Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #47637

Uncovering the Cognitive Roots of Misconceptions in Physics Education for Engineering Students Through Transitional Diagnostic Models

Mr. Amirreza Mehrabi, Purdue Engineering Education & Electronic and Computer Engineering

I am Amirreza Mehrabi, a dual-degree student pursuing a Ph.D. in Engineering Education and a second M.Sc. in Computer Engineering with a specialization in Artificial Intelligence and Machine Learning at Purdue University, West Lafayette. My current research, under the guidance of Prof. Jason Wade Morphew in the ENE department, focuses on cognitive fatigue during adaptive assessments. I completed my first master's degree in Engineering Education at the UNESCO Chair of Engineering Education at the University of Tehran, where I developed a deep interest in human adaptation to technology and modeling human behavior using machine learning and cognitive research. My academic journey began with a B.Sc. in Industrial Engineering from Sharif University of Technology, where I earned the prestigious Gold Medal in the Industrial Engineering Olympiad (Iran, 2021), the highest-level recognition in the field. My research and professional goals are driven by a passion for leveraging technology to enhance educational experiences, optimize adaptive learning systems, and promote equitable access to quality education through data-driven innovations.

Dr. Jason Morphew, Purdue University at West Lafayette (PPI)

Dr. Jason Morphew is an assistant professor at Purdue University in the School of Engineering Education. He serves as the director of undergraduate curriculum and advanced learning technologies for SCALE and is affiliated with the INSPIRE research institute for Pre-College Engineering and the Center for Advancing the Teaching and Learning of STEM. He serves as the course curator for the Freshman semester engineering design course that serves over 2,500 freshman engineering students every year. His award-winning teaching has been recognized for his teaching in the First Year Engineering program and is the Dr. Morphew has also recently taught courses focused on the pedagogy of integrated STEM and educational research methodology. Dr. Morphew's research focuses on the application of principles of learning derived from cognitive science and the learning sciences to the design and evaluate technology-enhanced learning environments. More specifically, his research examines the impact of technologies such as augmented-reality, gesture-based digital environments, microelectronics, and artificial intelligence on learning, interest, identity, motivation, and decision making in STEM. His research views learning through self-regulated learning, constructivist, and embodied cognition lenses.

Uncovering the Cognitive Roots of Misconceptions in Physics Education for Engineering Students Through Transitional Diagnostic Models

Amirreza Mehrabi *
School of Engineering Education &
Department of Electronic and
Computer Engineering
Purdue University
West Lafayette, IN 47907, USA
amehrabi@purdue.edu

Jason Wade Morphew
School of Engineering Education
Purdue University
West Lafayette, IN 47907, USA
jmorphew@purdue.edu

Abstract

This study investigates the identification and persistence of misconceptions among engineering students in foundational STEM courses, focusing on physics concepts assessed through the Force Concept Inventory (FCI). Misconceptions, defined as systematic and deeply rooted alternative understandings, hinder students' ability to master complex topics and apply knowledge effectively. Traditional models such as Item Response Theory and Cognitive Diagnostic Models are limited in their ability to track misconceptions over time, failing to capture how these erroneous beliefs evolve or persist across assessments. To address this gap, we employ a Transition Diagnostic Classification Model (TDCM) that incorporates a Q-matrix to map misconceptions to test items and monitor their transitions as distinct cognitive attributes over successive evaluations. Using data from 1,529 engineering students who completed pre- and post-tests in the Force Concept Inventory, the TDCM reveals the persistence and evolution of misconceptions in areas such as Force and Motion and Vector Addition. Misconceptions in Force and Motion, often aligned with intuitive but incorrect reasoning, exhibit strong persistence, while misconceptions in Vector Addition are more frequently acquired but less stable. These findings align with Conceptual Change Theories, which emphasize the coherence and resistance of misconceptions as cognitive structures embedded in students' mental models. By analyzing transition probabilities and reliability metrics, the TDCM offers actionable insights for educators, facilitating targeted interventions. This study demonstrates the TDCM's effectiveness in enhancing conceptual understanding, supporting data-driven strategies to address persistent misconceptions, and improving outcomes in engineering education.

*Corresponding author

Introduction

Misconceptions, deeply embedded in students' cognitive frameworks, present significant challenges in education, particularly within STEM fields such as engineering. These misconceptions arise not as random errors but as coherent alternative understandings that conflict with established scientific principles, often shaped by prior knowledge and intuitive reasoning[1, 2]. The alternative conceptions that students construct tend to be robust and persist even after instruction, hindering students' ability to engage with more complex topics or to apply scientific and engineering principles in real-world scenarios. The research underscores the importance of identifying and addressing student misconceptions to foster meaningful learning [3, 4]. Conceptual Change Theories explain the persistence of misconceptions as coherent cognitive structures that resist change without targeted interventions. Within constructivist learning theories, misconceptions are not simply incorrect explanations that can be corrected simply by providing students with the correct explanations. Rather, misconceptions are both systematic explanations students build to explain the world, but also emergent and contextual [1]. This means that supporting conceptual change requires targeted and theoretically informed approaches to teaching [5, 6].

