Academic Journal of Education
ISSN: 3080-9940 (Print), ISSN: 3080-9959 (Online) | Vol. 3, No. 2, 2026

Construction and Empirical Analysis of a Multi-Source Assessment System for Theoretical Mechanics under an Outcome-Based Education Framework

Shuang Zhou

School of Mechanical and Power Engineering, Yingkou Institute of Technology, Yingkou, Liaoning, 115014, China

Abstract: To overcome the limitations of using a single aggregate score to distinguish differences in knowledge mastery, engineering application, and learning-process performance in Theoretical Mechanics, this study developed a multi-source assessment framework based on outcome-based education (OBE). The framework links course outcomes, assessment components, evidence sources, and instructional improvement. Data were collected from two intact classes of Mechanical Design, Manufacturing and Automation in the first semester of the 2025-2026 academic year and included stage tests, homework, learning-management-system (LMS) quizzes, an integrated project, and item-level final-examination scores. The mean final-examination score was 51.17 ± 22.89. The cohort attainment levels for Course Outcomes 1, 2, and 3 were 0.513, 0.436, and 0.553, respectively. Mean stage-test and LMS-quiz scores showed moderate positive correlations with final-examination performance ($r = 0.461$ and 0.406, respectively; $p < 0.01$). Project scores were high with limited dispersion and were not significantly correlated with final-examination scores. The multi-source assessment identified kinematics and multi-step calculation tasks as the main learning bottlenecks and revealed insufficient discrimination in project scoring. Accordingly, tiered learning tasks, an analytic project rubric, and a data-informed diagnostic feedback cycle are proposed.

Keywords: Theoretical Mechanics; Outcome-based Education; Multi-source Assessment; Course Outcome Attainment; Project-based Learning.

1. Introduction

Theoretical Mechanics is a foundational course in many engineering programs and plays a critical role in professional training. The course is characterized by abstraction, rigor, logical coherence, and engineering relevance; consequently, student learning cannot be evaluated adequately through a single final examination. Traditional assessment compresses knowledge retention, modeling ability, computational rigor, and engineering application into one aggregate score. This makes it difficult to identify where students encounter difficulty or to translate assessment results into targeted instructional improvement.

Outcome-based education requires alignment among course outcomes, learning activities, and assessment evidence. Biggs [1] introduced constructive alignment in 1996 and emphasized that learning activities and assessment tasks should be designed around intended learning outcomes. Black and Wiliam [2] subsequently showed that formative assessment improves learning only when the resulting evidence is used to adjust subsequent teaching and learning. More recently, Luo et al. [3] established a whole-process assessment approach within an OBE framework, combined it with a three-dimensional evaluation-feedback process, and improved teaching quality through a modular teaching model organized around conceptual links and integrated instructional methods. Yi et al. [4] developed a process-based assessment system for Theoretical Mechanics under an OBE framework. They found similar trends in process-assessment and examination scores, indicating that process data can reflect students' learning status and support course outcome attainment.

Existing reforms in Theoretical Mechanics have introduced project-based learning, engineering cases, and blended instruction [5-7]. Yao et al. [8] developed a four-component instructional system integrating engineering cases, everyday-life contexts, software tools, and creative teaching aids. By combining process-based assessment with competency-attainment evaluation, they increased course outcome attainment from 0.72 to 0.88 over three teaching cycles. Wang et al. [9] explored reform in five areas: blended instruction, mind maps, contextualized examples and exercises, curriculum-based ideological education, and diversified assessment. Nevertheless, empirical studies on how multiple sources of evidence can jointly diagnose course outcomes remain limited.

Our previous work implemented project-based learning in Theoretical Mechanics and used authentic mechanical-system problems to connect statics, kinematics, and dynamics [10]. Building on that work, the present study integrates stage tests, homework, LMS quizzes, an integrated project, and the final examination within one evidence framework. It examines score distributions, interrelationships among assessment components, and their diagnostic value for course outcomes.

2. Multi-Source Assessment Framework and Statistical Methods

2.1. Course Outcomes and Assessment Principles

The course syllabus specifies three course outcomes (COs). CO1 requires students to master the fundamental concepts, principles, and methods of statics and to construct and analyze mechanical models. CO2 focuses on fundamental kinematics and its application to simple mechanical-engineering problems. CO3 requires students to master fundamental dynamics and develop logical reasoning, mathematical modeling, and the integrated application of knowledge to

problem solving.

The assessment framework follows three principles. First, alignment requires each source of evidence to correspond to an explicit course outcome. Second, multiple evidence sources capture both learning processes and end-of-course performance. Third, feedback utility requires assessment results to identify specific knowledge modules and competency components so that they can inform the next cycle of instructional improvement.

