Article

Design and First-Round Implementation of Project-Based Reform in an Application-Oriented Course: A Case Study of an Undergraduate Applied Statistics Course

Qing Luo[1,*]

[1] School of Business, Geely University, Chengdu, China
* Correspondence: Qing Luo, School of Business, Geely University, Chengdu, China

Abstract: To address the theory–practice gap in statistics courses for economics and management students at application-oriented universities, we reformed an undergraduate Applied Statistics course around authentic project tasks and examined the reform in a three-group quasi-experiment. The reform is scheduled as a two-year (four-semester) implementation, and this paper reports its first-semester pilot. The redesigned course follows a "1 + 7 + N" structure-one tool-orientation module (SPSS), seven project task packages aligned with the course's knowledge modules, and supporting teaching resources-assessed through individual quizzes, team deliverables, process portfolios, and a final defense. A pilot class (n = 89), a regular undergraduate comparison class (n = 57), and a junior-college transfer (zhuanshengben) class (n = 73) completed identical baseline instruments in Week 1. Interim findings show that: (1) the three groups differed significantly in self-rated prerequisite knowledge (Ms = 2.54, 2.84, and 1.97), with transfer students showing a clear articulation gap; (2) pilot-class students held high expectations for project-based learning (M = 3.60), especially for authentic cases and enterprise practice (M = 3.84); and (3) first-round projects were completed as designed, with instructor rubric scores of team deliverables rising from Project 1 (M = 81.7) to Project 2 (M = 85.9), inter-group peer evaluation differentiating team products (M = 90.0, range 82.5–93.7), and within-team peer evaluation indicating sound team functioning. The study offers an operable pathway for project-based course reform; causal effects remain to be tested against post-course data.

Keywords: application-oriented course; project-based reform; applied statistics; quasi-experimental design; teaching evaluation

Received: 08 October 2026
Revised: 21 October 2026
Accepted: 12 November 2026
Published: 18 November 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Application-oriented universities are positioned to cultivate students' practical abilities and professional competencies, and their course teaching is expected to serve this goal [1, 2]. However, statistics courses in economics and management programs have long faced three prominent problems. First, content organization follows the logic of disciplinary knowledge, with concepts, formulas, and derivations dominating the classroom, leaving students "unable to apply what they have learned." Second, teaching is dominated by lecturing, while hands-on software practice and authentic data processing remain weak, producing students who "can solve examination problems but cannot analyze real-world situations." Third, assessment relies mainly on final paper-and-pencil examinations, which fail to capture students' ability to solve practical problems. These problems are compounded by the intake profile: economics and management students have relatively weak mathematical foundations and report anxiety toward abstract statistical concepts, while the enrollment of junior-college-to-undergraduate transfer (zhuanshengben) students further intensifies the within-class differentiation of learning conditions.

Project-based learning (PBL) offers a well-documented response to these problems. Reviews and meta-analyses indicate that PBL improves academic achievement relative to traditional instruction (average effect size d+ = 0.71) and yields positive affective and behavioral outcomes in higher education, while also identifying implementation conditions such as teacher support, group dynamics, assessment design, and technological support [3–7]. In statistics education specifically, the GAISE College Report recommends teaching statistics as an investigative process, integrating real data with context and purpose, fostering active learning, and using technology to explore concepts and analyze data [8]. Empirical studies of project-based statistics courses report improved student attitudes, engagement, and understanding across diverse settings, including community colleges, engineering programs, and technology-enhanced collaborative designs [9–12]. Two gaps remain, however. First, most evidence comes from North American and European universities; little is known about project-based statistics teaching in Chinese application-oriented universities, where students' quantitative preparation is weaker and more heterogeneous. Second, vertical transfer students-who enter undergraduate programs from junior colleges and are documented to experience transfer shock and preparation gaps [13]-have rarely been examined as a distinct group in course-level reform studies. In addition, the peer-assessment literature warns that summative peer ratings in group work are subject to leniency and validity threats, which must inform the design of process assessment [14, 15].

