Model comparison statistics, including log-likelihood (G^2), AIC, and BIC, were used to identify the model that best balances parsimony and fit. As shown in Table 7, the 2-status model had the lowest BIC, while the 5-status model yielded the lowest AIC. In line with recent simulation research, which cautions that BIC may underestimate the number of profiles in moderate samples, we gave greater consideration to AIC in selecting the optimal model.^[26] Additionally, likelihood-ratio tests (ΔG^2) indicated significant model improvements up to the 5-status solution. The 6-status model, by contrast, did not result in a statistically significant improvement over the 5-status model ($\Delta G^2(26) = 21.72$, $p = .704$). The five-status model was thus selected for its statistical fit, improved conceptual resolution, and alignment with learning progression theory, which emphasizes the importance of distinguishing transitional stages in conceptual development.

The 5 latent statuses were interpreted as: Naïve, Lower-Mixed, Upper-Mixed, Transitional Hybrid, and Near-Scientific. These groups reflect varying degrees of conceptual accuracy and context sensitivity, and they mirror some of the distinctions observed in the F&M domain, though with finer differentiation between mixed profiles.

The Naïve status is characterized by the highest probabilities across all items for responses indicating that, although both objects interact and exert forces on each other, the heavier, larger, or more active object always exerts a greater force. For instance, in item 4, 83.1% of students in this group selected option A, reflecting a common misconception influenced by the *Mass* contextual feature—namely, that the truck with greater mass exerts more force on the smaller car (Table 8). A similar response pattern appears in items 15 and 16, where 32.0% and 43.8% of students, respectively, selected option B, again suggesting a mass-based misconception. Furthermore, a notable proportion of students in this status selected option D (27.3% for item 15 and 21.9% for item 16), indicating the belief that only the more active object (i.e., the car pushing the truck) exerts a force—an error likely triggered by the *Active/Pushing* contextual feature. It should also be noted that students' choice of option B in item 15 may stem from the influence of the *Acceleration* context, leading to the mistaken belief that the accelerating object exerts more force than the one being acted upon. Taken together, these results suggest that students in the Naïve status are susceptible to multiple conceptual features and hold various context-triggered misconceptions regarding Newton's Third Law.

The remaining four statuses show a higher likelihood of correct reasoning on at least one item. The Lower-Mixed and Upper-Mixed statuses demonstrate similar correct response probabilities on item 16 (49.2% and 52.0%, respectively), but a deeper examination reveals important differences. Both statuses

Table 7. Comparisons of models for the NTL data

Number of statuses	G^2	DF	AIC	BIC	ΔG^2	ΔDF	p
2	1797.74	249966	1863.74	2001.06			
3	1678.67	249946	1784.67	2005.21	119.07	20	0.000
4	1630.98	249924	1780.98	2093.07	47.69	22	0.001
5*	1559.62	249900	1757.62	2169.58	71.36	24	0.000
6	1537.90	249874	1787.90	2308.05	21.72	26	0.704

Note: Asterisk (*) denotes the selected mode

Table 8. Item-response probabilities and prevalence of latent statuses (NTL data)

			Item-response Probabilities of Latent Status				
			Naïve	Lower -mixed	Up per-mixed	Tran. hybrid	Near-scientific
Item 4.	To determine the action and reaction forces when a large truck collides head-on with a small car.						
A.	The larger truck exerts more force.		0.831	0.863	0.919	0.220	0.044
B.	The larger truck exerts less force.		0.000	0.021	0.010	0.002	0.000
C.	No force.		0.037	0.000	0.000	0.000	0.000
D.	Only the larger truck exerts a force.		0.055	0.011	0.002	0.003	0.000
E.	*The action and reaction forces are equal.		0.077	0.105	0.069	0.775	0.956
Item 15.	To determine the action and reaction forces when a small car is pushing a truck and speeding up.						
A.	*The action and reaction forces are equal.		0.210	0.119	0.085	0.428	0.948
B.	The larger truck exerts more force.		0.320	0.074	0.133	0.033	0.035
C.	The more active car or the car “causing” the motion exerts more force.		0.197	0.676	0.772	0.539	0.000
D.	Only the more active car exerts a force.		0.273	0.131	0.010	0.000	0.017
E.	No force.		NA	NA	NA	NA	NA
Item 16.	To determine the action and reaction forces when a small car pushes a truck and moves at a constant speed.						
A.	*The action and reaction forces are equal.		0.079	0.492	0.520	0.840	0.915
B.	The larger truck exerts more force.		0.438	0.032	0.022	0.000	0.028
C.	The more active car or the car “causing” the motion exerts more force.		0.000	0.331	0.451	0.138	0.044
D.	Only the more active car exerts a force.		0.219	0.124	0.001	0.004	0.000
E.	No force.		0.264	0.021	0.006	0.018	0.013
Item 28.	Determining the action and reaction forces when a heavier student is pushing a lighter student.						
A.	No force.		0.021	0.157	0.005	0.005	0.082
B.	Only heavier or more active student exerts a force.		0.058	0.479	0.031	0.000	0.058
C.	The heavier or more active student exerts less force.		0.052	0.364	0.058	0.024	0.012
D.	The heavier or more active student exerts more force.		0.698	0.000	0.797	0.246	0.000
E.	*The action and reaction forces are equal.		0.171	0.000	0.109	0.725	0.848
●	Prevalence of Statuses	Pre-test	0.051	0.142	0.611	0.138	0.058
at		Post-test	0.064	0.053	0.385	0.409	0.089

