Advances in Educational Technology and Psychology (2026)
Clausius Scientific Press, Canada
DOI: 10.23977/aetp.2026.100416
ISSN 2371-9400 Vol. 10 Num. 4

Task-Bound, Process-Traceable and AI-Interactive Teaching Reform for the Course Design of Mechanical Design

Chunxing Gu[1,a,*], Di Zhang[2,b]

[1]*School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China*
[2]*School of Mechanical Electronic Technology, Shanghai Jianqiao University, Shanghai, China*
[a]*chunxinggu@foxmail.com*, [b]*dizhang2018@foxmail.com*
[*]*Corresponding author*

Keywords: Artificial Intelligence, Task-Bound Design; Process Traceability; Engineering Education Accreditation; Curriculum Reform

Abstract: Course design marks the first independent design practice for undergraduate mechanical engineering students. However, the delivery of this task consistently falls short of its intended competency-building objectives: the decision-making and trade-off processes leave almost no traceable evidence for instructors to assess, and the professional values integrated into Chinese curricula under the framework of Curriculum Ideological and Political Education (CIPED) are confined to opening lectures or concluding report sections, with no demonstrable connection to actual design practice. This paper describes a reform that is task-bound, process-traceable and AI-interactive while leaving the assessment structure untouched. Each design dimension is bound to a named step of a nine-step protocol for the standard two-stage cylindrical-gear reducer project, so every judgment must appear as something that can be opened and checked: a line of the calculation sheet, an interactive 3D model, a passage of the report, or a field of the title or calculation block. Two AI roles work inside that workflow, and each leaves a trace. The AI tutor requires students to justify, in their own words, the candidate solutions they adopt or reject; the AI design agent renders the submitted model as an interactive view against which a named safety-relevant decision can be inspected. A diagnostic rubric, with a pre-registered multi-batch rule that can return “not established”, turns the question of whether this formation reaches the design work into an inspection of deliverables already graded. The reducer is the main case, transmission-system designs the verification cases.

1. Introduction

The course design of mechanical design is where undergraduates stop following worked examples and start producing a design of their own. Given a duty, a transmission ratio and a service life, they choose a transmission scheme, calculate parameters, size and check the parts, build a 3D

model and issue drawings [1, 2]. What the course really assesses is design judgment: how parameters are fixed, how schemes are weighed, how decisions that bear on safety are defended. None of that is legible in the compliant drawing handed in at the end [3].

Chinese universities have worked mainly on what is taught and with what tools. Miao, Kou and Wang [3] use a neural-network surrogate for reducer parameter design and evaluation, so that trade-offs among strength, cost and efficiency can be explored without repeated hand calculation, and note that the project still runs on handbook lookup while industry has moved to parametric design and topology optimization. Wu et al. [4] add a teaching platform, SolidWorks and a parametric part library to a before-, in- and after-class structure. He et al. [5] argue from the Industry 4.0 angle that AI can keep content current and support adaptive tutoring, and describe an MIT case in which a generative model produces bicycle-frame variants under mass, cost and stiffness constraints [6].

A second line of work acts on the process. Song et al. [7] run an engineering-driven cycle of six stages in a parallel-programming course and confine a large language model to the role of assessor, which reads students' code and analysis documents against a three-dimension rubric; gains over two academic years are reported. For project work of this kind, reforms built on outcome-based education and on accreditation have reorganized objectives, task cards and assessment [8, 9], and discipline-level work argues that what the course teaches must follow technological change [10].

What these reforms share is a change in what is taught, how it is presented, or who does the judging, and two gaps remain. Professional values still enter as an opening lecture or a closing paragraph, and nothing in the deliverable separates a genuine ethical judgment from a sentence that asserts one. Where a model is asked to judge the design process [7], the judgment becomes the object of trust, since the criteria are ones the instructor did not write and cannot inspect. This paper takes a different route: AI is used to make the student's reasoning leave checkable artifacts, while the judging, and the grading, stay where they were. A nine-step protocol is set out for the two-stage cylindrical-gear reducer, with transmission-system designs as the verification cases; the decision rule is fixed in advance, so that the reform can be shown to fail.