In China, the transformation of local undergraduate institutions toward application-oriented education is a national policy direction [1, 2], and the development of applied universities has been increasingly documented [16]. The Educational Philosophy of Geely University (Ji Xiao Zi [2026] No. 17) articulates the university's "local, industry-oriented, and application-focused" positioning and its educational objectives of "solid theoretical foundations, strong practical abilities, and high comprehensive quality," and proposes deepening the institution-level project-system talent cultivation model [17]. Project-based talent cultivation must ultimately be implemented within individual courses, and the project-based reform of courses is precisely its micro-level carrier and implementation foundation. Against this background, the course team, supported by a university-level educational reform project, implemented a project-based reform of the Applied Statistics course and launched the first-round pilot in September 2026. This paper reports the overall design of the reform and the early evidence from its first-round implementation, focusing on two questions: (1) how the project-based reform of the course was systematically designed and implemented; and (2) what characteristics emerged in the learning baselines of students from different educational backgrounds and in the implementation of project tasks during the first round.

2. Overall Design of the Project-Based Reform

2.1. Course System: The "1 + 7 + N" Structure

Following a backward-design logic [18] and outcome-based education principles [19], the course was restructured around the question "what statistical competencies are required to complete an authentic data-analysis task?" into a system of "one tool-orientation module + seven project task packages + N types of teaching resources." The tool-orientation module delivers concentrated training in basic SPSS operations at the beginning of the semester, removing tool barriers for subsequent projects. The seven project task packages correspond one-to-one with the course's knowledge modules-statistical survey, data organization and presentation, descriptive statistics, parameter estimation, hypothesis testing, analysis of variance, and linear regression-and are completed sequentially by student teams, forming a complete chain of "posing questions–collecting data–analyzing data–interpreting results–making recommendations." The N types of teaching resources include project task sheets, datasets, case packages, SPSS operation guides, assessment rubrics, and exemplar student work, supporting "learning by doing." Project contexts are drawn preferentially from campus life and major-related

fields, with articulation to disciplinary competitions also considered. The design is project-based rather than problem-based in the sense that each package is anchored in a tangible product rather than in an open-ended problem scenario alone [20]. The N types of resources are maintained as versioned teaching work packages that will be iteratively optimized after each implementation round on the basis of classroom evidence.

2.2. Implementation Progression and Assessment Mechanism

Project implementation follows a progression of "cognitive experience–imitative operation–simulated integration–authentic application": in the first two projects, the instructor provides complete data and operational demonstrations for students to imitate; in the middle projects, scaffolding is gradually withdrawn and teams design their own schemes; and the final project requires students to independently complete the full workflow around an authentic problem and to present a public defense. The assessment mechanism accordingly consists of four components: individual quizzes to ensure attainment of knowledge objectives; team project deliverables scored by the instructor with a rubric (100-point scale) across six dimensions-data quality, method selection, software operation, result interpretation, decision recommendations, and report standardization; process portfolios comprising classroom observation records, within-team peer evaluation, and student reflective journals; and presentations with defense to examine decision communication and teamwork. Given the documented leniency and validity threats of summative peer assessment [14, 15], within-team peer evaluation is deliberately used for qualitative corroboration and anomaly detection rather than as a high-stakes differentiating measure. In addition, after each project defense, every team rated the deliverables of all other teams in its class on a 100-point scale (inter-group peer evaluation), providing a peer-based ranking of team products.