Note: Item-response probabilities are constrained to be equal on pre- and post-test occasions. Bold font indicates the enormous item-response probabilities given Latent Status for an item. Starred options are the keys for corresponding items.

show strong misconceptions activated by the *Mass* context, as indicated by the high probabilities of selecting option A for item 4 (86.3% and 91.9%, respectively). On item 15—which involves *Mass*, *Active/Pushing*, and *Acceleration* features—students in both groups most frequently selected options reflecting that the more active or accelerating object exerts more force (e.g., 67.6% for Lower-Mixed and 77.2% for Upper-Mixed).

However, the groups diverge in item 28. Students in the Lower-Mixed status most commonly selected option B (47.9%), suggesting the belief that only the heavier student applies a force to the lighter one. In contrast, students in the Upper-Mixed status selected option D (79.7%), indicating the belief that the heavier student exerts a greater force—an interpretation consistent with their previous responses. Thus, while the Upper-Mixed status reflects a more stable but incorrect application of mass-dominant reasoning, the Lower-Mixed group reveals a more varied misconception pattern, potentially combining multiple intuitive frameworks. To summarize: (1) students in the Lower-Mixed status may believe either that the more massive or accelerating object always exerts more force, or that only the heavier/active object applies force; (2) students in the Upper-Mixed status appear more consistently committed to the idea that the heavier or causative object always exerts a greater force.

Students in the Transitional Hybrid status are more likely to respond correctly to two items—item 4 and item 28—demonstrating an improved but still partial understanding of Newton’s Third Law. Nevertheless, they remain susceptible to the *Acceleration* context. Specifically, 53.9% selected option C in item 15, reflecting the misconception that the object causing the acceleration exerts a greater force. This pattern suggests that these students may still be incorrectly applying Newton’s Second Law in Third Law contexts.

Finally, students in the Near-Scientific status showed the highest probability of responding correctly across all four items. These students consistently recognized that both interacting objects exert equal and opposite forces on each other, regardless of mass, motion, or apparent activity. For each of the four NTL items, this group demonstrated approximately 90% probability of selecting the scientifically correct response, indicating a robust and context-independent understanding of Newton’s Third Law.

The prevalence of latent statuses was further examined. At pre-test, the most prevalent group was Upper-Mixed (61.1%), followed by Lower-Mixed (14.2%) and Transitional Hybrid (13.8%). After instruction, the Transitional Hybrid group became the most common (40.9%), while Upper-Mixed remained substantial (38.5%), and Near-Scientific increased modestly to 8.9%.

However, the groups diverge in item 28. Students in the Lower-Mixed status most commonly selected option B (47.9%), suggesting the belief that only the heavier student applies a force to the lighter one. In contrast, students in the Upper-Mixed status selected option D (79.7%), indicating the belief that the heavier student exerts a greater force—an interpretation consistent with their previous responses. Thus, while the Upper-Mixed status reflects a more stable but incorrect application of mass-dominant reasoning, the Lower-Mixed group reveals a more varied misconception pattern, potentially combining multiple intuitive frameworks. To summarize: (1) students in the Lower-Mixed status may believe either that the more massive or accelerating object always exerts more force, or that only the heavier/active object applies force; (2) students in the Upper-Mixed status appear more consistently committed to the idea that the heavier or causative object always exerts a greater force.

5. Discussion

5.1 Latent statuses of students’ understanding of force concepts measured by the FCI and their relationship to the proposed LPs

This study employed LTA on selected FCI items to identify students’ conceptual understanding patterns of F&M and NTL. The LTA results revealed four and five qualitatively distinct latent statuses for F&M and NTL, respectively, representing a progression from naïve to near-scientific reasoning. The overall patterns align closely with the learning progression (LP) levels previously proposed by Alonzo and Steedle for F&M and by Neumann et al. for NT^[13,31]

Based on students’ responses to four F&M-related items, four latent statuses were identified. The most advanced group—labeled Near-Scientific—corresponds well to Level 4 of the F&M-LP, in which students are expected to recognize that net force is proportional to acceleration, not velocity.^[13] The Mixed and Near-Scientific statuses share some features of Level 3, wherein students typically begin to understand force interactions on stationary objects but may still harbor misconceptions in other contexts.^[13] For example, students in both groups demonstrated high probabilities of correctly identifying balanced forces acting on a stationary chair (item 29, *No Motion* context), consistent with the cognitive benchmarks of Level 3.

