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: *For a long time, the evaluation method for the mechanical engineering course has been closed-book written tests; their unitary nature has limited the cultivation of students' engineering ability, and as a result, there is a weak correlation between course grades and students' actual professional competence. Due to the structural flaws in the assessment dimensions of core mechanical engineering courses and the lack of alignment between the traditional assessment system and the achievement of course objectives, this paper proposes three reform directions: the construction of a formative assessment system based on ability stratification, the systematic embedding mechanism of project-oriented assessment modules, and the dynamic redistribution of process weights. To ensure the good operation of the above paths, this paper puts forward a plan to align assessment standards with the technical specifications of the mechanical industry, an organisational structure for multi-stakeholder collaborative assessment, and a teaching feedback closed-loop method based on assessment data. Combine formative assessment and project-based assessment, adjust the proportion of the two types of assessments according to the actual circumstances, build an independent assessment system that can be changed in light of changes in teaching difficulties and student performance, and improve 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 goals of the Mechanical Engineering course are to help students obtain some basic engineering knowledge, design systems, organise processes and solve problems. Currently, most assessment methods are still in the form of final closed-book written tests, and the focus is on memorisation and reproduction of individual knowledge. Therefore, the students tend to focus on memorising exercises that are typical of the exam and lose the ability to think logically and solve non-routine problems. There is a mismatch between the form of assessment and the demands of engineering ability; at the same time, there are structural defects in the dimensions of evaluation and an underestimation of the achievement of course objectives. Therefore, we need to explore reform paths for the assessment methods of mechanical engineering courses. Reforms will add more ways to increase the objectivity of assessment, such as ability stratification, project-based learning and dynamic weighting; at the same time, to support the operation of the evaluation system and achieve the goals of engineering education, standard alignment, multi-stakeholder cooperation and data feedback will also be improved.

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

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

For a long time, the traditional form of testing for mechanical engineering courses has been closed-book written tests, and the questions mainly cover the restatement of principles, derivation of formulas and solution of simplified computational problems. The previous goals were only to find out whether the students had memorised some knowledge points; they did not cover all the abilities students should have in mechanical engineering, such as system design, process planning and fault diagnosis. In preparation for the test, students tend to concentrate their study on typical problems in mechanics and thus have weak analytical and decision-making abilities when faced with irregular engineering problems. The form of assessment does not meet the demands of engineering ability, so course grades are not representative of students' actual professional skills.

According to the internal logic of engineering education, the ability of mechanical engineers should cover all relevant areas, such as scheme design, model construction and analysis, manufacturing process evaluation, technical documentation, etc. A single paper-and-pencil test cannot obtain the actual operation data and judgement quality of students in the areas of computer-aided design, finite element simulation and optimisation of assembly processes. At the same time, this evaluation form does not cover the hidden qualities of cooperation, engineering ethics and techno-economic trade-offs. Prolonged use of only one way will likely result in teaching behavior centred on examinations, reduce the proportion of experiments, course projects and project-based learning in the evaluation system, and thus limit the systematic structure of students' engineering competence[1].

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

Most of the main mechanical courses, such as Mechanical Principles, Mechanical Design and Fundamentals of Mechanical Manufacturing Technology, have various levels of ability goals, including kinematic analysis, strength verification, tolerance and fit, and process planning formulation. Only a small number of the general assessment criteria that are quantifiable and standardised have been added to the evaluation system so far, such as the calculation of degrees of freedom for mechanisms, selection of gear parameters and time quota for simple processes. At present, there are no particular assessment instruments or scores for the high-level ability dimensions of scheme innovativeness, structural manufacturability analysis and trade-off decision-making under multiple constraints. Therefore, there will be no serious suspension of the "analysis", "synthesis" and "evaluation" levels in the course objectives.

The absence of assessment dimensions can also be seen in the artificial division between the quality of the Design process and the quality of the final result. Most mechanical design courses only use the final engineering drawings or calculation specifications for grading; they do not save the iterative process of the design scheme, the reasons for changing parameters, or the coherence in the design logic. Trial and error or the use of empirical formulas by students in their design work cannot be tracked or corrected because there is no process-oriented assessment mechanism. With the development of modern mechanical systems in recent years, many new skills have emerged, but the traditional tests have not kept up with this pace. The above deficiency in structure will reduce the diagnostic power of the assessment and make it difficult to find problems in time.