2.2. Assessment Components and Evidence Mapping

Table 1. Multi-Source Assessment Evidence Framework for Theoretical Mechanics

Assessment component	Recommended weight	Primary evidence	Focus of observation	Supported COs
Stage tests	15%	Three comprehensive tests	Fundamental concepts, computational methods, and stage-wise mastery	1, 2, 3
Homework	15%	Five assignments	Derivations, computational rigor, and graphical representation	1, 2
Integrated project	10%	Integrated design based on an engineering case	Force analysis, motion analysis, dynamics calculations, and engineering application	3
LMS quizzes	Process evidence	Thirteen knowledge-point quizzes	Learning progress, immediate mastery, and early warning	1, 2, 3
Final examination	60%	Item-level scores mapped to course outcomes	Knowledge structure, diagrams, integrated calculations, and standardized presentation	1, 2, 3

The overall course grade consisted of 40% process-based assessment and 60% final examination. Table 1 summarizes the components of the multi-source assessment. Process evidence included three stage tests, five homework assignments, 13 LMS quizzes, and one integrated project. The integrated project was based on an engineering case involving a mechanical system. Students were required to analyze the forces acting on individual components and the overall system; calculate friction, constraint forces, and acceleration; draw the relevant free-body diagrams; and prepare a comprehensive design report. The project therefore assessed students' ability to integrate statics, kinematics, and dynamics in engineering problem solving. End-of-course data comprised the total final-examination score, each student's item-level scores mapped to the course outcomes, and the overall course grade. In addition to the overall grade, the original score from each evidence source was retained to prevent a weighted aggregate score from obscuring differences in learning.

2.3. Participants and Statistical Methods

The participants were 59 students enrolled in Theoretical Mechanics during the first semester of the 2025-2026 academic year. They were drawn from two intact classes, Classes 2403 and 2404, in the Mechanical Design, Manufacturing and Automation program.

For each assessment variable, the sample size, mean, standard deviation, median, and pass rate were calculated. Cohort attainment for each course outcome was defined as the mean student score on the items mapped to that outcome divided by the maximum score available for those items. The proportion of students with individual attainment of at least 0.60 was also calculated as a diagnostic indicator rather than as a new grading rule. Pearson correlation analysis was used to examine associations between process evidence and final-examination scores. The two intact classes were compared using Welch's independent-samples t test, with Cohen's d reported as the effect size. Statistical significance was set at 0.05. All correlations were interpreted as associations and not as evidence of causality.

3. Results and Analysis

3.1. Score Characteristics Across Assessment Components

Table 2. Descriptive Statistics for the Assessment Evidence

Assessment evidence	n	Mean	SD	Median	Minimum	Maximum	Pass rate
Mean stage-test score	63	83.89	4.88	84.00	71.67	94.67	100.0%
Homework	63	81.43	12.57	84.60	33.80	98.00	95.2%
Integrated project	61	87.92	4.95	89.00	72.00	97.00	100.0%
LMS quizzes	63	83.09	6.03	83.37	55.11	94.25	98.4%
Final examination	63	51.17	22.89	54.00	5.00	96.00	38.1%
Overall course grade	63	55.52	24.94	64.00	5.00	94.00	60.3%

Note: A score of 60 or above was classified as a pass

Table 2 presents the mean, standard deviation, median, minimum, maximum, and pass rate for each source of assessment evidence. Process-based scores were generally high, whereas final-examination scores had the lowest mean and the greatest dispersion. The integrated project score was 87.92 ± 4.95, and all students scored at least 60. The final-examination score was 51.17 ± 22.89, with a pass rate of 38.1%. These distributions suggest that process-based tasks supported broad participation, whereas the concentrated project scores provided limited discrimination among performance levels. By contrast, the final examination more clearly differentiated students' ability to construct models independently and complete multi-step calculations under time constraints.

3.2. Course Outcome Attainment

According to the examination blueprint, final-examination items were mapped to CO1, CO2, and CO3, with maximum scores of 36, 34, and 30, respectively. Table 3 reports the mean score, standard deviation, cohort attainment, and individual attainment for each course outcome. Figure 1 compares cohort attainment with the diagnostic threshold of 0.60. Cohort attainment was below 0.60 for all three outcomes. CO2 was the weakest outcome, with cohort attainment of 0.436, a mean score of 14.83, and only 25.4% of students attaining at least 0.60. CO3 had the highest mean attainment, but its standard deviation was 9.98 points, indicating substantial variation in performance on integrated dynamics problems.

Table 3. Final-Examination Course Outcome Attainment

Course outcome	Max. score	Mean score	SD	Cohort attainment	Students attaining >=0.60
CO1 Statics and mechanical modeling	36	18.46	8.16	0.513	33.9%
CO2 Kinematics and engineering analysis	34	14.83	7.77	0.436	25.4%
CO3 Dynamics and integrated application	30	16.59	9.98	0.553	45.8%

Note: Cohort attainment equals the mean student score for a course outcome divided by the maximum score for the corresponding examination items; n = 59.