However, further distinctions are notable. While Mixed students continued to exhibit misconceptions in other contexts (e.g., Force and Motion). Meanwhile, the Naïve group demonstrated misconceptions characteristic of Level 2 (e.g., associating force with motion and vice versa), but also displayed confusion consistent with Level 3 (e.g., believing constant force produces constant speed).^[13] For example, students in this status indicated that a box without ap-

Table 9. Transition Probabilities Among Latent Statuses (NTL data)

	Latent Status				
	<i>Naïve</i>	<i>Lower-mixed</i>	<i>Upper-mixed</i>	<i>Tran. hybrid</i>	<i>Near-scientific</i>
<i>Naïve</i> : showing various misunderstandings of NTL activated by all contextual features.	0.507	0.000	0.235	0.258	0.000
<i>Lower-mixed</i> : showing limited scientific understanding of NTL and misunderstandings activated by some contexts.	0.000	0.365	0.342	0.000	0.293
<i>Upper-mixed</i> : showing limited scientific understanding of NTL and misunderstandings activated by some contexts (less than the lower-mixed status).	0.039	0.000	0.530	0.431	0.000
<i>Transitional hybrid</i> : showing partially scientific understanding of NTL and misunderstandings activated by few contexts.	0.102	0.000	0.004	0.894	0.000
<i>Near-scientific</i> : showing almost consistent understanding with the scientific statement of NTL and minor misunderstanding.	0.000	0.031	0.000	0.151	0.818

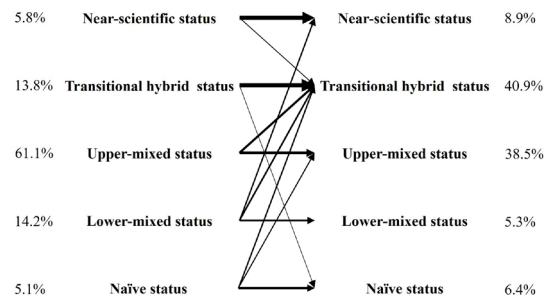


Figure 2. Latent status distributions and transition probabilities from pre- to post-test for Newton’s Third Law understanding. Each node indicates the proportion of students in each latent status before and after instruction. Arrows represent transitions occurring in at least 5% of cases. Arrow thickness reflects transition probability. The figure illustrates predominant shifts from misconception-driven reasoning to increasingly scientific understanding.

plied force will stop moving (item 27, *No Force* context), and that a constant force causes constant velocity (item 22, *Force* context).

These findings indicate that while the lower and upper anchors of the F&M-LP are well-supported, the intermediate levels may benefit from further differentiation. In particular, the Transitional Hybrid status—positioned between the Mixed and Near-Scientific statuses—appears to capture a distinct developmental stage not explicitly identified in the original LP. With more advanced learners, finer-grained distinctions may emerge at the upper end of the progression. Thus, this study supports revising the F&M-LP to include an additional intermediate level between Levels 3 and 4.

Turning to Newton’s Third Law, five latent statuses emerged based on students’ responses to four relevant FCI items. The most advanced status, Near-Scientific, closely aligned with Level 5 of the NTL-LP proposed by Neumann et al., in which students recognize that all forces in an interaction occur in equal and opposite pairs—regardless of mass, motion, or contextual cues. [31,37]

The relationship between the remaining statuses and Levels 1–4 of the NTL-LP, however, is less straightforward. According to Neumann et al.’s original LP, Level 1 describes students who believe that one object applies a force to another based on intrinsic properties such as mass. [31,37] Interestingly, this belief was most prominently observed in the Lower-Mixed status rather than in the Naïve status. For instance, students in the Lower-Mixed group were more likely to claim that only the heavier or more active object applies force (item 28), suggesting they held a context-activated misconception consistent with Level 1. In contrast, students in the Naïve group often recognized that both objects apply forces but incorrectly believed the magnitudes differed depending on mass or motion. These patterns reflect conceptual features associated with Levels 2 and 3 of the NTL-LP, where students acknowledge mutual forces but interpret them as unequal due to observable behaviors or physical properties. [31,37]

Therefore, the boundaries between Levels 1 through 3 in the original NTL-LP appear to overlap in practice. This warrants reconsideration and possible revision of the NTL-LP to better reflect how these ideas manifest simultaneously in students’ thinking.