2. Three Weak Points in the Current Course Design

The course is nominally aimed at engineering competence, yet neither its formation nor its observation has a stable structure. Three weaknesses follow.

2.1. Design Task and Competence Formation Drift Apart

The course is normally run as a pipeline: assign the task, consult the handbook, calculate the parameters, produce the drawing. Students can walk through it and still be unable to say what a parameter means inside the complete machine, or why one scheme was preferred to another [1]. At the nodes that matter most, namely transmission-ratio allocation, gear-module selection and shaft-system layout, there is no explicit path for judgment, and the complaint follows: the student can calculate and can draw, but cannot explain the design [3, 11].

2.2. Design Decisions Leave No Observable Trace

A good reducer design is distinguished by how parameters are fixed, safety margins justified and schemes compared, and none of it shows on the surface of a submitted deliverable. Assessment falls back on result-based indicators: whether the report is correct, whether the drawings follow convention, whether the machine functions [3, 12]. An instructor can decide whether a strength requirement is met; deciding whether the student understood the design is harder. Reasoning behind key parameters is dropped because nothing at process level captures it, and grading drifts toward

holistic impression [3].

2.3. Cultivation Elements Have No Checkable Link to Assessment

CIPED content, engineering ethics, green and low-carbon design and professional responsibility appear as one opening lecture plus a closing sentence in the report. Nothing connects them to the work, so a claim that safety, standards and responsibility were considered carries no more weight than a sentence that mentions them and nothing else [5]. Evaluation of engineering quality is likewise disconnected from the graded deliverables, and the results rarely feed back into the next task [7].

All three shortcomings can be traced to one fundamental gap: neither the development of core competencies nor the integration of affiliated educational elements has been systematically embedded in the course design, and no traceable pathway connects the design process to assessment outcomes. A reform therefore has to rebuild the workflow, the mode of observation and the feedback loop together.

3. A Task-Bound, Process-Traceable, AI-Interactive Reform Scheme

Figure 1 shows the scheme. Four principles define it, and none of them adds content: each is a condition on how the work is carried and judged.

(1) Task-bound: each design dimension is attached to a named step of the reducer workflow, so that task, competence and value are carried by the same piece of work instead of three separate documents;

(2) Process-traceable: the output of each step must be a checkable deliverable — a calculation-sheet line, an interactive 3D model, a report passage, a title-block or calculation-block field — rather than a statement that cannot be opened;

(3) AI-interactive: an AI tutor, which requires students to justify the solutions they take from it, and an AI design agent, which renders the submitted model as an interactive view for re-checking a safety decision; both leave a record;

(4) Assessment-neutral: no new grading item, no change of weight, no alteration to the goal-to-indicator mapping; the only addition is one executable check on deliverables that are already graded.