2.3. Research Design: A Three-Group Quasi-Experiment

The reform is planned as a two-year (four-semester) implementation, with project tasks and assessment instruments revised after each round. To examine the reform, the first-round implementation adopted a quasi-experimental design with a pilot class and comparison classes. The experimental group comprises the pilot classes undergoing project-based reform (two 2024-cohort auditing classes, the pilot group, coded S). Project teams (3–7 members each) were formed entirely through student self-selection, without instructor intervention. The regular comparison group (coded D) consists of regular undergraduate parallel classes not undergoing the reform in the same semester, with identical syllabus, textbook, and contact hours. Because junior-college transfer students differ markedly from regular undergraduates in prerequisite foundations, program structure, and age composition [13], a separate transfer group (coded Z) was established for testing the heterogeneity of reform effects across student sources. All three groups completed a pre-course baseline and expectations questionnaire and the same knowledge pre-test in Week 1, and will complete a post-course questionnaire and post-test in Week 17. The core measurement items of the three groups' questionnaires are identical; the pilot-group questionnaire adds items specific to project-based learning, and the transfer-group questionnaire adds background items on the junior-college stage as covariates. Questionnaire data from the unreformed 2025 classes serve as a historical baseline. All multi-item scales showed acceptable to high internal consistency in each group (Cronbach's α = .72–.96; details in Section 3). The knowledge pre-test was administered to establish baseline equivalence for the planned covariance analysis and will be analyzed together with the post-test after the course concludes; all quantitative evidence reported in this interim report is therefore self-reported. Analyses were conducted in Python 3.12 with SciPy, using two-tailed tests with α = .05.

3. Early Evidence from the First-Round Implementation

3.1. Baseline: Significant Differences in Prerequisite Foundations Across the Three Groups

In Week 1, 219 valid pre-course questionnaires were collected: 89 from the pilot group, 73 from the transfer group, and 57 from the regular comparison group. The groups were comparable in gender composition ($\chi^2(2) = 4.63$, $p = .099$) but differed sharply in prior exposure to statistics ($\chi^2(4) = 116.71$, $p < .001$): the proportion of students who had never studied statistics-related content was 7.9% in the pilot group, 3.5% in the regular comparison group, and 47.9% in the transfer group.

Self-rated prerequisite knowledge (1–5 scale; 1 = "completely unaware," 5 = "very familiar"; higher scores indicate better mastery) shows that all three groups fell between the anchors "aware" (2) and "moderate" (3) across the five dimensions of descriptive statistics, probability foundations, probability distributions, software tools, and programming and statistical software, but intergroup differences were evident (Table 1, Figure 1). Welch's t-tests on the five-item composite showed that the pilot group scored significantly lower than the regular comparison group, $t(133.3) = -2.60$, $p = .010$, mean difference = -0.31, 95% CI [-0.54, -0.07], Cohen's $d = -0.43$, whereas the transfer group scored significantly lower than both undergraduate groups (vs. pilot: $t(142.0) = 4.27$, $p < .001$, $d = 0.69$; vs. regular comparison: $t(127.2) = 6.45$, $p < .001$, $d = 1.10$). All pairwise differences remained significant after Holm correction for multiple comparisons, and Mann–Whitney U tests yielded the same significance pattern (all $p \leq .002$). The prerequisite scale showed good internal consistency in each group (Cronbach's $\alpha = .86$, .87, and .92 for Groups S, D, and Z, respectively).

Table 1. Key Pre-Course Baseline Indicators of the Three Groups (Means, 1–5 Scale)

Indicator	Pilot group S (n = 89)	Transfer group Z (n = 73)	Comparison group D (n = 57)
Self-rated prerequisite knowledge (mean of 5 items)	2.54	1.97	2.84
Expectations for course content (mean of 5 items)	3.72	3.41	3.54
Anticipated learning difficulties (mean of 5 items)	3.42	3.58	3.25
Confidence in completing the course	3.13	2.82	3.16

Note. Prerequisite knowledge and confidence: 1 = lowest, 5 = highest. Expectations: 1 = "not at all expected," 5 = "very much expected." Difficulties: 1 = "not difficult at all," 5 = "very difficult." Values are recomputed from the raw response data.