The Transitional Hybrid status aligned well with Level 4 of the NTL-LP, in which students attempt to apply Newton’s Second Law to Third Law scenarios. For example, students in this group frequently selected item 15’s option C, which implies that the object “causing” the acceleration exerts more force. However, as noted in previous research, this misconception can also stem from

Level 3 thinking—that is, inferring force magnitude from the observed effect of motion. [31,37] Importantly, students in the two Mixed statuses also showed this error but were less consistent in their responses, and their misconceptions were compounded by reasoning typical of lower levels (e.g., misunderstanding of mass effects). This suggests that while option C may reflect Second Law misapplication, it may also indicate less sophisticated misconceptions, particularly in the Mixed groups.

In light of this, the current study proposes that Neumann et al.’s NTL-LP should be extended by inserting additional intermediate levels between Levels 3 and 4, analogous to the proposed revision for the F&M-LP. These additions would better capture the nuances of students’ context-dependent reasoning and provide more actionable benchmarks for instructional design and assessment.

5.2 The relationship between students’ progressively sophisticated conceptual understanding and the dependence on the conceptual features of force concepts

The results of this study provide strong evidence for a developmental relationship between students’ conceptual understanding of force concepts and their dependence on specific contextual features. Across both F&M and NTL domains, students progressed from a high dependence on surface-level contextual cues (e.g., mass, motion, causality) to more generalized and context-independent reasoning as their understanding advanced from Naïve to Near-Scientific.

For the F&M domain, students in the Naïve status exhibited substantial misconceptions across all four contextual categories (*Motion*, *Force*, *No Force*, and *No Motion*). Students in the Mixed status exhibited reduced misconceptions but were still affected by three contextual categories—particularly the *Motion* (item 13), *Force* (item 22), and *No Force* (item 27) contexts. In contrast, those in the Near-Scientific status showed minimal context-triggered errors, with the *Motion* context remaining the most challenging. These findings suggest a hierarchy of conceptual accessibility among contextual features: students tend to first develop accurate reasoning in the *No Motion* context, followed by improvements in the *Force* and *No Force* contexts, with *Motion* remaining the last and most difficult to integrate accurately. This progression aligns closely with prior empirical work that has emphasized context effects in student reasoning. [13,22]

A similar pattern was observed for NTL. The four NTL-related FCI items used in this study involved three key contextual features—*Mass*, *Active/Pushing*, and *Acceleration*—all known to influence students’ interpretations. [51] Students in the Naïve and both Mixed statuses consistently demonstrated misconcep-

tions triggered by either the *Mass* or *Acceleration* contexts. In particular, these students often misattributed force magnitude based on which object appeared more active, larger, or responsible for the motion. However, students in the Transitional Hybrid status exhibited a narrower dependence: their misconceptions were primarily confined to the *Acceleration* context, often misapplying Newton's Second Law in Third Law situations (e.g., inferring that the accelerating object exerts more force). By contrast, students in the Near-Scientific group demonstrated minimal sensitivity to any of the contextual features, consistently applying Newtonian reasoning across diverse scenarios.

These findings support a key premise from the KiP perspective: that conceptual reasoning is highly context-dependent at early stages and becomes more coherent and transferable over time.^[66] As students progress conceptually, they are increasingly able to map physical phenomena to appropriate theoretical models, integrating fragmented elements into more structured, central-idea-oriented knowledge networks. This is consistent with the Conceptual Framework Model, which posits that novices' reasoning is composed of loosely connected, context-activated ideas, while experts organize their understanding around central conceptual nodes that unify reasoning across contexts.^[24]

Taken together, these findings suggest that the extent of a student's dependence on contextual features serves as an important diagnostic indicator of their position along the conceptual progression. The latent statuses identified via LTA offer a nuanced lens for evaluating this relationship, providing strong empirical support for modeling conceptual understanding and context-dependence as interwoven dimensions of learning.

From an assessment perspective, these findings highlight the need for diagnostic instruments that are deliberately designed around contextual features. Several studies have taken this approach. For example, Bao et al. developed a context-based multiple-choice assessment specifically to probe how various contextual cues affect students' understanding of NTL.^[51] More recently, researchers working within the conceptual framework paradigm have expanded this approach to develop multiple context-sensitive instruments across physics domains.^[17,22–25] The current study further illustrates the utility of LTA as a methodological tool for capturing the dynamic interplay between conceptual development and context activation. This approach provides a powerful means for refining learning progressions, informing targeted instruction, and designing assessments that reveal the deeper structure of student thinking.

5.3 The transitions of latent statuses of students' understanding of force concepts over time

The results of this study indicate that, overall, students tended to remain in the same latent status from pre- to post-test—with the notable exception of those initially classified in the Naïve status for Force and Motion (F&M). Students in the Naïve group exhibited high transition probabilities into adjacent, more advanced statuses, suggesting a degree of conceptual movement even in the absence of specific intervention. For students in other statuses, however, transitions to higher levels occurred with less than 50% probability, indicating relatively limited upward progression during the instructional period.