Different from other disciplines, engineering requires students to integrate foundational scientific principles, technological innovation, and computational reasoning to address complex and ill-defined real-world problems. This combination makes misconceptions particularly problematic in engineering contexts as conceptual misconceptions may not be evident when students are calculating values, but rather emerge when reasoning about a problem conceptually [7, 8, 9]. In addition, errors in basic physics can lead to flawed assumptions in technology and design reasoning [10]. Misconceptions in science like physics, such as those assessed in the Force Concept Inventory (FCI), are not isolated errors but reflect structured cognitive frameworks that must be understood in relation to the conceptual and cognitive interrelations they embody. Addressing deeply rooted misconceptions necessitates a comprehensive understanding of how misconceptions form, persist, and evolve. It is critical to examine how physics misconceptions held by engineering students and interrelated with each other and how instruction impacts the ways that engineering students experience conceptual change during instruction. This leads us to examine a method for assessing transitions in student misconceptions using adaptive formative assessment to inform targeted and individualized instruction.

This study leverages Transitional Cognitive Diagnostic Models (TCDMs) to investigate which misconceptions persist or dissipate with instruction among first-year engineering students in a physics context. By mapping test items to specific misconceptions and analyzing transitions between cognitive states, the study identifies patterns of persistence and resolution, providing a dynamic understanding of how misconceptions evolve over time [11, 12, 13]. These insights help us address which misconceptions are persistent and which undergo conceptual change after traditional instruction. This leads to the second key question: How do misconceptions in foundational physics concepts, such as Force and Motion and Vector Addition, persist or evolve in engineering students?

Literature review

Two key frameworks explain the nature and persistence of misconceptions in learning; Constructivist Learning Theories, and Conceptual Change Theories. Together, they reveal the cognitive mechanisms underlying misconceptions and inform strategies for their identification and remediation in education.

Constructivist learning theories

Constructivist theory explicit that learning is an active process where individuals construct conceptual knowledge by integrating new information into existing cognitive structures [14, 15, 16, 17]. Conceptual understanding is formed by using the learner's existing conceptual explanation (or schema) to actively decide what information to which the learner should attend, whether the new information agrees with the existing explanation, and how to modify their conceptual explanation [18, 19, 20].

When exposed to new information, learners attempt to engage in assimilation, the process of integrating the new information into their existing conceptual explanations. If the information does not align with their conceptual explanations, learners either reject the new information or engage in accommodation, restructuring or expanding conceptual explanations to account for new information. Accommodation is essential for conceptual change to occur, however, accommodation is effortful and often resisted by learners. The rejection of discrepant information is likely when the information is ambiguous and the learners' prior knowledge appears to explain phenomena effectively [18]. Scaffolding and reflective tasks are critical in helping learners engage in accommodation, guiding them to recognize inconsistencies in their reasoning and revise their frameworks [21, 22]. From this theoretical lens, misconceptions are alternative conceptions that conflict with scientifically accepted knowledge that arise from learners' attempts to make sense of their experiences using existing cognitive structures. Misconceptions are not arbitrary mistakes or aberrant wrong ideas; rather they are largely coherent mental models shaped by everyday observations and intuitive reasoning. As such, misconceptions are a natural part of learning.

Misconceptions occur when learners attempt to assimilate knowledge into pre-existing schema without altering the underlying framework. For instance, a student might assimilate the concept of gravitational force into their understanding of weight, incorrectly assuming that weight determines the rate of fall [21, 22]. Once formed, misconceptions persist because they are logical within the learner's current schema and they often allow learners to make correct predictions in specific contexts, making them difficult to unlearn. Misconceptions can persist, even in the face of evidence, as existing conceptual understanding guides attention and helps learners interpret new information [23]. Resistance to accommodation thus results in persistent misconceptions, particularly in cases where existing schemata are deeply entrenched, or connected to one's identity.

Conceptual change theories

Conceptual change theories, rooted in constructivism, focus on replacing entrenched misconceptions with scientifically accurate conceptions [4, 24, 25]. These theories emphasize cognitive conflict, where contradictions between prior beliefs and new evidence prompt conceptual reevaluation. However, perception is shaped by existing mental models, causing students to misinterpret even clear demonstrations [26, 27]. Conceptual change occurs when new explanations are intelligible, plausible, and useful, but misconceptions persist when these criteria are unmet. Instructional strategies like discrepant events, guided inquiry, and scaffolding aid this transition [27]. In engineering education, hands-on, context-rich learning helps students refine abstract concepts, while iterative engagement with theory and practice fosters conceptual change [4, 25]. Effective instruction involves identifying misconceptions and using targeted scaffolding—such as simulations or design projects—to refine students' mental models through experimentation and validation [28].

Methods of tracking misconceptions

Goris and Dyrenfurth [28] emphasized that misconceptions in engineering education arise from the interconnected domains of science, technology, and problem-solving, deeply rooted in students' prior experiences and mental models. These misconceptions are resistant to change, requiring targeted diagnostic methods and interventions guided by conceptual change theories, such as cognitive conflict. Goris and Dyrenfurth [28] highlight the necessity of creating dissatisfaction with flawed reasoning and replacing it with intelligible, plausible, and fruitful concepts. This perspective aligns with the potential of adaptive online assessment systems, which can dynamically address individual misconceptions in real-time through targeted feedback, interactive simulations, and iterative problem-solving tasks. Diverse statistical and diagnostic frameworks have been proposed to diagnose student misconceptions. Herrmann-Abell et al. [29] utilized Rasch modeling to align student abilities and item difficulties on a common scale, complemented by option probability curves to visualize how misconceptions evolve with ability levels. Goguadze et al. [30] applied Bayesian networks to probabilistically map misconceptions and their underlying causes, achieving high diagnostic accuracy through pretest-posttest data. Mevarech and Zemira [31] integrated diagnostic modeling with mastery learning strategies, combining Rasch analysis and targeted corrective instruction to reduce persistent misconceptions.