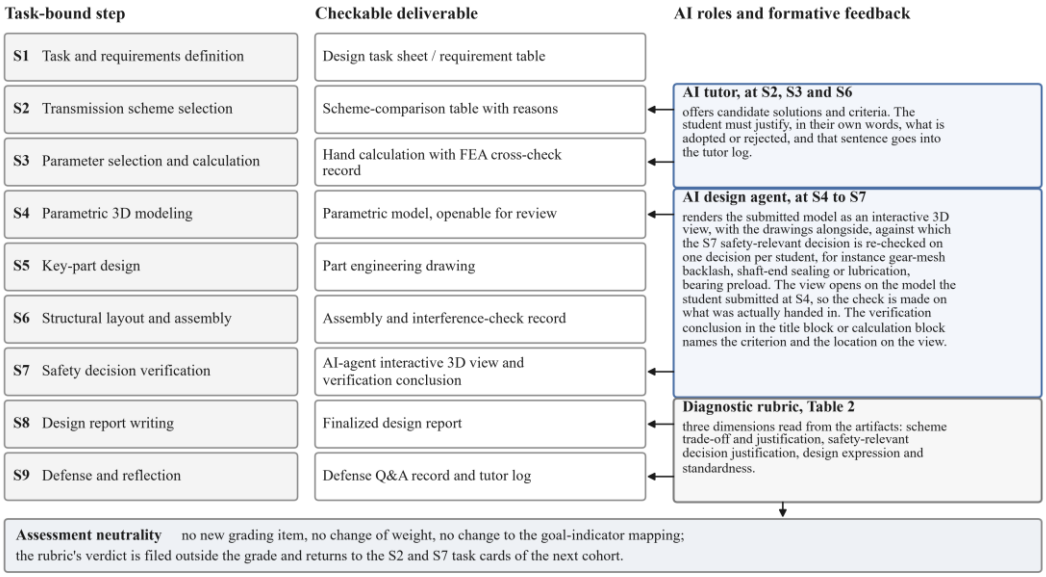


Figure 1: Overall framework of the task-bound, process-traceable, AI-interactive reform.

3.1. A Task-Bound Nine-Step Workflow

The scheme turns on a protocol table. Table 1 binds each design dimension to one of the nine steps of the reducer workflow and, for every step, records the principal and secondary dimension, the deliverable in which the result must appear, and where the judgment becomes visible. Each step is assigned exactly one principal dimension. A given dimension may be referenced in multiple steps—for example, rigor features in steps S3, S5 and S7—on the condition that each instantiation is reflected in a separate, identifiable deliverable.

Table 1: Task-bound nine-step design workflow, with the gear reducer as an example.

Step	Principal dimension	Secondary dimension	Checkable deliverable	Where the judgment becomes visible
S1 Task and requirements definition	Requirement and constraint expression	Communication	Design task sheet / requirement table	Design basis and constraints passage of the report
S2 Transmission scheme selection	Scheme trade-off and justification	Standards and awareness	Scheme-comparison table with reasons	Scheme comparison section of the calculation sheet
S3 Parameter selection and calculation	Rigor in calculation and verification	Scheme trade-off	Hand calculation with FEA cross-check record	Verification line; title-block or calculation-block field
S4 Parametric 3D modeling	Engineering expression	Autonomy	Parametric model, openable for review	Model file and drawing
S5 Key-part design	Rigor in part design	Safety-relevant decision	Part engineering drawing	Drawing and tolerance annotations
S6 Structural layout and assembly	Engineering expression	Scheme trade-off	Assembly and interference-check record	Assembly model with explanatory passage
S7 Safety decision verification	Safety and responsibility	Rigor	AI-agent interactive 3D view and verification conclusion	Safety-verification line of the title or calculation block
S8 Design report writing	Expression and communication	Rigor	Finalized design report	Relevant passages of the report
S9 Defense and reflection	Communication and reflection	Autonomy	Defense Q&A record and tutor log	Tutor log

The safety and responsibility dimension carries what the course previously handled as engineering ethics and CIPED content, and it is not a discussion of ethics appended to the project. At S7 the student makes a verifiable statement about one safety decision while looking at the model already submitted, which is what attaching a cultivation element to real design work means.

3.2. Two AI Roles and the Records They Leave

AI carries checkable intermediates here. Its value lies in making the students' deliverables and logs openable and reviewable, not in what it produces on its own.

(1) AI tutor, which demands justification rather than answers: at scheme comparison, parameter verification and structural layout, the student puts a question and receives candidate solutions or criteria, and must state in the deliverable, in their own words, why the solution was adopted or rejected; that sentence goes into the tutor log. Whether the AI answer was right is not recorded, and the log turns the exchange into a process record that can be read later.

