Exploration and Analysis of Reform in Assessment Methods for Mechanical Engineering Courses

Youwei Wan*

Hainan Vocational University of Science and Technology, Haikou, 571126, China
**Corresponding author: wanyouwei15@163.com*

Abstract: *Mechanical engineering course assessment methods have long been dominated by closed-book written examinations, whose unitary form imposes constraints on the cultivation of engineering competence, resulting in a low correlation between course grades and students' genuine professional competence. In response to the structural deficiencies in the assessment dimensions of core mechanical engineering courses and the misalignment between the traditional assessment system and the attainment of course objectives, this paper proposes three reform paths: the construction of a formative assessment system based on ability stratification; the systematic embedding mechanism of project-driven assessment modules; and the dynamic redistribution of process weights. To ensure the effective operation of the above paths, this paper designs a strategy for aligning assessment standards with technical specifications of the mechanical industry, an organizational framework for multi-stakeholder collaborative assessment, and a teaching feedback closed-loop method driven by assessment data. By integrating formative assessment and project-based assessment elements and reconstructing assessment weight adjustment rules, the assessment system is enabled with self-adaptive capability to teaching difficulties and student performance, thereby enhancing the engineering equivalence and diagnostic accuracy of mechanical engineering course evaluation.*

Keywords: *Mechanical engineering courses, Assessment methods, Formative assessment, Ability stratification, Dynamic weights.*

1. Introduction

The teaching objectives of mechanical engineering courses emphasize the development of comprehensive engineering capabilities such as system design, process planning, and fault diagnosis. However, the current assessment methods still predominantly rely on final closed-book written examinations, which focus on the memorization and reproduction of discrete knowledge points. This evaluative orientation leads students to gravitate toward mechanical training with typical exercises during exam preparation, resulting in insufficient analytical and decision-making abilities when confronting unstructured engineering problems. The mismatch between assessment formats and the demands of engineering competence, coupled with the structural deficiencies in evaluation dimensions and the miscalculation of course objective attainment, renders the traditional system incapable of accurately reflecting learning outcomes and guiding instructional improvement. Therefore, it is evidently necessary to explore reform paths for the assessment methods of mechanical engineering courses. Such reform requires the introduction of strategies such as ability stratification, project-driven approaches, and dynamic weights while maintaining evaluation objectivity, and the construction of operational support mechanisms from three aspects — standard alignment, multi-stakeholder collaboration, and data feedback—so as to achieve substantive support of the evaluation system for engineering education objectives.

1. Review of the Current Assessment Status of Mechanical Engineering Courses and Attribution of Problems

1.1 The Constraint of Unitary Assessment Formats on the Cultivation of Engineering Competence

The traditional assessment of mechanical engineering courses has long been dominated by closed-book written examinations, with question types concentrated on the restatement of principles, derivation of formulas, and solution of simplified computational problems. This form emphasizes the memorization and reproduction of discrete knowledge points, whereas comprehensive abilities such as system design, process planning, and fault diagnosis required in mechanical engineering are difficult to measure effectively through standardized examination papers. During exam preparation, students tend to invest their efforts in mechanical training with typical exercises, resulting in relatively low analytical and decision-making abilities when confronting unstructured engineering problems. The mismatch between assessment formats and the demands of engineering competence leads to a low correlation between course grades and students' genuine professional competence.

From the perspective of the intrinsic logic of engineering education, mechanical engineering competence comprises multiple inseparable dimensions, including scheme conception, modeling and analysis, manufacturing process evaluation, and technical documentation writing. A unitary paper-and-pencil test fails to capture students' operational performance and judgment quality in links such as computer-aided design, finite element simulation, or assembly process optimization. Meanwhile, this assessment format neglects the observation of implicit abilities such as teamwork, engineering ethical judgment, and techno-economic trade-off. Long-term reliance on a unitary format tends to induce teaching behaviors to shift toward an exam-oriented direction, weaken the due weight of experiments, course projects, and project-based learning components in the evaluation system, and consequently impose a systematic constraint on the structure of students' engineering competence[1].