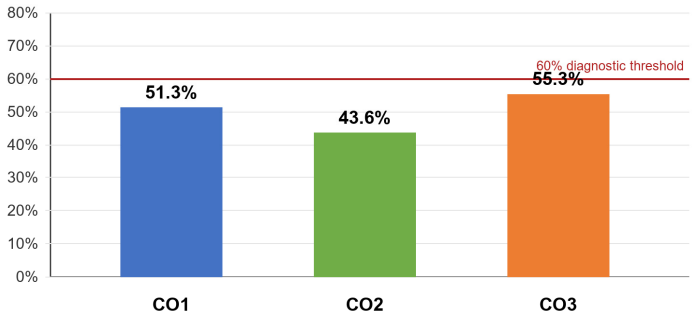


Figure 1. Course Outcome Attainment in the Final Examination

Table 4. Item-Level Performance in the Final Examination

Item category	Maximum score	Mean score	Score rate	Zero-score rate	Full-score rate
CO1 foundational item	6	3.66	61.0%	0.0%	23.7%
CO1 diagram item	10	5.68	56.8%	3.4%	11.9%
CO1 calculation item	20	9.12	45.6%	15.3%	8.5%
CO2 foundational item	4	2.92	72.9%	5.1%	50.8%
CO2 integrated item	30	11.92	39.7%	6.8%	3.4%
CO3 foundational item	10	6.81	68.1%	5.1%	27.1%
CO3 integrated item	20	9.78	48.9%	28.8%	20.3%

The item-level results further show that performance on foundational tasks was consistently higher than performance on integrated calculation tasks. The score rate for the CO2 foundational item was 72.9%, whereas the integrated item

mapped to the same outcome had a score rate of only 39.7%. The score rate for the CO1 calculation item was 45.6%, and the zero-score rate for the CO3 integrated item was 28.8%. These findings indicate that students' main difficulty was not recalling isolated concepts but abstracting a model from a kinematic context, selecting an appropriate theorem, and completing a sequence of calculations.

3.3. Associations Between Process Evidence and Final-Examination Performance

Table 5. Correlations Between Process-Based Scores and Final-Examination Scores

Variable	n	Pearson r	p value	Interpretation
Mean of three stage tests	63	0.461	<0.001	Moderate positive correlation
LMS quizzes	63	0.406	0.001	Moderate positive correlation
Homework	63	0.200	0.116	Not significant
Integrated project	61	-0.052	0.692	Not significant

Note: p values are based on two-sided tests. Correlation coefficients represent associations and do not imply causation

Mean stage-test and LMS-quiz scores were significantly and moderately positively correlated with final-examination scores. These two sources of evidence can therefore support mid-semester identification of students at risk of poor performance. Homework scores were only weakly associated with final-examination scores, potentially because of correction opportunities, external assistance, and variation in task difficulty. Project scores were not correlated with final-examination scores; however, this finding does not indicate that project-based learning was ineffective. The project and the closed-book examination assessed different capabilities, and the restricted dispersion of project scores (SD = 4.95) may have attenuated the correlation. The more important implication is that the project rubric should discriminate more clearly among modeling quality, computational procedure, and engineering interpretation.

3.4. Between-Class Differences

The mean final-examination scores for Classes 2403 and 2404 were 45.03 ± 21.22 and 57.13 ± 23.19, respectively. Class 2404 scored 12.09 points higher, a statistically significant difference ($t = -2.161$, $p = 0.035$, Cohen's $d = 0.54$). At the outcome level, the mean attainment of CO1 was 0.601 in Class 2404 and 0.421 in Class 2403 ($p = 0.002$, $d = 0.86$). The between-class difference was not significant for CO2 ($p = 0.181$), whereas the difference for CO3 approached significance ($p = 0.075$). Because students were not randomly assigned to classes, these differences may also reflect prior preparation, classroom climate, attendance, and other factors. They are therefore not interpreted as causal effects of any single instructional intervention. Nevertheless, the results show that identical pacing and tasks do not ensure comparable learning outcomes across classes and that tiered support should be informed by class-level and individual-level evidence.

4. Instructional Improvement Driven by Assessment Results

4.1. Establishing an Evidence Chain for Course Outcomes

In subsequent teaching cycles, stage tests, homework, the integrated project, and final-examination items should be consistently mapped to the three course outcomes. Feedback after each assessment should include not only an aggregate score but also class-level outcome attainment and the distribution of error types. For components of CO2 that remain below the diagnostic threshold of 0.60, the outcome should be decomposed into teachable subcompetencies, including reference-frame selection, relationships among the three types of motion, application of the composition theorems, and solution of geometric constraints. This mapping would allow each assessment finding to correspond directly to a subsequent instructional action.