Methods

Setting of study

In this study, we analyze data from 1,529 Engineering Students from six institutions who enrolled in a physics class that used the Force Concept Inventory (FCI) on the LASSO website [32] test to assess conceptual understanding. All students completed both the pre- and post-tests, with no missing responses or absences. The data was preprocessed for analysis in the TDCM library in R [33], including adjustments such as creating a Q-matrix and reversing responses, ensuring a robust dataset for tracking changes in understanding and misconceptions over the course. The Q-matrix (Table 1) is a binary framework mapping test items to misconceptions in *Force and Motion*, *Vector Addition*, *Friction*, and *Acceleration and Velocity*. Each row represents an item, each column a misconception, with 1 indicating the item assesses that misconception. Since mastering even one misconception can lead to incorrect responses, we utilized the Deterministic

Inputs, Noisy ”Or” Gate (DINO) model with reverse response logic, where a student answers incorrectly if they have mastered at least one misconception [12, 13, 34, 35].

Table 1: Q Matrix. Full version available in github link

Item number	Force and Motion	Vector Addition	Friction	Acceleration and Velocity
1	0	1	0	1
2	0	1	1	0
3	1	1	0	1
4	0	1	0	1
5	1	1	1	0
6	0	0	0	1
7	1	1	1	0
8	1	1	0	0
...	...	...	...	...
30	1	1	0	1
31	1	1	0	1

Modeling

Addressing misconceptions requires aligning learning objectives with students' reasoning, as misconceptions often manifest as systematic errors that demand detailed analysis in relation to learning goals. TDCMs extend diagnostic classification frameworks by incorporating latent transition analysis to evaluate changes in attribute mastery over time. Building on the DINO model logic, TDCMs enable the study of mastery and non-mastery transitions in a pretest/posttest setup. It combines the advantages of diagnostic classification models with longitudinal analysis, making it suitable for assessing interventions that target specific skills or attributes [12, 13, 34, 35]. TDCMs model the probability of an item response as:

$$P(X_i = 1|\alpha_c) = \frac{\exp(\lambda_{i,0} + \sum_a \lambda_{i,a}\alpha_a + \sum_{(a,b)} \lambda_{i,ab}\alpha_a\alpha_b)}{1 + \exp(\lambda_{i,0} + \sum_a \lambda_{i,a}\alpha_a + \sum_{(a,b)} \lambda_{i,ab}\alpha_a\alpha_b)},$$

Where $\lambda_{i,0}$ is the intercept, $\lambda_{i,a}$ represents the main effects of attribute mastery, and $\lambda_{i,ab}$ accounts for interaction effects [34, 35]. In practice, TDCMs specify mastery status transitions, using pre- and posttest data to estimate transition probabilities. For instance, the probability of transitioning from non-mastery to mastery for a given attribute (α) is modeled as $\tau_{\alpha_2|\alpha_1}$, where α_1 and α_2 represent the pretest and posttest statuses, respectively. This structure enables targeted analysis of intervention effects, distinguishing the effectiveness of instructional strategies across multiple attributes [13, 34, 35].

Results

Model fit statistics

Model fit was assessed using five statistical measures, as presented in Table 2. The Mean Absolute Deviation of Correlations (MADcor) evaluates the alignment between model-predicted and observed item relationships, while the residual covariance measure (100*MADRESIDCOV) assesses unexplained dependencies and the assumption of local independence [9, 36, 33]. The Standardized Root Mean Square Residual (SRMSR) quantifies the average discrepancy between observed and predicted response probabilities, and the Mean Absolute Deviation for Q3 statistics (MADQ3) identifies local item dependencies not captured by the model. Lastly, the Root Mean Square Error of Approximation (RMSEA) evaluates item-level fit while penalizing model complexity [9, 36, 33].

Table 2: Summary of Global Fit Statistics

Statistic	Value
MADcor	0.065
SRMSR	0.084
100*MADRESIDCOV	1.40
MADQ3	0.046
Mean RMSEA	0.0855

The results, detailed in Table 3 and Table 4, indicate a well-fitting model. A MADcor of 0.065 confirms strong alignment between the skill structure and observed data, while a residual covariance of 1.40 suggests the model adequately accounts for inter-item relationships, supporting local independence [9, 36, 33]. The SRMSR of 0.084, though slightly above ideal thresholds, indicates minimal residuals, suggesting the model attributes responses appropriately. Furthermore, a MADQ3 value of 0.045 reinforces that most item dependencies are accounted for, affirming the robustness of the attribute structure in diagnosing student skill mastery [9, 36, 33]. The RMSEA for the items ranged from 0.04 to 0.11 and had a mean value of 0.085. These results indicated that most items demonstrated an acceptable fit RMSEA < 0.1, thus supporting the assumption that the attributes successfully captured the cognitive processes required for these items. Two items, Item 22 and Item 30, had elevated RMSEA values, signaling potential misfitting of the model with these items or gaps in how these items are linked to the attribute structure. These outliers require further investigation to ensure consistent fit quality across all items.