1.2 Structural Deficiencies in the Assessment Dimensions of Core Mechanical Engineering Courses

Core mechanical courses such as Mechanical Principles, Mechanical Design, and Fundamentals of Mechanical Manufacturing Technology encompass ability objectives covering multiple levels, including kinematic analysis, strength verification, tolerance and fit, and process planning formulation. However, the current assessment system typically extracts only those components that are quantifiable and standardizable as evaluation criteria, such as the calculation of degrees of freedom of mechanisms, selection of gear parameters, or time quota for simple processes. For higher-order ability dimensions such as scheme innovativeness, structural manufacturability analysis, and trade-off decision-making under multiple constraints, there is a lack of corresponding assessment tools and scoring criteria. This selective evaluation leads to the substantial suspension of the "analysis," "synthesis," and "evaluation" levels within the course objectives.

The deficiency in assessment dimensions is also manifested in the artificial separation between the quality of the design process and the quality of the final outcome. In most mechanical design courses, only the final engineering drawings or calculation specifications are taken as grading objects, while observation points for the iterative paths of design schemes, the rationale for parameter modifications, and the coherence of design logic are not established. The trial-and-error methods or mechanical application of empirical formulas adopted by students in their design work cannot be identified or corrected due to the absence of a process-oriented assessment mechanism. Moreover, interdisciplinary competencies such as control logic, sensor selection, and coordination with embedded software in modern mechanical systems remain almost entirely unaddressed within traditional assessment dimensions. These structural deficiencies not only weaken the diagnostic function of the assessment but also leave course improvement efforts without precise data support.

1.3 The Misalignment Between the Traditional Assessment System and the Attainment of Course Objectives

The calculation of course objective attainment depends on the mapping relationship between assessment results and preset ability indicators. In the traditional assessment system, several questions in one examination paper often simultaneously target multiple objectives at different levels, and the weighting of each objective has not been systematically calibrated. For example, a comprehensive problem involving gear strength calculation may simultaneously assess three abilities — formula memorization, parameter table lookup, and numerical computation—but the scoring only provides a total score without separating the actual attainment level of each individual ability. This coarse-grained evaluation approach confines attainment analysis to the level of total score statistics, making it difficult to pinpoint specific ability deficiencies.

The deeper misalignment is reflected in the disconnection between assessment content and the actual competence requirements of engineering practice. The teaching objectives of mechanical engineering courses typically include "the ability to design a simple mechanical transmission device," yet the design problems in traditional examination papers are mostly simplified and idealized working conditions, neglecting essential links in real-world design such as lubricant selection, sealing structure configuration, and thermal balance verification. The attainment measured by the assessment system reflects students' problem-solving ability under simplified models rather than their engineering design capability oriented toward real-world constraints. This systemic misalignment leads to an artificially inflated phenomenon in course objective attainment, conceals substantive deficiencies in the teaching process, and causes subsequent course improvement measures to lose their precise targeting[2].

2. Core Reform Paths for Assessment Methods of Mechanical Engineering Courses

2.1 Construction of a Formative Assessment System Based on Ability Stratification

The construction of the formative assessment system takes the mechanical engineering competency framework as the underlying structure, categorizing ability objectives into four progressive levels according to cognitive complexity: memory recall, principle application, comprehensive analysis, and engineering innovation. Each level corresponds to differentiated observation indicators and assessment tools: the memory recall level employs online self-assessment and concept mapping; the principle application level assigns small-scale computational or selection tasks; the comprehensive analysis level introduces chapter-by-chapter review of design calculation documents; and the engineering innovation level involves open-ended scheme demonstration. By embedding short-cycle assessment nodes within teaching units or weeks, time-series data of students' ability development are obtained, capturing individual students' pass rates and error patterns at different ability levels, thereby precisely identifying stagnation points at specific levels. This stratified framework also provides ability-level anchors for subsequent dynamic weight adjustments, enabling the assessment results to be directly mapped to the decomposed items of course objectives.