This limited degree of conceptual change is not unexpected, as the current study did not implement any targeted instructional interventions designed to explicitly promote students' progression along the learning progression. The findings, therefore, reflect naturalistic shifts that may result from general course exposure rather than deliberate pedagogical strategies.

Nonetheless, existing research has identified instructional methods that can effectively support conceptual progression. In particular, studies grounded in the conceptual framework approach have shown that instruction which emphasizes knowledge integration around a central disciplinary idea can facilitate more coherent understanding and reduce students' reliance on surface-level contextual features.^[17,22–25] Instructional designs that explicitly highlight the central scientific idea and its connections to various related concepts have been found to help students restructure their knowledge and recognize commonalities across different problem contexts.

Additionally, this study revealed a systematic pattern in the order of contextual features to which students are sensitive during their progression. For example, in the F&M domain, students appeared to first integrate understanding in the *No Motion* context, followed by *Force* and *No Force*, and only later in *Motion*. In the NTL domain, dependency on features such as *Mass*, *Active/Pushing*, and *Acceleration* gradually diminished as students developed more scientific reasoning. These findings suggest that instructional interventions could be designed to scaffold learning across contextual features, gradually expanding students' ability to apply central ideas across increasingly challenging or misleading contexts.

In sum, while the present study was not intervention-based, it offers valuable insights into the trajectory of conceptual change and the context-dependence of reasoning. Future instructional designs could build on these insights by

sequencing learning activities to first consolidate understanding in less ambiguous contexts before introducing those known to activate persistent misconceptions. Doing so may promote smoother and more robust transitions along the conceptual continuum.

5.4 Limitations

This study has several limitations that should be acknowledged when interpreting the findings. First, only eight items from the FCI were included in the analysis—four targeting F&M and four targeting NTL. This limitation was largely due to the analytic framework adopted in the study. Specifically, LTA was conducted using students' full response patterns across five options per item, rather than dichotomous scoring. With four items and five response options each, the number of unique response patterns becomes $5^4 = 625$. Given a sample size of 474 students across pre- and post-tests, the maximum number of potential response patterns is 948 (2×474), which is sufficient to cover the 625 possibilities. However, increasing the number of items exponentially inflates the response pattern space. For example, analyzing eight F&M items would yield $5^8 = 390,625$ possible response combinations—far exceeding the number of participants, and thereby violating assumptions of response pattern sparsity required for stable LTA modeling. Furthermore, although the F&M learning progression was originally developed based on middle school populations, prior research has demonstrated that its conceptual distinctions remain observable among university students. Even so, applying a K–12-derived model to post-secondary learners may entail interpretive constraints that should be acknowledged.

Future studies could address this limitation by using larger-scale datasets, such as those available through repositories like PhysPort, which hosts more than 20,000 matched pre–post student responses to the 1995 version of the FCI.^[67] Such datasets would enable the inclusion of more items, thus offering a more comprehensive representation of students' conceptual structures across a broader range of contexts.

Second, a conceptual overlap among contextual features in some NTL items introduced interpretive ambiguity. Many of the selected items activated multiple conceptual features simultaneously—for example, *Mass*, *Active/Pushing*, and *Acceleration*—making it difficult to isolate which specific feature was responsible for triggering students' misconceptions. This may have contributed to some ambiguity in interpreting intermediate latent statuses. Similar challenges have been noted in prior studies, including mixed conceptual activation within profiles and topic-based variation affecting transition patterns.^[21,26] This limits the precision of the contextual feature diagnosis in the NTL domain. In addition, none of the available NTL items from the FCI addressed velocity-based scenarios, which may have resulted in an underrepresentation of misconceptions specific to that context. However, as our primary research goal was to investigate how patterns of context-sensitive reasoning vary across latent groups rather than to diagnose each misconception type in isolation, this omission is unlikely to substantially affect the validity of our inferences. Future work could employ diagnostic instruments that are more carefully designed to target a single contextual feature per item. One promising example is Bao et al.'s context-based multiple-choice survey, in which each item is explicitly designed to probe a distinct feature of NTL reasoning.^[51] This level of diagnostic specificity would help disentangle the effects of overlapping contextual variables on students' conceptual activation patterns.

6. Conclusion

This study contributes to our understanding of how contextual features interact with students' conceptual development in physics, specifically within the domains of F&M and NTL. By employing LTA, we were able to model the progression of students' conceptual understanding alongside their contextual dependence, revealing distinct latent statuses that corresponded with varying degrees of reasoning sophistication and context sensitivity.

The findings indicate that as students progress from naïve to near-scientific understanding, they exhibit decreasing reliance on context-dependent cues and increasingly consistent application of core scientific principles. This pattern was observed across both F&M and NTL domains, supporting the view that context-independence is a hallmark of conceptual integration. Furthermore, the identified latent statuses generally aligned with existing learning progression models, though our analysis also revealed the need for additional intermediate levels to better capture the nuanced shifts in students' reasoning.