1.3 The Disconnect Between the Traditional Assessment System and the Achievement of Course Objectives

The calculation of course objective achievement is based on mapping the results of assessment to the set ability indicators. In the old assessment system, many questions in a single test paper often cover several objectives at different levels, and the weights of these objectives have not been organised systematically. For example, in the general problem of calculating gear strength, all three skills—formula memorisation, parameter table lookup and numerical calculation—need to be used, but only the total score is given; we do not know how well each individual skill performed. A coarse-grained evaluation method can only give us a general score and cannot tell us why someone has relatively poor ability.

The larger the misalignment, the more distant the assessment content is from the actual competence requirements of engineering practice. The primary teaching aim of the mechanical engineering course is to teach students how to design a simple mechanical transmission device; however, the design problems in the traditional examination papers are often too basic and idealistic, lacking the necessary links to actual design, such as lubricant selection, configuration of the sealing structure and verification of thermal balance. According to the results of the assessment system, the students can solve problems with simplified models but lack real-life engineering design skills. As a result of this system mismatch, there is an artificial increase in the frequency of the occurrence of the phenomenon that course objectives have been achieved; real problems in the teaching process go unnoticed, and subsequent course improvement measures lack specificity[2].

2. Main Reform Directions for Assessment Methods in Mechanical Engineering Courses

2.1 Construction of a Formative Assessment System Based on Ability Stratification

The construction framework of the formative assessment system is based on the competence model for mechanical engineering, and the cognitive goals have been divided into four levels according to the difficulty of cognition: memory recall, principle application, comprehensive analysis and engineering innovation. The indices of observation and evaluation at all levels are different; for example, online self-assessments and concept maps are used to recall memory at the observation level, small-scale computational or selection tasks are given in the principle application stage, design calculation documents are reviewed chapter by chapter in the comprehensive analysis, and open-ended scheme demonstrations are carried out at the engineering innovation level. Add a short-cycle assessment node in the teaching unit or week to obtain time-series data of students' learning progress, and by observing the passing rates and types of errors for students at all ability levels over time, one can determine where the difficulties for students lie. The stratification above will provide the basis for the ability level in the next dynamic weight adjustment, and thus the assessment results can be directly mapped to the decomposed items of the course objectives.

The first goal of this system will be to organise the hierarchy of assessment rubrics. The bottom level of the assessment for the design problem of the shaft system in the Mechanical Design course is to confirm that the proper standard components have been selected and the dimensions are complete; the intermediate level is to verify the reasonableness of the position and fixing method; and the top level is to evaluate the optimisation capacity of the scheme under the constraints of space and thermal deformation, as well as an additional check on the completeness of the logic chain in the design specifications. The Pass Threshold of each level is set differently. Students who do not meet the low-level requirements will not be allowed to take the high-level test, and the number of their retests and typical errors will also be recorded. Unbalanced development patterns are not without flaws in ability, as shown by the previous total-score weighting method; rather, formative assessment can find all sorts of weaknesses in many places, offer personalised courses and learning support, and obtain detailed information on learning results.

2.2 Systematic Embedding Mechanism of Project-Driven Assessment Modules

The module for project-based assessment is organised according to the whole process of mechanical product development, and the assessment items that were previously distributed among the different parts of the course have now been arranged in a logical order of technical tasks. A typical module can be named "Design and Verification of a Reducer Transmission Scheme under Specified Operating Conditions", and it will cover all the above functions simultaneously: allocation of kinematic parameters, verification of gear strength, shafting structural layout, selection of lubrication method, etc. There are several necessary delivery nodes in this module, and each node has different scoring standards and normative requirements; for example, the scheme review report, the first draft of the calculation specification, the three-dimensional model and the two-dimensional engineering drawing are all such nodes. The module score is the sum of the scores of several independent problems in the old test paper, and it can also indicate a student's general design ability and systematic thinking in solving coupled sub-problems based on the dependency relationship among nodes[3].