A key operation of this system lies in the hierarchical coding of assessment rubrics. For the shaft system design task in the Mechanical Design course, the low-level assessment focuses on the correctness of standard component selection and the completeness of dimensioning; the intermediate-level assessment concentrates on the rationality analysis of positioning and fixing methods; and the high-level assessment evaluates the capability of scheme optimization under spatial constraints and thermal deformation conditions, with an additional examination of the completeness of the logic chain in the design specification. Each level is set with an independent passing threshold. Students who fail to meet the low-level requirements are not permitted to proceed to the higher-level assessment, while their number of retests and typical error patterns are also recorded. This asymmetric progression mechanism avoids the masking effect of traditional total score weighting on ability deficiencies, enables formative assessment to genuinely serve ability diagnosis and differentiated teaching intervention, and provides fine-grained data sources for the calculation of course ability attainment.

2.2 Systematic Embedding Mechanism of Project-Driven Assessment Modules

The project-driven assessment module takes the entire process of mechanical product development as its organizing thread, integrating the assessment nodes originally scattered across various chapters of the course into a coherent technical task chain. A typical module can be set as "design and verification of a reducer transmission scheme under specified operating conditions," which simultaneously covers multiple ability domains such as kinematic parameter allocation, gear strength verification, shafting structural layout, and lubrication method determination. The module incorporates several mandatory delivery nodes, including the scheme review report, the first draft of the calculation specification, the three-dimensional model, and the two-dimensional engineering drawing, with each node corresponding to explicit scoring criteria and normative requirements. The module score replaces the scores of several isolated problems in the traditional examination paper, forming an integrative measurement of comprehensive design ability, while simultaneously examining students' systematic thinking in solving coupled sub-problems through the dependency relationships among the nodes[3].

Systematic embedding requires the module to be fixed in both time and space within the course teaching cycle and to form a temporal coupling with the lecture content. Taking the Fundamentals of Mechanical Manufacturing Technology course as an example, the project module can be set as "process planning and fixture scheme design for a given part." The module is initiated after the theoretical instruction on process datum selection and is completed by the end of the chapter on fixture positioning error analysis. During the module progression, multiple intermediate checkpoints are established, corresponding to subtasks such as blank selection, operation sequencing, allowance determination, and positioning scheme, with each checkpoint having a minimum passing standard. By decomposing the assessment module into independently evaluable subcomponents, the cognitive load of a single evaluation is reduced, while teachers are provided with a technical window to observe the completeness of students' design decision chains, and it also supports the separate tracing and remediation of specific sub-abilities within the module.

2.3 Dynamic Redistribution of Process Weights in Mechanical Design and Manufacturing Courses

The dynamic redistribution of process weights aims to break through the rigid structure of fixed proportional allocation (e.g., 30% for regular performance and 70% for final examination) and to establish weight adjustment rules that match the characteristic stages of course progression. In the early stage of the course (the period of concept introduction and basic skills training), the coefficients of in-class quizzes and unit assignments in the final overall grade are increased, while the preset weight of the later concentrated assessment is reduced. As the course enters the comprehensive design stage, the weights are gradually shifted toward project modules and design documents, and in the final phase, the cumulative impact of the low-weight nodes in the earlier stages is corrected through weighted backward adjustment. This mechanism adopts a sliding window weighting method, which dynamically adjusts the contribution proportion of each assessment based on the coefficient of variation of students' most recent three formative assessments, while simultaneously introducing a group difficulty response factor so that the weight allocation reflects the evolutionary trend of collective difficulties during the teaching process[4].

In specific implementation, a weight response function is established for mechanical design and manufacturing courses. The weight of a given assessment node is set to be positively correlated with its corresponding ability-level coefficient and the score dispersion of the student cohort at that node, and is multiplied by the node's prior discrimination parameter. For example, in the CNC machining technology course, the weight of the manual programming node is higher than that of the G-code simulation node, because the former involves more algorithmic logic and constraint judgments and is more critical for subsequent learning. After identifying collective difficulties of the student cohort in the tool path optimization stage in the middle of the course, the system automatically raises the upper weight limit of subsequent related assessment nodes, simultaneously compresses the weights of the already-passed high-score nodes, and reallocates the retest weight coefficients for the nodes that have not met the standards. The dynamic redistribution strategy endows the assessment system with self-adaptive regulation capability in response to teaching difficulties and student performance, avoids evaluation distortion caused by static weights, and provides differentiated final grade composition schemes for students with different learning paths.