Importantly, the study underscores the value of explicitly considering contextual features in both instructional design and diagnostic assessment. Instruction aimed at promoting deep conceptual learning should not only focus on central scientific ideas, but also on helping students generalize their understanding across varied problem contexts. Similarly, assessments that systematically vary contextual features can more effectively uncover students' misconceptions and

inform targeted intervention.

In sum, this research demonstrates the utility of LTA as a methodological tool for tracing conceptual development and provides empirical support for the integration of context-based reasoning within learning progression frameworks. Future instructional and assessment efforts should leverage these insights to foster more robust, transferable scientific understanding in learners.

Acknowledgments

Yue Ming and Ying Nie contributed equally to this work. This work was supported by the National Social Science Foundation of China (Grant No. CHA200261). Any opinions expressed in this work are those of the authors and do not necessarily represent those of the funding agencies.

References

- [1] L. Bao, K. Koenig. "Physics education research for 21st century learning." *Discip Interdiscip Sci Educ Res* **2019**, *1*, 1, 2.
- [2] "Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century." **2012**, 13398.
- [3] M. Alonso. "Problem solving vs. conceptual understanding." **1992**, *60*, 9, 777–78.
- [4] A. A. diSessa, N. M. Gillespie, J. B. Esterly. "Coherence versus fragmentation in the development of the concept of force." *Cognitive Science* **2004**, *28*, 6, 843–900.
- [5] R. Duit, D. F. Treagust. "Conceptual change: A powerful framework for improving science teaching and learning." *International Journal of Science Education* **2003**, *25*, 6, 671–88.
- [6] J. Minstrell. "Facets of students' knowledge and relevant instruction." **1992**, 110–28.
- [7] S. Vosniadou. "Capturing and modeling the process of conceptual change." *Learning and Instruction* **1994**, *4*, 1, 45–69.
- [8] S. Vosniadou. "On the Nature of Naïve Physics." **2002**, 61–76.
- [9] Y. Hadzigeorgiou. "Young Children's Ideas About Physical Science Concepts." **2015**, 67–97.
- [10] "Taking Science to School: Learning and Teaching Science in Grades K-8." **2007**.
- [11] C. W. Anderson. "Conceptual and empirical validation of learning progressions." **2008**.
- [12] M. Heritage. "Learning progressions: Supporting instruction and formative assessment." **2008**.
- [13] A. C. Alonzo, J. T. Steedle. "Developing and assessing a force and motion learning progression." *Science Education* **2009**, *93*, 3, 389–421.
- [14] A. A. diSessa. "A Friendly Introduction to 'Knowledge in Pieces': Modeling Types of Knowledge and Their Roles in Learning." **2018**, 65–84.
- [15] A. A. DiSessa. "Knowledge in pieces." **1988**.
- [16] Z. Liu, S. Pan, X. Zhang, L. Bao. "Assessment of knowledge integration in student learning of simple electric circuits." *Phys. Rev. Phys. Educ. Res.* **2022**, *18*, 2, 020102.
- [17] L. Xie, Q. Liu, H. Lu, Q. Wang, J. Han, X. Feng, L. Bao. "Student knowledge integration in learning mechanical wave propagation." *Phys. Rev. Phys. Educ. Res.* **2021**, *17*, 2, 020122.
- [18] D. Tong, J. Liu, Y. Sun, Q. Liu, X. Zhang, S. Pan, L. Bao. "Assessment of student knowledge integration in learning work and mechanical energy." *Phys. Rev. Phys. Educ. Res.* **2023**, *19*, 1, 010127.
- [19] S. T. Lanza, L. M. Collins. "A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior." *Developmental Psychology* **2008**, *44*, 2, 446–56.
- [20] M. Wu, F. Ouyang. "Using an integrated probabilistic clustering approach to detect student engagement across asynchronous and synchronous online discussions." *J Comput High Educ* **2025**, *37*, 1, 299–326.
- [21] E. Schlatter, I. Molenaar, A. W. Lazonder. "Learning scientific reasoning: A latent transition analysis." *Learning and Individual Differences* **2021**, *92*, 102043.
- [22] Y. Nie, Y. Xiao, J. C. Fritchman, Q. Liu, J. Han, J. Xiong, L. Bao. "Teaching towards knowledge integration in learning force and motion." *International Journal of Science Education* **2019**, *41*, 16, 2271–95.
- [23] R. Dai, J. C. Fritchman, Q. Liu, Y. Xiao, H. Yu, L. Bao. "Assessment of student understanding on light interference." *Phys. Rev. Phys. Educ. Res.* **2019**, *15*, 2, 020134.
- [24] L. Bao, J. C. Fritchman. "Knowledge integration in student learning of Newton's third law: Addressing the action-reaction language and the implied causality." *Phys. Rev. Phys. Educ. Res.* **2021**, *17*, 2, 020116.
- [25] W. Xu, Q. Liu, K. Koenig, J. Fritchman, J. Han, S. Pan, L. Bao. "Assessment of knowledge integration in student learning of momentum." *Phys. Rev. Phys. Educ. Res.* **2020**, *16*, 1, 010130.