Table 3: Item-Level RMSEA Values

Item	RMSEA Value
Item 1	0.08137
Item 2	0.0922
Item 3	0.0779
Item 4	0.0891
Item 5	0.0715
⋮	⋮
Item 22	0.1715
⋮	⋮
Item 30	0.1152
Item 31	0.0606

Table 4: Item Parameters

Parameter	**Value**
λ_0	-2.433
$\lambda_{1,1}$	–
$\lambda_{1,2}$	–
$\lambda_{1,3}$	–
$\lambda_{1,4}$	–
$\lambda_{2,12}$	–
$\lambda_{2,13}$	–
$\lambda_{2,14}$	–
$\lambda_{2,23}$	–
$\lambda_{2,24}$	2.065
$\lambda_{2,34}$	–
$\lambda_{3,123}$	–
$\lambda_{3,124}$	–
$\lambda_{3,134}$	–
$\lambda_{3,234}$	–
$\lambda_{4,1234}$	–
λ_0	-1.131
$\lambda_{4,1234}$	1.788
λ_0	-1.823
$\lambda_{3,124}$	2.043
λ_0	-1.745
$\lambda_{2,24}$	2.4
λ_0	-0.964
$\lambda_{4,1234}$	2.678

Reliability

Reliability measures model consistency and diagnostic clarity in identifying misconceptions for each misconception (Table 5). These measures were calculated using metrics such as the point-biserial correlation (*pt bis*), information gain (*info gain*), tetrachoric correlation (*polychor*), average maximum transition posterior (*ave max tr*), and the proportion of examinee posterior probabilities exceeding specific thresholds ($P(t > k)$). Additionally, weighted versions of point-biserial and information gain (*wt pt bis* and *wt info gain*) were included. specifically, *pt bis* assesses the strength of the relationship between observed responses and latent traits, *info gain* quantifies the informational value of transitions, *polychor* evaluates the correlation between transitions under the assumption of continuous latent traits, and *ave max tr* measures the consistency of classification confidence [9, 36, 33].

Table 5: Reliability of Different Categories

Reliability Metrics				
Category	**pt bis**	**info gain**	**polychor**	**ave**
Force and Motion	0.500	0.388	0.803	0.831
Vector Addition	0.205	0.218	0.358	0.542
Friction	0.447	0.371	0.770	0.817
Acceleration and Velocity	0.822	0.507	0.971	0.935
Reliability Metrics				
Category	**max tr**	**P**$(t > .6)$	**P**$(t > .7)$	**P**$(t > .8)$
Force and Motion	0.941	0.789	0.598	0.260
Vector Addition	0.293	0.268	0.249	0.224
Friction	0.923	0.750	0.626	0.433
Acceleration and Velocity	0.959	0.922	0.873	0.787
Reliability Metrics				
Category	**P**$(t > 0.9)$	**wt pt bis**	**wt info gain**	
Force and Motion	0.621	0.512	0.512	
Vector Addition	0.249	0.261	0.261	
Friction	0.593	0.493	0.493	
Acceleration and Velocity	0.859	0.648	0.648	

Acceleration and Velocity consistently shows the highest reliability values across all metrics. For instance, *pt bis* = 0.822, *polychor* = 0.971, and *ave max tr* = 0.935 indicate a high level of stability and diagnostic clarity. Proportions exceeding thresholds, such as $P(t > 0.9)$ = 0.787, further confirm the model's strong confidence in classifying misconceptions in this domain. These results suggest that misconceptions in *Acceleration and Velocity* are persistent and readily identifiable, making this domain a critical target for instructional interventions. Its high reliability underscores the effectiveness of the model in capturing the latent constructs and providing actionable diagnostic information (Table 5).

In contrast, *Vector Addition* exhibits the lowest reliability values, with *pt bis* = 0.205, *polychor* = 0.358, and *ave max tr* = 0.542, along with $P(t > 0.9)$ = 0.224. These values reflect low reliability and highlight challenges in diagnosing misconceptions consistently in this domain. The abstract and multidimensional nature of *Vector Addition* likely contributes to this inconsistency, indicating the need for more refined diagnostic tools. *Force and Motion* and *Friction* show moderate reliability, with *pt bis* values of 0.5 and 0.447, respectively. Metrics such as *polychor* = 0.803 for *Force and Motion* and *ave max tr* = 0.817 for *Friction* further support this moderate classification. These findings suggest that while the model performs reasonably well in these areas, there is room for improvement in capturing and diagnosing misconceptions more consistently (Table 5).

Item parameters and misconception prevalence

Item parameters in Table 4 reflect both the sensitivity of test items to specific misconceptions and their effectiveness in diagnosing students' misconception status. The baseline parameter (λ_0) represents the likelihood of an incorrect response for students without misconceptions. The predominantly negative values suggest that misconception-free students are generally less likely to answer incorrectly [12] More negative values, such as $\lambda_0 = -2.433$, indicate items with low sensitivity to misconception-free students, meaning these items strongly differentiate between students with and without misconceptions. Conversely, items with less negative λ_0 values, such as $\lambda_0 = -0.964$, suggest a greater chance of incorrect responses even among those who do not hold the misconception, indicating that some items may be more difficult or susceptible to general errors rather than misconception-driven mistakes [12].