(2) AI design agent, which turns the model into a reviewable view: the agent renders the submitted model as an interactive 3D view with the drawings alongside, against which one safety decision, for instance gear-mesh backlash, shaft-end sealing or bearing preload, is re-checked. The

student then completes one safety-verification line in the title or calculation block, naming the criterion and where on the view it holds.

Together the two roles leave a two-level record, a tutor log at process level and a completed field at deliverable level, from which an instructor can judge a student's reasoning and the student can review their own path [13].

3.3. A Diagnostic Rubric for the Competence Dimensions

Judging whether a decision is well argued requires criteria that two readers can apply to the same artifact, and Table 2 supplies them for three dimensions drawn from the course goals. The rubric supports formative diagnosis only and leaves the grading structure as it stands [11]. The dimensions named in Table 1 do not map one-to-one onto rubric rows; three are decided directly from the artifacts, namely the trade-off and justification of a scheme, the justification of a safety decision, and the expression and standardness of the deliverable. Rigor, standards awareness and requirement expression are judged through those same rows. Autonomy and reflection, at S4 and S9, are read from the log and the defense record.

Table 2: Rubric for the design-decision competence, used for formative feedback.

Competence dimension	Insufficient	Partially sufficient	Sufficient
Scheme trade-off and justification	Lists conclusions or copies the handbook; no grounds, or grounds that cannot be verified	States the direction of the trade-off, but grounds are vague and lack criteria	Makes a verifiable trade-off against explicit criteria of strength, efficiency, cost and manufacturability
Safety-relevant decision justification	No engineering statement on the safety decision, or one detached from the deliverable	Points out the decision, but grounds are incomplete or the comparison location is unclear	Gives a complete justification for a concrete safety decision that can be verified on the interactive view or the drawing
Design expression and standardness	Deliverable cannot be opened, fields are missing, conclusions have no anchor	Deliverable is largely complete, but the route from evidence to conclusion is unclear	Deliverable is verifiable, fields are complete, and conclusions follow the stated criteria

3.4. Assessment Neutrality and the Closed Loop

Assessment neutrality is preserved: no grading item is added, no weight changes, the goal-to-indicator mapping is untouched, and the AI roles and the nine-step table are embedded in deliverables students already produce. One thing is added on the assessment side, a question a second reader must answer from the graded deliverables alone: is this design statement attached to a real technical judgment that can be opened and compared against the view or the drawing? Section 5.1 gives it operational form.

The loop closes where it began: what one batch shows informs the next batch, and feedback shapes the justification a student makes in the following step. Keeping assessment neutral is what makes the scheme portable, since it can be dropped into a running course instead of asking for a rebuild.

4. Teaching Practice and Case Analysis

The reform was piloted in the course design of mechanical design for mechanical-engineering undergraduates, with the reducer as the central task. Two consecutive offerings took part, one teaching class of about 80 students in each, and every offering ran for two weeks; three instructors

served as reviewers, two of whom taught the course. The reducer is analyzed as the main case; a planetary transmission system and a bicycle-frame lightweighting problem serve as the verification cases, testing whether the scheme survives a change of transmission type and design object.

4.1. Main Case: Two-Stage Cylindrical-Gear Reducer

The project runs through the nine steps of Table 1, with the instructor, the students and the AI tools taking different roles at different stages.

(1) Requirements, scheme, calculation, S1 to S3. The project opens from a real engineering setting with a given power, ratio and life. Students complete the requirement table at S1, then weigh cylindrical against bevel gearing and single against two-stage layouts at S2, writing down the reasons. They calculate parameters and cross-check the hand calculation by finite element analysis (FEA) at S3, which lands rigor on the calculation sheet and the title-block field. Many students arrive focused on the numbers, with little sense of the trade-offs among efficiency, volume and cost, which makes S2 the place for targeted explanation.

(2) Modeling, part design, assembly, S4 to S6. Gears and shafts are generated by parametric modeling at S4; the key-part drawing with tolerance annotations is completed at S5; the assembly model at S6 supports an interference check with a written record. The AI design agent supplies the interactive view at this stage, letting students look at gear-mesh backlash and seal structures, and prepares the carrier that S7 uses.