- [26] P. A. Edelsbrunner, M. Flaig, M. Schneider. "A Simulation Study on Latent Transition Analysis for Examining Profiles and Trajectories in Education: Recommendations for Fit Statistics." *Journal of Research on Educational Effectiveness* **2023**, *16*, 2, 350–75.
- [27] K. Neumann, T. Viering, W. J. Boone, H. E. Fischer. "Towards a learning progression of energy." *J Res Sci Teach* **2013**, *50*, 2, 162–88.
- [28] J. Yao, Y. Guo. "Validity evidence for a learning progression of scientific explanation." *J Res Sci Teach* **2018**, *55*, 2, 299–317.
- [29] J. C. Hadenfeldt, K. Neumann, S. Bernholt, X. Liu, I. Parchmann. "Students' progression in understanding the matter concept: STUDENTS' PROGRESSION IN UNDERSTANDING MATTER." *J Res Sci Teach* **2016**, *53*, 5, 683–708.
- [30] B. Morgan, W. Baggett, V. Rus. "Error Analysis as a Validation of Learning Progressions." **2014**.
- [31] I. Neuman, G. W. Fulmer, L. L. Liang, K. Neumann. "Annual meeting of the National Association for Research in Science Teaching (NARST), Rio Grande, Puerto Rico." **2013**.
- [32] H. Jin, P. Van Rijn, J. C. Moore, M. I. Bauer, Y. Pressler, N. Yestness. "A validation framework for science learning progression research." *International Journal of Science Education* **2019**, *41*, 10, 1324–46.
- [33] G. W. Fulmer, L. L. Liang, X. Liu. "Applying a Force and Motion Learning Progression over an Extended Time Span using the Force Concept Inventory." *International Journal of Science Education* **2014**, *36*, 17, 2918–36.
- [34] J. T. Steedle, R. J. Shavelson. "Supporting valid interpretations of learning progression level diagnoses." *J Res Sci Teach* **2009**, *46*, 6, 699–715.
- [35] A. M. Just, A. Vorholzer, C. Von Aufschnaiter. "Employing a Force and Motion Learning Progression to Investigate the Relationship between Task Characteristics and Students' Conceptions at Different Levels of Sophistication." *Education Sciences* **2023**, *13*, 5, 444.
- [36] A. Thissen-Roe, E. Hunt, J. Minstrell. "The DIAGNOSER project: Combining assessment and learning." *Behavior Research Methods, Instruments, & Computers* **2004**, *36*, 2, 234–40.
- [37] G. W. Fulmer. "VALIDATING PROPOSED LEARNING PROGRESSIONS ON FORCE AND MOTION USING THE FORCE CONCEPT INVENTORY: FINDINGS FROM SINGAPORE SECONDARY SCHOOLS." *Int J of Sci and Math Educ* **2015**, *13*, 6, 1235–54.
- [38] D. Hestenes, M. Wells, G. Swackhamer. "Force concept inventory." **1992**, *30*, 3, 141–58.
- [39] D. J. Low, K. F. Wilson. "The role of competing knowledge structures in undermining learning: Newton's second and third laws." **2017**, *85*, 1, 54–65.
- [40] K. F. Wilson, D. J. Low. "On Second Thoughts...: Changes of Mind as an Indication of Competing Knowledge Structures." **2015**, *83*, 9, 802–08.
- [41] S. Zhou, C. Zhang, H. Xiao. "Students' Understanding on Newton's Third Law in Identifying the Reaction Force in Gravity Interactions." *EURASIA J MATH SCI T* **2015**, *11*, 3.
- [42] Y. Ding, G. Zhu, Q. Bian, L. Bao. "Analysis of students' conceptual change in learning Newton's third law with an integrated framework of model analysis and knowledge integration." *Phys. Rev. Phys. Educ. Res.* **2024**, *20*, 2, 020141.
- [43] C. L. Smith, M. Wiser, C. W. Anderson, J. Krajcik. "FOCUS ARTICLE: Implications of Research on Children's Learning for Standards and Assessment: A Proposed Learning Progression for Matter and the Atomic-Molecular Theory." *Measurement: Interdisciplinary Research & Perspective* **2006**, *4*, 1–2, 1–98.
- [44] D. Hammer, T.-R. Sikorski. "Implications of Complexity for Research on Learning Progressions: RESEARCH ON LEARNING PROGRESSIONS." *Sci. Ed.* **2015**, *99*, 3, 424–31.
- [45] R. Lehrer, L. Schauble. "Learning Progressions: The Whole World is NOT a Stage: LEARNING PROGRESSIONS." *Sci. Ed.* **2015**, *99*, 3, 432–37.
- [46] P. Potvin, G. Cyr. "Toward a durable prevalence of scientific conceptions: Tracking the effects of two interfering misconceptions about buoyancy from preschoolers to science teachers." *J Res Sci Teach* **2017**, *54*, 9, 1121–42.
- [47] L.-M. Brault Foisy, E. Ahr, J. Blanchette Sarrasin, P. Potvin, O. Houdé, S. Masson, G. Borst. "Inhibitory control and the understanding of buoyancy from childhood to adulthood." *Journal of Experimental Child Psychology* **2021**, *208*, 105155.
- [48] J. Lin, Y. Xing, Y. Hu, J. Zhang, L. Bao, K. Luo, K. Yu, Y. Xiao. "Inhibitory control involvement in overcoming the position-velocity indiscriminatory misconception among college physics majors." *Phys. Rev. Phys. Educ. Res.* **2023**, *19*, 1, 010112.
- [49] Y. Wen, J. Lin, Y. Ming, J. Zhang, X. Wu, L. Bao, K. Yu, Y. Xiao. "Role of inhibition in overcoming interferences of misconception under similar fea-