3. Operational Support for the Reform of Assessment Methods in Mechanical Engineering Courses

3.1 Strategy for Aligning Assessment Standards with Technical Specifications of the Mechanical Industry

The formulation of assessment standards should refer to the current technical standard system of the mechanical industry, covering multiple dimensions such as drawing annotation specifications (GB/T 1182 for geometrical tolerancing), material selection criteria (GB/T 3077 for alloy structural steels), process documentation requirements (JB/T 9169 for process planning format), and numerical control programming standards (GB/T 18726 for NC machining codes). The measurable parameters in the above specifications are to be converted into deduction items or grade thresholds in the scoring rubrics. For example, in the assessment of the Mechanical Drawing course, the evaluation of view representation schemes no longer relies solely on graphic completeness, but introduces quantitative indicators of specification compliance, such as section line spacing, dimension chain closure, and surface roughness notation placement. The core of the alignment strategy lies in extracting the mandatory clauses and recommended clauses from industry specifications, corresponding respectively to the mandatory items and bonus items in the assessment system, thereby forming an evaluation scale with engineering equivalence.

To achieve systematic alignment, a bidirectional mapping matrix between assessment standards and industry specifications should be constructed. The row vectors of the matrix correspond to the assessment nodes of the course, and the column vectors correspond to the involved technical specification items, with each matrix element annotated with the application level of the specification (recognition, interpretation, or execution). Taking the assessment of process planning formulation in the Fundamentals of Mechanical Manufacturing Technology course as an example, the location datum annotation on the operation sketch must comply with GB/T 24737, the tolerance grade of process dimensions must match GB/T 1800, and the selection of cutting parameters must refer to the recommended ranges in the Machining Process Handbook. The assessment scoring sheet directly cites the specification numbers and conformance criteria, and the process documents submitted by students are subject to compliance verification. This strategy directly aligns the assessment results with students' mastery of the technical language of the industry, thereby reducing the semantic conversion loss between campus evaluation and the engineering practice environment.

3.2 Organizational Structure and Operational Logic of Multi-Stakeholder Collaborative Assessment

The organizational structure of multi-stakeholder collaborative assessment is composed of the lead course instructor, lab instructors, graduate teaching assistants, and student representatives, forming a clear division of evaluation functions at different levels. The lead course instructor is responsible for designing the overall framework of the assessment rubric and conducting the final review of high-weight nodes. The lab instructors undertake process observation and raw data recording in performance-based assessments. The graduate teaching assistants execute standardized tasks, such as preliminary compliance checks on drawing assignments and format verification of calculation documents. The student representatives participate in rubric piloting and scoring consistency verification during peer evaluation sessions. Information transfer among the various stakeholders adopts an evaluation work order transfer mechanism, in which each work order specifies the evaluation object, deadline, delivery standards, and dispute resolution pathway, ensuring that the division of responsibilities is clear and traceable.

The operational logic is built upon the principles of role interlocking and process separation. Role interlocking requires that the scoring for the same assessment node be independently completed by at least two parties, and if the discrepancy exceeds a preset tolerance threshold, a third-party review is initiated. Taking the project defense in the Mechanical Design course as an example, the lead instructor assesses the principle feasibility of the design scheme, the lab instructor verifies the prototype assembly and motion functions, the teaching assistant reviews the document format and drawing completeness, and three students from the same group anonymously evaluate each other's contribution and collaboration performance. Process separation means that evaluation actions and teaching assistance actions are performed by different parties, avoiding the conflict of interest in which the instructor both trains and scores. Under this structure, the scoring matrix can simultaneously output individual ability scores and a team collaboration coefficient, with the latter computed based on the statistical convergence of score deviations among the various parties, rather than on subjective impressions[5].