(3) Safety verification and writing, S7 and S8. S7 is the pivot. A student selects one safety decision from the assembly model, draws a verification conclusion against the interactive view, and completes one safety-verification field naming the criterion and the location of the comparison. The tutor records the justification in the student's own words, not a restatement of the AI answer. These justifications are written into the final report at S8, so the deliverable contains at least one verifiable engineering statement rather than a claim about one.

(4) Defense and diagnostic feedback, S9. The instructor applies the rubric in Table 2 to each student's decisions, giving formative feedback while leaving the grade alone, and draws on the tutor log for process evidence. Weaknesses that recur, scheme comparisons without criteria for instance, are written into the S2 task card for the next cohort.

4.2. Verification Cases: Planetary Transmission and Bicycle Frame

To verify the generalizability of the proposed scheme beyond the reducer case, two supplementary design tasks were adopted for validation: a planetary transmission system design [2], and a bicycle-frame lightweighting problem solved using topology optimization [14]. The latter is identical to the design case employed by He et al. to illustrate generative design methodologies [5]. Both differ in design object, constraints and toolchain, yet the decisions map onto the three dimensions of Table 2 without modification, and the rubric returns consistent judgments across objects.

5. Effectiveness and Decision Criteria of the Reform

What is proposed here is a teaching design and a checkable protocol, not a claim of measured improvement. Its effectiveness criteria are therefore registered in advance across several batches, so that the reform cannot certify itself from one favorable run.

5.1. Pre-Registered Decisions and Failure Examples

The decisions below were registered before the first offering. The two instructors judged every safety statement independently; the third reviewer, who did not teach the course, judged a random sample of 40 statements in each offering.

(1) Decision 1, checkability established: two reviewers, working only from the graded deliverables (calculation sheet, drawings, 3D model, report, title and calculation blocks, tutor log), independently decide whether a design statement is attached to a real technical judgment, and their verdicts of "sufficient" or "present but not attached" agree under a stated criterion, with a stated rule for disagreement.

(2) Decision 2, AI interaction established: the tutor log and the interactive view are used by reviewers to re-check at least one safety decision in at least one of the two batches.

(3) Decision 3, assessment neutrality established: the grading structure, the weights and the goal-to-indicator mapping are unchanged, and the checkability verdict is filed as a review record outside the grade.

Failure examples are given for contrast. A report may say that the design took full account of safety and responsibility, yet the sentence cannot be matched to any decision on the drawings, model or calculation sheet, and is recorded as "present but not attached". A tutor log may repeat the AI solution without the grounds for the trade-off, which does not count toward "sufficient" at S2. The safety-verification field may be blank, or say only "verified", in which case Decision 2 is not met.

Where Decisions 1 and 2 hold across at least two batches and Decision 3 also holds, the outcome reported is that a checkable generation-and-feedback loop has been established. Where any one fails, the outcome is "not established", with the instances that failed.

5.2. Observations from the Formative Diagnosis

Used formatively, the rubric separates what the previous assessment could not: a comparison with criteria from one with only a direction, and a safety field that names a criterion and a location from one that says "verified". Neither judgment needs a new grade; both need the deliverable to be openable. The operation log and interactive design view provide retrospectively traceable process evidence, facilitating the core transition the course aims to achieve: a shift from mere functional implementation to explicit justification of design decisions [13]. This paper reports the protocol and the diagnostic capability it creates.

5.3. Wider Value

The scheme's value extends beyond this course in three ways. It can be embedded in any running course design without a rebuild. Because the AI roles are confined to carrying checkable intermediates, the risks of AI over-claiming and student over-reliance stay bounded [15, 16]. Furthermore, the one-to-one mapping among tasks, design dimensions and deliverables generates the process-level evidence required for program accreditation, which formally evaluates how graduate attribute indicators are achieved [11, 12].