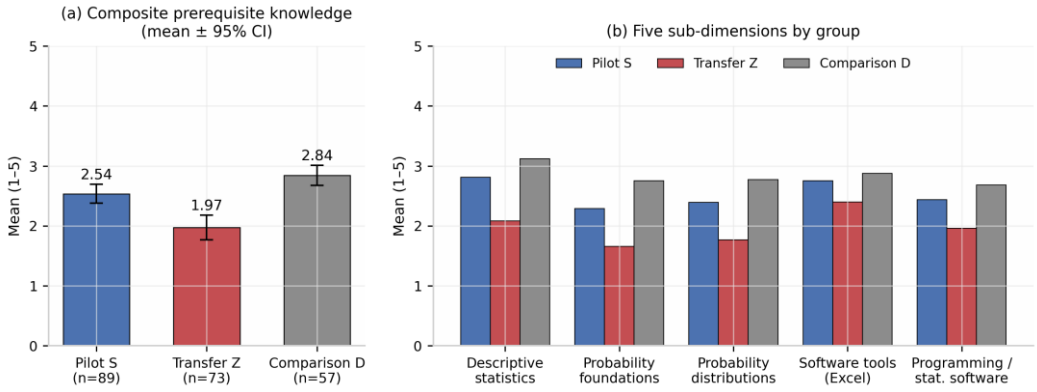


Figure 1. Pre-course Prerequisite Knowledge Across the Three Groups: (A) Composite Means with 95% Confidence Intervals; (B) Means of the Five Sub-Dimensions. Group S = Pilot Classes; Group Z = Transfer (Zhuanshengben) Comparison; Group D = Regular-Undergraduate Comparison.

The baseline differences suggest two points. First, the pilot and comparison groups are not strictly equivalent; post-course comparisons must therefore use pre-course levels

as covariates in analysis of covariance to avoid misreading starting-point differences as reform effects. Second, the learning discontinuity of the transfer group is quantitatively confirmed: 61.6% of this group had not taken statistics or mathematics courses during the junior-college stage, and their confidence in completing the course (2.82) was the lowest of the three groups (Kruskal–Wallis H = 6.41, p = .041; vs. pilot, p = .030; vs. regular comparison, p = .028). Expectation levels were otherwise comparable across groups (Kruskal–Wallis H = 5.57, p = .062), indicating that the groups differ in preparation rather than in motivation. This finding provides direct evidence for adjusting junior-college-to-undergraduate articulation teaching.

3.2. Pilot-Group Students' Expectations and Concerns about Project-Based Learning

Pilot-group students' expectations for project-based learning were generally high (mean of five items = 3.60, α = .89; Figure 2): the highest expectation was for "systematic use of SPSS in project tasks" (3.75), followed by "completing data collection and analysis through teamwork" (3.73) and "completing project tasks with authentic campus or enterprise data" (3.70), while "presenting learning outcomes through project reports and public defense" was significantly lower (3.31) than each of the other four items (paired t(88) = -5.49 to -2.42, all p < .05). Regarding expectations for course content (α = .93), "authentic case studies and enterprise project practice" (M = 3.84) scored significantly higher than "derivation and proof of statistical theory" (M = 3.52), paired t(88) = 4.27, p < .001, dz = 0.45, indicating that students' demand structure is consistent with the direction of the project-based reform. It should be noted, however, that pilot-group students were aware of participating in the reform pilot, so these expectations may be partly inflated by expectancy (Hawthorne-type) effects; the post-course questionnaire will provide a more sober reassessment.

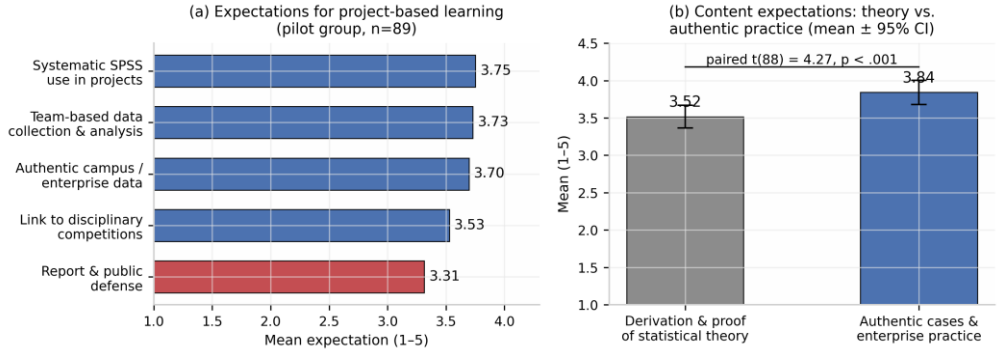


Figure 2. Pilot-group Expectations for Project-Based Learning: (A) Mean Ratings of the Five Expectation Items; (B) Preference for Authentic Case and Enterprise Project Practice over Derivation and Proof of Statistical Theory (Paired-Samples T-Test). Error Bars Show 95% Confidence Intervals.