Higher-order parameters (λ_k) illustrate how misconceptions influence incorrect responses and how their interactions compound error likelihood. For instance, $\lambda_{2,24} = 2.065$ suggests that students holding misconception 2 and misconception 4 are significantly more prone to errors on this item, demonstrating how multiple misconceptions reinforce one another. Similarly, $\lambda_{4,1234} = 2.678$ highlights an item where the presence of misconception 4, particularly when interacting with other misconceptions, strongly increases the likelihood of an incorrect response [12]. The interaction effects observed in parameters such as $\lambda_{3,124} = 2.043$ suggest that certain items may be particularly diagnostic in identifying complex misconception patterns. These findings underscore the importance of considering both individual and interacting misconceptions when designing assessments, as certain items are more effective at detecting conceptual misunderstandings than others. Identifying these items is crucial for refining instructional approaches and developing targeted interventions to address persistent misconceptions [12].

Misconception persistence

Proficiency proportions indicate the evolution of misconceptions over time, offering a diagnostic lens into student learning [33]. A reduction in proportions over time reflects the success of interventions, while minimal change indicates the need for more effective strategies to address entrenched misconceptions. The proficiency proportions from Pre to Post all showed a decline over time, reflecting an overall decrease in the number of students that endorse misconceptions in introductory physics [33]. For example, ”Force and Motion” decreased from 0.650 at Pre to 0.393 at Post, while ”Acceleration and Velocity” decreased from 0.722 to 0.418 (see Figure 1). The relatively high probabilities of students transitioning from a misconception state to a

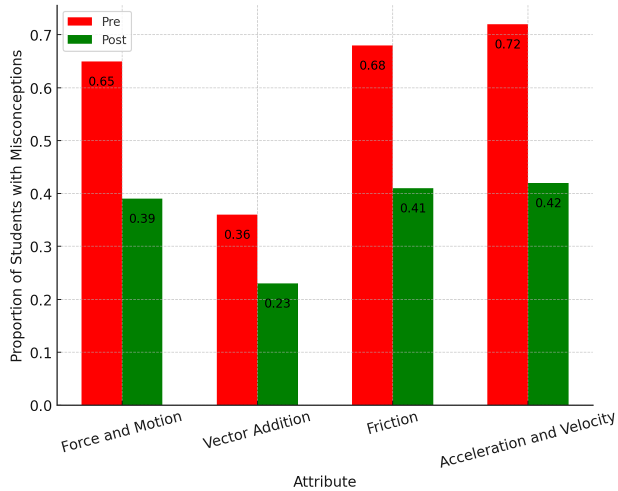

Figure 1: Proficiency proportion [33, 12]

misconception-free state demonstrate that targeted instructional efforts can be moderately effective at correcting misconceptions [33]. Transition probabilities help teachers identify persistent misconceptions and track learning progress, while transition posteriors provide student-specific insights for targeted interventions [11].

According to Table 6, for students who begin non-proficient ($Pre[0]$), the students of remain non-proficient ($Post[0]$) are high across all domains, with values such as $P = 0.974$ for *Force and Motion* and $P = 0.825$ for *Vector Addition*, indicating the persistence of misconceptions. Conversely, transitions from non-proficiency to proficiency ($Post[1]$) are low, such as $P = 0.026$ for *Force and Motion*. Students who start proficient ($Pre[1]$) have a moderate to high population rate of retaining their proficiency ($Post[1]$), with values like $P = 0.590$ for *Force and Motion* and $P = 0.566$ for *Acceleration and Velocity*. These patterns suggest that while misconceptions are difficult to overcome, maintaining proficiency is more stable, highlighting the need for interventions that specifically target transitions to proficiency (Table 6).

Table 6: Transition Probabilities Across Time Points

	Force and Motion		**Vector Addition**		**Friction**		**Acceleration and Velocity**	
Pre/post	**post [0]**	**post [1]**	**post [0]**	**post [1]**	**post [0]**	**post [1]**	**post [0]**	**Post [1]**
Pre [0]	0.974	0.026	0.825	0.175	0.972	0.028	0.968	0.032
Pre [1]	0.410	0.590	0.680	0.320	0.413	0.587	0.434	0.566

Note: Pre(0) = Students who begin as non-proficient; Pre(1) = Students who begin as proficient; Post(0) = Students who end as non-proficient; Post(1) = Students who end as proficient [33]

Table 7 represent the likelihood of correctly classifying each student into their actual misconception status across different physics concepts. The table provides insight into how well students are classified and their probability of being placed into an alternative misconception status. For Force and Motion (FM), students who were initially misconception-free ($P(X_{00})$)

were classified correctly with probabilities ranging from 0.450 (Student 2) to 0.867 (Student 1). However, Student 2 had a 23% chance ($P(X_{01}) = 0.230$) of being misclassified as having developed a misconception. Similarly, for Vector Addition (VA), the probability of remaining misconception-free varies, with Student 1 at 0.772 and Student 2 at 0.490, indicating classification stability differences among students. In Friction (FR) and Acceleration and Velocity (AV), classification into misconception states ($P(X_{11})$) remains relatively low, with Student 1 in Friction having only a 4.0% chance ($P(X_{11}) = 0.040$) of retaining a misconception, while Student 2 has a higher probability of 17%. These results suggest that while some students are correctly classified according to their misconception status, others have a substantial probability of being categorized differently [33].