6. Conclusions

The course design of mechanical design fails in three linked ways: the task has drifted from competence formation, the decision process is hard to observe, and the cultivation elements have no checkable link to the work. The reform proposed here answers all three with one move, attaching

the judgment to an artifact a second reader can open: a nine-step workflow, two AI roles, and a rubric with a pre-registered rule that can answer “not established”. What changes is not the content or the tools but what is inspected. Two limits should be admitted: the evidence comes from two offerings at one institution, and the agreement criterion across reviewers is still loose. Tightening it, extending the cases to other design objects, and linking the diagnostic feedback to the task cards automatically are the next steps.

References

[1] Pu, L.G., Chen, G.D. and Wu, L.Y. (2013) Mechanical Design. 9th Edition, Higher Education Press, Beijing. (in Chinese)

[2] Budynas, R.G. and Nisbett, J.K. (2015) Shigley's Mechanical Engineering Design. 10th Edition, McGraw-Hill, New York.

[3] Miao, J.W., Kou, Z.Q. and Wang, K. (2026) Reform and Exploration of AI-Enabled Teaching in the Course Design of Mechanical Design. Journal of Hubei Open Vocational College, 39, 145-146, 150. (in Chinese)

[4] Wu, R.L., He, C.M., Li, H.H., Ma, L.W. and Zhang, H.X. (2025) Exploration of the Teaching Model of the Mechanical Design Course Driven by Digital Technology + Artificial Intelligence under the Background of New Engineering. China Modern Educational Equipment, 19, 71-73. (in Chinese)

[5] He, S., Wang, J.R., Lu, T. and Liu, K.H. (2026) Research on the Reform and Practice of AI-Empowered Mechanical Design Courses. Journal of Higher Education, 13, 18-21. (in Chinese)

[6] Regenwetter, L., Srivastava, A., Gutfreund, D., et al. (2023) Beyond Statistical Similarity: Rethinking Metrics for Deep Generative Models in Engineering Design. Computer-Aided Design, 165, Article 103609.

[7] Song, Z.R., Deng, Q.N., Liang, X.Y. and Guan, H.B. (2026) Engineering-Driven and AI-Empowered Teaching Reform for Parallel and Distributed Programming. Computer Education, 10, 72-77. (in Chinese)

[8] Jiao, W.M., Xu, Y. and Li, J.Y. (2024) Teaching Reform of the Mechanical Design Course Project Based on the OBE Concept. University, 2, 101-104. (in Chinese)

[9] Xiao, S.H., Li, W., Jiang, F.G., et al. (2023) Research on the Reform of the Mechanical Design Course Project under the Concept of Engineering Education Accreditation. China Modern Educational Equipment, 15, 78-80, 90. (in Chinese)

[10] Wang, S.T., Ling, L., Chen, J.Y., et al. (2025) Reform and Practice of the Mechanical Design Theory and Method Course from the Perspective of Technological Change. Research in Higher Education of Engineering, 1, 34-39. (in Chinese)

[11] Biggs, J. and Tang, C. (2011) Teaching for Quality Learning at University. 4th Edition, McGraw-Hill and Open University Press, Maidenhead.

[12] ABET (2023) Criteria for Accrediting Engineering Programs, 2024-2025. ABET, Baltimore.

[13] Woolf, B.P. (2008) Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning. Morgan Kaufmann, Burlington.

[14] Bendsøe, M.P. and Sigmund, O. (2003) Topology Optimization: Theory, Methods and Applications. Springer, Berlin.

[15] Zawacki-Richter, O., Marín, V.I., Bond, M. and Gouverneur, F. (2019) Systematic Review of Research on Artificial Intelligence Applications in Higher Education—Where Are the Educators? International Journal of Educational Technology in Higher Education, 16, Article 39.

[16] Holmes, W., Bialik, M. and Fadel, C. (2019) Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign, Boston.