- ture saliency: An eye-tracking study of the projectile motion problem." *Phys. Rev. Phys. Educ. Res.* **2024**, *20*, 2, 020121.
- [50] H. Jin, J. N. Mikeska, H. Hokayem, E. Mavronikolas. "Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature." *Science Education* **2019**, *103*, 5, 1206–34.
- [51] L. Bao, K. Hogg, D. Zollman. "Model analysis of fine structures of student models: An example with Newton's third law." **2002**, *70*, 7, 766–78.
- [52] E. F. Redish. "Educational Assessment and Underlying Models of Cognition." **2004**, 221–64.
- [53] L. Bao, E. F. Redish. "Model analysis: Representing and assessing the dynamics of student learning." *Phys. Rev. ST Phys. Educ. Res.* **2006**, *2*, 1, 010103.
- [54] M. C. Linn. "The Knowledge Integration Perspective on Learning and Instruction." **2005**, 243–64.
- [55] P. Eaton, S. D. Willoughby. "Confirmatory factor analysis applied to the Force Concept Inventory." *Phys. Rev. Phys. Educ. Res.* **2018**, *14*, 1, 010124.
- [56] J. Han, L. Bao, L. Chen, T. Cai, Y. Pi, S. Zhou, Y. Tu, K. Koenig. "Dividing the Force Concept Inventory into two equivalent half-length tests." *Phys. Rev. ST Phys. Educ. Res.* **2015**, *11*, 1, 010112.
- [57] D. Hestenes, I. Halloun. "Interpreting the Force Concept Inventory: A Response to March 1995 Critique by Huffman and Heller." **1995**, *33*, 8.
- [58] T. F. Scott, D. Schumayer, A. R. Gray. "Exploratory factor analysis of a Force Concept Inventory data set." *Phys. Rev. ST Phys. Educ. Res.* **2012**, *8*, 2, 020105.
- [59] Y. Xiao, G. Xu, J. Han, H. Xiao, J. Xiong, L. Bao. "Assessing the longitudinal measurement invariance of the Force Concept Inventory and the Conceptual Survey of Electricity and Magnetism." *Phys. Rev. Phys. Educ. Res.* **2020**, *16*, 2, 020103.
- [60] I. Neumann, G. W. Fulmer, L. L. Liang. "Analyzing the FCI based on a force and motion learning progression." **2013**.
- [61] Y. Gao, X. Zhai, B. Andersson, P. Zeng, T. Xin. "Developing a Learning Progression of Buoyancy to Model Conceptual Change: A Latent Class and Rule Space Model Analysis." *Res Sci Educ* **2020**, *50*, 4, 1369–88.
- [62] M. Flaig, B. A. Simonsmeier, A.-K. Mayer, T. Rosman, J. Gorges, M. Schneider. "Reprint of 'Conceptual change and knowledge integration as learning processes in higher education: A latent transition analysis.'" *Learning and Individual Differences* **2018**, *66*, 92–104.
- [63] W. L. Romine, A. N. Todd, T. B. Clark. "How Do Undergraduate Students Conceptualize Acid–Base Chemistry? Measurement of a Concept Progression." *Science Education* **2016**, *100*, 6, 1150–83.
- [64] H. Akaike. "A new look at the statistical model identification." *IEEE Trans. Automat. Contr.* **1974**, *19*, 6, 716–23.
- [65] G. Schwarz. "Estimating the Dimension of a Model." *Ann. Statist.* **1978**, *6*, 2.
- [66] J. P. Smith Iii, A. A. diSessa, J. Roschelle. "Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition." *Journal of the Learning Sciences* **1994**, *3*, 2, 115–63.
- [67] "PhysPort data explorer." *Physport Data Explor.*