Table 7: Transition Posteriors for Different Skills for each student

Skill	Transition	Student 1	Student 2	Student 3	Student 4	...	Student 1,529
FM	X_{00}	0.8676	0.4200	0.5100	0.8120	...	0.8450
FM	X_{01}	0.0451	0.2300	0.1300	0.0950	...	0.0600
FM	X_{10}	0.0321	0.1500	0.2100	0.0320	...	0.0380
FM	X_{11}	0.0552	0.2000	0.1500	0.0610	...	0.0570
VA	X_{00}	0.7720	0.4900	0.4650	0.7440	...	0.7600
VA	X_{01}	0.0950	0.1400	0.1550	0.1150	...	0.1070
VA	X_{10}	0.0650	0.1800	0.2050	0.0720	...	0.0580
VA	X_{11}	0.0680	0.1900	0.1750	0.0690	...	0.0750
FR	X_{00}	0.9020	0.4100	0.4850	0.8650	...	0.8930
FR	X_{01}	0.0270	0.2300	0.1650	0.0420	...	0.0330
FR	X_{10}	0.0250	0.2000	0.1800	0.0380	...	0.0260
FR	X_{11}	0.0460	0.1600	0.1700	0.0550	...	0.0480
AV	X_{00}	0.8900	0.4700	0.4500	0.8580	...	0.8800
AV	X_{01}	0.0410	0.1700	0.1900	0.0580	...	0.0450
AV	X_{10}	0.0290	0.2000	0.1950	0.0450	...	0.0360
AV	X_{11}	0.0400	0.1600	0.1650	0.0390	...	0.0390

Note:FM: Force and Motion; VA: Vector Addition; FR: Friction; AV: Acceleration and Velocity. X_{00} (No Mastery $\rightarrow$ No Mastery): The transition posterior probability that a student did not have mastery at Pre and still does not have mastery at Post; X_{01} (No Mastery $\rightarrow$ Mastery): The transition posterior probability that a student did not have mastery at Pre but gained mastery by Post; X_{10} (Mastery $\rightarrow$ No Mastery): The transition posterior probability that a student had mastery at Pre but lost mastery by Post; X_{11} (Mastery $\rightarrow$ Mastery): The transition posterior probability that a student had mastery at Pre and still has mastery at Post [33].

Discussion

Student conceptual understanding is constructed in both formal and informal educational settings by individuals actively trying to make sense of the world. These conceptions can align with or diverge from scientific consensus. However, these conceptions are both contextual and dynamic [1], evolving through a complex interplay of perception, assimilation, and accommodation

[26, 14, 16], which are shaped by the nature of the misconceptions and the educational interventions. For conceptual change to occur, instructors need to know students' conceptions and misconceptions. However, assessing student misconceptions is challenging, particularly for large enrollment introductory STEM courses where it may be impossible for instructors to examine every students incorrect answers on every assessment. The TDCM framework provides a nuanced lens to examine students' misconceptions as well as conceptual change [26, 14, 16].

As shown in Table 6, misconceptions are generally persistent, with low transition rates to scientifically accurate conceptions ($P(X_{01})$), particularly for Vector Addition, where only 32% of students initially holding misconceptions transitioned to correct conceptions. Additionally, a high probability of students remaining in misconception states ($P(X_{11})$) was observed, aligning with prior research on the persistence of alternative conceptions in physics education [11, 37, 38]. Despite these challenges, Force and Motion, Friction, and Acceleration and Velocity exhibit stronger retention of correct conceptions, with over 97% of students who started misconception-free remaining so at post-instruction. However, Vector Addition remains an exception, with a substantial proportion of students retaining misconceptions. These findings underscore the need for targeted interventions, such as embodied learning activities, to support conceptual change [39]. This aligns with prior research demonstrating that introductory physics course instructors sometimes erroneously believe that conceptual understanding is developed through solving calculation-based problems [8, 7].

The results from this study also emphasize the role of diagnostic precision in shaping the understanding of misconception dynamics. The holistic perspective, supported by global fit statistics and detailed item-level analyses, underscores the intricate relationship between diagnostic accuracy, instructional effectiveness, and the inherent characteristics of misconceptions. Model fit statistics indicate that the TDCM approach is appropriate for the data. However, the reliability scores indicate room for improvement in the individual items on the assessment. The moderately high-reliability scores for *Force and Motion* and Acceleration and Velocity suggest diagnostic adequacy to detect and track these misconceptions. However, there remains room for refinement to enhance precision, particularly for *Vector Addition* misconceptions due to lower reliability scores. These scores reflect the need for more refined assessment strategies to address the multidimensional nature of this domain. Future research needs to continue developing assessments to identify specific misconceptions as well as assess canonical conceptions to more effectively inform targeted interventions.

Conclusion

By the time students enroll in foundational STEM courses, they have developed explanations that have allowed them to successfully navigate the world around them. These alternative conceptions can be resistant to change. This is especially true in engineering and physics courses that engage students with everyday phenomena, or involve ambiguous counter evidence [27, 23]. This underscoring the need for innovative instructional strategies tailored to address their intuitive but incorrect appeal.

Conceptual Change Theories are a robust framework for understanding the development, persistence, and evolution of misconceptions in STEM education. Conceptual Change Theories emphasize that misconceptions are not random errors but coherent cognitive structures that provide learners with a sense of internal consistency, even when scientifically inaccurate. The theory identifies three key criteria for replacing misconceptions with accurate concepts: *intelligibility* (the new concept must make sense to the learner), *plausibility* (the new concept must seem believable within the learner's worldview), and *fruitfulness* (the new concept must be useful and applicable). Misconceptions persist when these criteria are not met, as learners are unlikely to abandon intuitive frameworks that appear logical and functional within their everyday experiences [40, 28]. However, instructors need to assess the scientific conceptions and the specific misconceptions held by students to plan instruction to motivate conceptual change.