4.2. Tiered Support for Kinematics and Complex Calculations

The score rate for the CO2 integrated item was only 39.7%, indicating a discontinuity when students moved from isolated knowledge components to integrated application. Subsequent instruction can organize tasks into three levels. Foundational tasks should reinforce reference frames, the three types of motion, and basic formulas. Application tasks should require vector diagrams, identification of known and unknown quantities, and stepwise calculations. Extension tasks should introduce parameter variation and comparisons among solution methods. Students who perform poorly on consecutive LMS quizzes or stage tests should complete short diagnostic exercises and receive targeted feedback before beginning the integrated project.

4.3. Improving the Discriminatory Power of Project Scoring

The high mean and narrow distribution of project scores indicate that the project supported completion of integrated learning tasks but that the rubric did not sufficiently distinguish performance levels. The rubric should therefore specify five dimensions: mechanical-model construction; force and motion analysis; equation formulation and computation; result verification and engineering interpretation; and report quality. Each dimension should include observable performance-level descriptors. Before scoring, project submissions representing high, medium, and low performance should be selected as anchor samples to calibrate the scoring criteria. Two instructors should then independently score a subset of the submissions to evaluate inter-rater consistency.

4.4. Closing the Assessment and Feedback Loop

Each teaching cycle can follow a five-step process: data collection, attainment calculation, diagnosis of weak components, task adjustment, and reassessment. In this study, stage tests and LMS quizzes provided useful early-warning signals. Instructors can use item-level evidence at mid-semester to assign remedial learning rather than waiting until final grades are available. Subsequent assessments should retain a consistent mapping between items and course outcomes so that changes in attainment can be compared across teaching cycles.

5. Conclusion and Limitations

This study constructed a multi-source assessment framework for Theoretical Mechanics using authentic teaching records. The framework maps stage tests, homework, LMS quizzes, an integrated project, and the final examination to course outcomes, thereby converting a single aggregate score into a diagnostic body of evidence. The results show that stage tests and LMS quizzes provide mid-semester signals of final learning performance. CO2 and the integrated calculation items across outcomes were the principal areas of weakness. Project scores provided distinct evidence of engineering-application capability, but the scoring rubric requires greater discriminatory power and inter-rater consistency.

This study has several limitations. First, the sample was drawn from one cohort in a single program at one institution, which limits generalizability. Second, the study used data from intact classes without random assignment or an independent control group; consequently, score differences cannot be attributed to a single instructional intervention. Future studies should use data from multiple successive cohorts and further evaluate rubric reliability, inter-rater consistency, and changes in attainment before and after instructional improvement.

Acknowledgments

Funding: Education and Teaching Reform Project of Yingkou Institute of Technology (JG202426).

References

[1] Biggs, J. (1996). Enhancing teaching through constructive alignment. Higher Education, 32(3), 347–364. https://doi.org/10.1007/BF00139671.

[2] Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi.org/10.1080/0969595980050102.

[3] Luo, L. B., Wu, Y. Q., & Zhao, L. (2025). Reform of Theoretical Mechanics laboratory teaching for mechatronics majors based on OBE. China Machinery, (36), 155–158. (in Chinese)

[4] Yi, H. A., Qin, J. H., & Liu, J. D. (2023). Construction of a process-based assessment system for Theoretical Mechanics based on OBE. Western China Quality Education, 9(2), 6–10. (in Chinese)

[5] Yang, W. M., Zhao, H. M., & Zhao, Y. C. (2017). Exploration and practice of project-based teaching reform in Theoretical Mechanics under the Double First-Class initiative. New Curriculum Research, (11), 76–77. (in Chinese)

[6] Meng, H., Duan, S. S., & Wang, J. J. (2025). Teaching reform and practice in engineering foundation courses under the emerging engineering education framework: A case study of Theoretical Mechanics. China Metalforming Equipment & Manufacturing Technology, 60(4), 170–174. (in Chinese)

[7] Zhang, T. Y., Wang, Y., Xiao, S. Y., et al. (2025). A blended teaching model for local application-oriented universities: A case study of Mechanics of Materials. The Theory and Practice of Innovation and Entrepreneurship, 8(10), 116–118. (in Chinese)

[8] Yao, X. G., Zhao, Q. R., Ma, J. S., et al. (2026). Reform of Theoretical Mechanics teaching for engineering education accreditation: From knowledge-centered to competency-centered classrooms. Shanxi Youth, (17), 139–141. (in Chinese)

[9] Wang, M., Chen, X. F., & Wu, C. Z. (2026). Teaching reform in Theoretical Mechanics under higher-education modernization. Journal of Higher Education, 12(3), 140–143. (in Chinese)

[10] Zhou, S. (2025). Integrating project-based learning into Theoretical Mechanics. Journal of Education, 1(9). (in Chinese)