Systematic embedding should place the module in time and space during the teaching process, and it needs to be consistent with the lecture content. The Fundamentals of Mechanical Manufacturing Technology course is a typical case; therefore, the project module can be set as "Process Planning and Fixture Scheme Design for a Specific Part". This module is started after the theoretical instruction on process datum selection and finished by the end of the chapter on fixture positioning error analysis. At any time in the module, there are many intermediate checkpoints for sub-tasks, such as blank selection, operation sequencing, allowance determination and positioning scheme; all of them must meet the minimum passing requirements. Break down the components of the assessment module into several small, independent items to reduce the amount of work students need to do at one time, and at the same time, teachers can systematically track all the changes students make in the design process and promptly find any deficiencies in these sub-skills.

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

Modification of process weights will no longer be proportional (e.g., 30% for regular performance and 70% for the final exam); instead, based on the different stages of course progress, we will set weight adjustment rules. In the early stage of the course (the period for introducing concepts and teaching basic skills), the coefficients of in-class quizzes and unit assignments in the final overall grade are increased, and the preset weight of the later concentrated assessment is reduced. As the course enters the general Design Stage, the focus gradually moves away from the weight in the project module and design document, and in the final period, the cumulative effect of the low-weight nodes in the early stages is corrected by weighted back-adjustment. A sliding-window weighting method is used to adjust the weight of each assessment according to the coefficient of variation of students' three most recent formative assessments, and at the same time, a group difficulty response factor is added to show changes in the collective difficulties of teaching[4].

A particular form of the weight-response function will be used in the mechanical Design and Manufacturing course. The weight of a given assessment node is set to be positively correlated with the corresponding ability-level coefficient and the score dispersion of the student cohort at that node, and is multiplied by the prior discrimination parameter of the node. For example, in the CNC Machining Technology course, the weight of the manual programming node is relatively high compared to that of the G-code simulation node because it has more algorithmic logic and constraint judgment, and it is more fundamental to the subsequent learning. After identifying the collective difficulties of the student group in the tool path optimisation stage in the middle of the course, the system will automatically increase the upper weight limit of the next related assessment nodes, reduce the weights of high-score nodes that have already passed, and reallocate the retest weight coefficients for nodes that have not met the standards. Based on changes in the difficulty of teaching and students' performance, one can adjust the structure of the assessment system dynamically; thus, a fixed weight should not be used, and different final-grade composition plans will be provided for students at all times.

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

3.1 Strategy to Align Assessment Standards with Technical Specifications of the Mechanical Industry

The development of the assessment norms should be based on the existing technical norm system for the mechanical industry and cover all relevant contents, such as the drawing annotation specifications (GB/T 1182) for geometrical tolerance, the material selection criteria (GB/T 3077) for alloy structural steel, the process documentation requirements (JB/T 9169) of process planning forms, and the numerical control programming standards (GB/T 18726) for NC machining codes. The measurable indicators in the above specifications should be turned into deduction items or grade thresholds for the scoring rubric. For instance, in the assessment of the Mechanical Drawing course, the evaluation of view representation schemes is no longer based solely on the completeness of graphics; quantitative indices for specification compliance have also been added, such as the spacing of section lines, closure of dimension chains and placement of surface roughness notations. At the core of the alignment strategy is to extract the relevant clauses and recommended clauses from industry specifications, and based on this, set up corresponding mandatory items and bonus items in the assessment system to create an evaluation scale with engineering equivalence.

A Bidirectional Mapping Matrix of the assessment standards and industry specifications should be built to organise the corresponding orders. The row vectors of the matrix are the assessment nodes for the course, and the column vectors are the corresponding technical specification items; each element in the matrix is marked with the application level of the specification (recognition, interpretation or execution). For example, in the assessment of process planning formulation for the course Fundamentals of Mechanical Manufacturing Technology, the location datum annotation on the operation sketch should comply with GB/T 24737, the tolerance grade of process dimensions must be in line with GB/T 1800, and the selection of cutting parameters should be based on the recommended range in the Machining Process Handbook. The assessment score sheet will show the specification number and conformity criteria directly, and the process documents submitted by the students will be checked for compliance. Therefore, the evaluation results will directly show how well students have learned the technical language of the industry and reduce the semantic conversion loss between school assessment and the engineering practice environment.