By employing Q-matrices within the TDCM, our study tracks transitions in misconception states over time, providing a dynamic view of learning. This temporal perspective reinforces the theory's assertion that conceptual change is not a one-time event but a gradual process requiring repeated exposure to conflicting evidence and iterative refinement of mental models. For example, the frequent acquisition of misconceptions in Vector Addition, despite its low persistence, reflects the ongoing cognitive struggle learners face when reconciling new information with pre-existing schemas. These transitions, captured quantitatively in our results, demonstrate the robustness of Conceptual Change Theories in explaining how misconceptions evolve and resist correction over time [40, 28, 30].

The study illustrates that misconceptions persist and evolve during introductory physics courses for engineering majors. In addition, this study identifies misconceptions that appear to be more resilient to change. These misconceptions appear to be linked to abstract concepts involving spatial reasoning or everyday scenarios that differ from the idealized physics contexts, such as, projectile motion problems that ignore air resistance. While this study demonstrates a powerful method for examining misconception change, and potentially revealing pathways for improving educational practices and reducing the prevalence of these cognitive barriers in engineering education, a limitation of this study is that we did not systematically examine the instructional context or content for the students in the study. Because of this, we did not aim to examine the causes or instruction that led to conceptual change or persistence. Rather our goal is to examine the potential for the TDCM model to measure misconceptions and conceptual change. Future research should utilize more focused studies that incorporate the TDCM method for examining the precise nature of conceptual change in introductory STEM courses.

Acknowledgment

This study was funded, in part, by the National Science Foundation, grant number 2141847. We extend our appreciation to LASSO for their support in both collecting and sharing data for this research. The R code and the Q-matrix used in this study are available at the following GitHub repository: `https://github.com/amehrabi67/TDCM-ASEE-2025`. The student response data are not publicly posted; to request access to the anonymized dataset, please email Dr. Jason Morphew at jmorphew@purdue.edu. Additionally, the full list of Force Concept Inventory (FCI) questions is available at [41] and [32] under the FCI section.

References

[1] D. E. Brown, “Students’ conceptions as dynamically emergent structures,” *Science & Education*, vol. 23, pp. 1463–1483, 2014.

[2] D. Hammer, “Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions?” *Journal of the Learning Sciences*, vol. 5, no. 2, pp. 97–127, 1996.

[3] R. Even and T. Wallach, “What does it mean to interpret students’ talk and actions?” in *Proceedings of the 30th International Conference for the Psychology of Mathematics Education*, vol. 1. International Group for the Psychology of Mathematics Education, 2006, p. 355.

[4] S. Vosniadou, “Reframing the classical approach to conceptual change: Preconceptions, misconceptions and synthetic models,” *Second international handbook of science education*, pp. 119–130, 2012.

[5] R. Liu, R. Patel, and K. R. Koedinger, “Modeling common misconceptions in learning process data,” in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 2016, pp. 369–377.

[6] L. C. Kaczmarczyk, E. R. Petrick, J. P. East, and G. L. Herman, “Identifying student misconceptions of programming,” in *Proceedings of the 41st ACM technical symposium on Computer science education*, 2010, pp. 107–111.

[7] J. W. Morphew and J. P. Mestre, “Exploring the connection between problem solving and conceptual understanding in physics,” *Revista de Enseñanza de la Física*, vol. 30, pp. 75–85, 2018.

[8] W. Fakcharoenphol, J. W. Morphew, and J. P. Mestre, “Judgments of physics problem difficulty among experts and novices,” *Physical Review Physics Education Research*, vol. 11, no. 2, p. 020104, 2015.

[9] A. Mehrabi, J. W. Morphew, B. N. Araabi, N. Memarian, and H. Memarian, “Ai-enhanced decision-making for course modality preferences in higher engineering education during the post-covid-19 era,” *Information*, vol. 15, no. 10, p. 590, 2024.

[10] X. Li, Y. Li, and W. Wang, “Long-lasting conceptual change in science education: the role of u-shaped pattern of argumentative dialogue in collaborative argumentation,” *Science & Education*, vol. 32, no. 1, pp. 123–168, 2023.

[11] M. A. Schellman and M. J. Madison, “Estimating the reliability of skill transitions in longitudinal diagnostic classification models,” *Journal of Educational and Behavioral Statistics*, p. 10769986241256032, 2024.

[12] D. Facal, J. Guàrdia-Olmos, and O. Juncos-Rabadán, “Diagnostic transitions in mild cognitive impairment by use of simple markov models,” *International Journal of Geriatric Psychiatry*, vol. 30, no. 7, pp. 669–676, 2015.

[13] O. V. Forlenza, B. S. Diniz, P. V. Nunes, C. M. Memória, M. S. Yassuda, and W. F. Gattaz, “Diagnostic transitions in mild cognitive impairment subtypes,” *International Psychogeriatrics*, vol. 21, no. 6, pp. 1088–1095, 2009.

[14] S. O. Bada, “Constructivism learning theory: A paradigm for teaching and learning,” *IOSR Journal of Research Method in Education*, vol. 5, no. 6, pp. 66–70, 2015.

[15] C. T. Fosnot, *Constructivism: Theory, perspectives, and practice*. Teachers College Press, 2013.

[16] M. J. Reddy, “A case of frame conflict in our language about language,” *Metaphor and Thought*, pp. 164–201, 1979.