3.3 Data-Driven Teaching Feedback Closed-Loop Design Method

The starting point of the teaching feedback closed loop is the raw scoring data generated from each assessment node, which, after structured cleaning, are stored in the course ability database. Each data record carries ability-attribute labels (e.g., gear strength calculation ability, process planning formulation ability) and a timestamp, supporting both cross-sectional (inter-class comparison) and longitudinal (individual growth tracking) analyses. Item Response Theory is employed to calibrate the parameters of assessment items, obtaining the discrimination and difficulty parameters for each item, based on which inefficient or invalid assessment entries are screened out. For the score distribution of the student cohort on specific ability labels, ability radar charts and heat maps are drawn to locate the knowledge and skill areas with high error rates. For example, if a class scores significantly lower on the label "selection of axial positioning methods in shaft system structural design" than on other labels, a specialized diagnosis for that ability point is triggered.

The terminal end of the closed loop consists of the generation of intervention signals directed at both teaching behaviors and learning behaviors. The system automatically produces two types of feedback based on preset deviation thresholds: teacher-oriented course adjustment recommendations, including slowing down classroom lectures on weak ability points, adding targeted exercises, or adjusting constraints for subsequent project modules; and student-oriented personalized learning path recommendations, covering indices of assessment nodes requiring retesting, supplementary learning material links, and optional module variants. The feedback signals are presented in the form of a data dashboard on the teaching management interface, accompanied by actionable adjustment parameters (e.g., exercise difficulty coefficients, retest window length). This closed-loop design enables automated mapping from assessment results to pedagogical decisions, shortens the feedback lag inherent in traditional evaluation models, and provides a quantifiable basis for continuous course iteration.

Conclusion

The reform of assessment methods for mechanical engineering courses needs to be advanced synergistically from three dimensions: ability diagnosis, task integration, and weight adjustment. The formative assessment system based on ability stratification can identify individuals' stagnation points at each level of memory recall, principle application, comprehensive analysis, and engineering innovation. The project-driven assessment modules achieve integrative measurement of comprehensive design ability through a coherent technical task chain. The dynamic redistribution of process weights adaptively adjusts the contribution ratio of each assessment node according to the course progression stage and student performance. The effective operation of the above paths relies on the bidirectional mapping between assessment standards and industry technical specifications, the role-interlocking mechanism of multi-stakeholder collaborative assessment, and the data-driven teaching feedback closed loop. Future directions may focus on developing intelligent assessment platforms to support real-time weight computation and feedback delivery, exploring data fusion methods for continuous ability tracking across courses and semesters, and further investigating operationalized representation techniques for non-technical abilities (such as engineering ethical judgment and collaborative communication) within the assessment system.

References

[1] Song D. L., et al. (2024). "Research on Assessment Methods of Mechanical Design Course Based on OBE Concept." University Education, (15), 61 – 65.

[2] Luo Y., & Luo Y. L. (2022). "Discussion on Project-Driven Teaching Method and Corresponding Course Assessment Methods — Taking 'Computer Simulation Technology' Course of Mechanical Engineering Major as an Example." Academics and Practice, (03), 215 – 218.

[3] Tang C. Z., & Liu H. W. (2021). "Practice and Exploration of Reform of Course Assessment Methods for Automotive Majors — Taking 'Fundamentals of Mechanical Design' as an Example." Automotive Practical Technology, 46(16), 210 – 212.

[4] Yu B., et al. (2021). "Exploration of Teaching Mode Reform for Postgraduate Professional Courses in Mechanical Engineering — Taking 'Modern Servo Control Technology' Course as an Example." Science Consultation (Educational Research), (15), 44 – 45.

[5] Yang N. N., & Lun F. Y. (2020). "Research and Practice on Assessment Methods of Mechanical Majors under the Background of Educational Informatization." Southern Agricultural Machinery, 51(15), 182 – 183.