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

The organisation structure of the multi-stakeholder collaborative assessment includes the lead course instructor, lab instructors, graduate teaching assistants, students and student representatives, etc., and all these parties are responsible for carrying out the evaluation tasks. The leader of the course is to build the general plan for the assessment rubric and to conduct a final check of the high-weight nodes. The lab teacher will observe the students' work and record the initial data. Graduate teaching assistants are in charge of the daily management, such as the first check of the compliance of drawing assignments and verification of the format of calculation documents. Student representatives will try out the rubric and confirm that the scores given by their peers are reasonable. Information transfer among all parties concerned in the problem will be organised in the form of evaluation work order transfer; each work order will specify the evaluation subject, deadline, delivery requirements and dispute resolution path, etc., so that responsibilities are clear and traceable.

The Operating Logic is based on the ideas of role interdependence and process separation. Role interlocking requires that the scores of the same assessment node be entered independently by at least two parties, and if there is a significant difference among them, a third party will carry out a review. For example, in the Mechanical Design course, the lead instructor will judge the principle feasibility of the design plan; the lab instructor will verify the assembly and motion functions of the prototype; the teaching assistant will check the document format and completeness of drawings; and three students from the same group will anonymously rate the contributions and cooperation of other students. Process separation is to separate the evaluation activities from those of teaching assistance so that the person in charge of teaching is not also responsible for grading. Based on the above structure, the scoring matrix can provide both individual ability scores and a team collaboration coefficient; the latter is determined by the degree of statistical convergence in the score deviations of all parties and is not based on subjective judgment[5].

3.3 Data-Driven Teaching Feedback Closed-Loop Design Method

The beginning of the teaching feedback closed-loop is the raw scoring data from all assessment nodes, which is then organised and stored in the course ability database. All the recorded data have an attribute label, such as the calculation capability of gear strength or the formulation ability of process planning, and a timestamp; thus, they can be used for cross-sectional analysis (comparison among classes) and longitudinal analysis (tracking the development of individuals). Item Response Theory is used to estimate the parameters of the assessment items, obtain discrimination and difficulty parameters for each item, and then inefficient or invalid assessment records can be excluded. Radar charts and heatmaps can be used to show the distribution of scores for all ability labels in the student group and identify areas in knowledge and skills that are frequently misunderstood. For example, if a class has a much lower score than the other classes in the category of "Selection of axial positioning methods in shaft system structural design", then a particular diagnosis of this weak link will be issued.

The end of the closed loop is the generation of intervention signals for both teaching behaviour and learning behaviour. According to the set deviation threshold, the two kinds of feedback given by the system are teacher-oriented course adjustment recommendations (e.g., slow down the speed of the class lecture in a weak area, add specific exercises, modify constraints for the next project module) and student-oriented personalised learning path recommendations (e.g., indicators of assessment nodes that need to be retaken, links to supplementary learning materials, optional module modifications). The feedback signal can be seen in the data dashboard of the teaching management window, and the corresponding adjustable parameters, such as the exercise difficulty coefficient and the retest window length, are also shown there. A closed-loop Design will be added to automatically link the results of the assessment with teaching activities, reduce the time lag in feedback of the old way, and provide specific data for continuous improvement of classes.

Conclusion

In accordance with the three directions of ability diagnosis, task integration and weight adjustment, the reform of assessment in the mechanical engineering course should be promoted. According to the division of abilities, formative assessment can determine at what stage a person has problems in memory recall, application of principles, overall analysis and engineering innovation. All the Design Skills will be covered in one large-scale technical problem of the project-based assessment modules.

Dynamically change the weights of the process according to the current stage of the course and the performance of the students. The good operation of the above paths is based on the bidirectional mapping of assessment standards and industry technical specifications, the role-interlocking mechanism of multi-stakeholder collaborative assessment, and the data-driven teaching feedback closed-loop. In the future, an intelligent assessment platform will be introduced to weigh scores and provide timely feedback; many ways will be developed to collect data on the continuous development of students' abilities in all courses and semesters, and how to integrate non-academic qualities, such as judgment of engineering ethics and cooperative communication, into the assessment system will also be studied in detail.

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.