[17] J. Morphew, A. Mehrabi, B. Van Dusen, and J. Nissen, “Ai-based concept inventories: Using cognitive diagnostic computer adaptivetesting in lasso for classroom assessment,” 2024.

[18] W. F. Brewer and C. A. Chinn, “Scientists’ responses to anomalous data: Evidence from psychology, history, and philosophy of science,” in *Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Philosophy of Science Association, 1994, pp. 304–313.

[19] W. F. Brewer, “The theory ladenness of the mental processes used in the scientific enterprise: Evidence from cognitive psychology and the history of science,” 2012.

[20] A. Mehrabi and J. Morphew, “Board 73: Ai skills-based assessment tool for identifying partial and full-mastery within large engineering classrooms,” in *2024 ASEE Annual Conference & Exposition*, 2024.

[21] J. P. Smith III, A. A. DiSessa, and J. Roschelle, “Misconceptions reconceived: A constructivist analysis of knowledge in transition,” *The journal of the learning sciences*, vol. 3, no. 2, pp. 115–163, 1994.

[22] Y. Qian and J. Lehman, “Students’ misconceptions and other difficulties in introductory programming: A literature review,” *ACM Transactions on Computing Education (TOCE)*, vol. 18, no. 1, pp. 1–24, 2017.

[23] K. E. P. Morphew, Jason W., Q. King-Shepard, P. Kwon, R. Lin, T. J. Nokes-Malach, and J. P. Mestre, “Seeing and doing are not believing: Investigating when and how conceptual knowledge impinges on observation and recall of physical motion,” *Journal of Experimental Psychology: Applied*, vol. 27, no. 2, pp. 307–323, 2021.

[24] L. Watson, “Children’s misconceptions and conceptual change,” *Australasian Journal of Early Childhood*, vol. 22, no. 2, pp. 12–16, 1997.

[25] A. A. DiSessa, *A history of conceptual change research*, 2014.

[26] W. F. Brewer and L. Loshkey, *Top-down and bottom-up influences on observation: Evidence from cognitive psychology and the history of science*. In A. Raftopoulos (Ed.), Cognitive penetrability of perception: Attention, action, strategies, and bottom-up constraints (pp. 31–47). Nova Science, 2005.

[27] C. A. Chinn and B. A. Malhotra, “Children’s responses to anomalous scientific data: How is conceptual change impeded?” *Journal of Educational Psychology*, vol. 94, pp. 327–343, 2002.

[28] T. V. Goris and M. J. Dyrenfurth, “Concepts and misconceptions in engineering, technology and science. overview of research literature,” in *Proc. amer. soc. eng. educ. il/in sectional conf*, 2012.

[29] C. F. Herrmann-Abell and G. E. DeBoer, “Using rasch modeling and option probability curves to diagnose students’ misconceptions.” *Grantee Submission*, 2016.

[30] G. Goguadze, S. Sosnovsky, S. Isotani, and B. M. Mclaren, “Towards a bayesian student model for detecting decimal misconceptions,” in *Proceedings of the 19th International Conference on Computers in Education, Chiang Mai, Thailand*, 2011, pp. 6–8.

[31] Z. R. Mevarech, “A deep structure model of students’ statistical misconceptions,” *Educational studies in mathematics*, vol. 14, no. 4, pp. 415–429, 1983.

[32] LASSO Education, https://lassoeducation.org/, Retrieved January 13, 2025.

[33] M. J. Madison and M. E. Cotterell, *Introduction to the TDCM Package*, January 2024.

[34] M. J. Madison, S. Chung, J. Kim, and L. P. Bradshaw, “Approaches to estimating longitudinal diagnostic classification models,” *Behaviormetrika*, vol. 51, no. 1, pp. 7–19, 2024.

[35] A. C. George, A. Robitzsch, T. Kiefer, J. Groß, and A. Ünlü, “The r package cdm for cognitive diagnosis models,” *Journal of Statistical Software*, vol. 74, pp. 1–24, 2016.

[36] C.-Y. Chiu, H.-F. Köhn, and H.-M. Wu, “Fitting the reduced rum with mplus: A tutorial,” *International Journal of Testing*, vol. 16, no. 4, pp. 331–351, 2016.

[37] M. S. Steinberg, D. E. Brown, and J. Clement, “Genius is not immune to persistent misconceptions: conceptual difficulties impeding isaac newton and contemporary physics students,” *International Journal of Science Education*, vol. 12, no. 3, pp. 265–273, 1990.

[38] A. Mehrabi, J. W. Morphew, and B. S. Quezada, “Enhancing performance factor analysis through skill profile and item similarity integration via an attention mechanism of artificial intelligence,” in *Frontiers in Education*, vol. 9. Frontiers Media SA, 2024, p. 1454319.

[39] M. McColgan, J. W. Morphew, G. Hassel, J. Bennett, and M. C. Kelly, “Understanding magnetism concepts through augmented reality: A qualitative analysis,” in *American Society for Engineering Education 2024 Proceedings*. American Society for Engineering Education, 2024, p. 47004.

[40] G. Özdemir and D. B. Clark, “An overview of conceptual change theories,” *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 3, no. 4, pp. 351–361, 2007.

[41] D. Hestenes, M. Wells, G. Swackhamer, I. Halloun, R. Hake, and E. Mosca, “Cuestionario sobre el concepto de fuerza,” *The Physics Teacher*, 1995.