Responses to open-ended items further corroborate this judgment. Among the "scenarios students hope course projects will involve," "auditing"-directly related to students' major-was mentioned 17 times, ranking first among all scenarios. Among "the greatest concerns about project-based learning," responses concentrated on worries about course difficulty and falling behind ("difficult," "don't know how," "failing the course," and similar statements totaled 28 mentions), as well as team coordination issues. This suggests that implementation needs to control the difficulty gradient of the first two projects, provide sufficient operational demonstrations, and curb free-riding through within-team peer evaluation and process portfolios.

3.3. Project Implementation and Representative Outcomes

The first-round implementation completed the teaching of Project 1 (statistical survey) and Project 2 (data organization and presentation). Taking the Market Survey Report on Public Satisfaction with Geely Automobile completed by one pilot-class team as an example: the team independently designed a 14-item questionnaire containing single-

choice, multiple-choice, and matrix-scale items; collected 47 responses via Wenjuanxing (an online survey platform), yielding 25 valid responses (53.2%) after applying the validity-screening criteria [author to specify: e.g., minimum completion time and attention-check items]; completed analyses of brand-awareness channels, product evaluations, and purchase intention using frequency statistics and scale means; and submitted a complete deliverable package including a member division-of-labor table, work-process records, the questionnaire sample, and a data-analysis report. In terms of deliverable form, students were able to complete the full statistical-survey workflow of "designing a questionnaire–collecting data–organizing data–writing a report"; the report was structurally complete, and the methods applied corresponded accurately to the course knowledge modules, meeting the expected objectives of Project 1 at the "cognitive experience–imitative operation" stage. Team topics concentrated on campus consumption, major-related industries, and local enterprises, consistent with the distribution of scenarios students expected in the pre-course questionnaire.

Beyond the representative example, the instructor scored all 15 teams' deliverables for Projects 1 and 2 with the six-dimension rubric (100-point scale per dimension; Table 3). The overall mean increased from 81.7 (SD = 3.1) for Project 1 to 85.9 (SD = 5.4) for Project 2, paired $t(14) = 5.68$, $p < .001$, mean difference = 4.2, 95% CI [2.6, 5.7], $dz = 1.47$, with 14 of the 15 teams improving. The two pilot classes did not differ significantly on either project (Project 1: Welch $t(12.3) = -0.50$, $p = .625$; Project 2: Welch $t(12.5) = 0.36$, $p = .723$), again indicating consistent implementation across classes. Dimension-level means show that students performed best on software operation and report standardization-capacities directly trained in the tool-orientation module and the imitative stage-whereas result interpretation and decision recommendations remained the weakest dimensions, pointing to the instructional emphasis of the middle-stage projects.

Table 3. Instructor Rubric Scores for Team Deliverables of Projects 1 and 2 (100-Point Scale, Means of 15 Teams)

Rubric dimension	Project 1	Project 2
Data quality	81.3	85.4
Method selection	80.5	86.4
Software operation	84.7	88.0
Result interpretation	79.0	84.4
Decision recommendations	80.0	83.9
Report standardization	84.9	87.2
Total (mean of six dimensions)	81.7	85.9

Note. Each dimension is scored 0–100; the total is the unweighted mean of the six dimensions. Class 2 contributed 8 teams and Class 3 contributed 7 teams.

3.4. Peer Evaluation of Group Collaboration

After Project 1, the two pilot classes conducted anonymous within-team peer evaluation; 15 team reports covering all 89 enrolled students were available for analysis (each student rated by teammates on six dimensions using a 1–5 scale; 526 valid ratings per dimension). The overall mean was 4.93 (SD = 0.19), with 4.95 (SD = 0.12) for Class 2 and 4.90 (SD = 0.25) for Class 3; the difference between the two classes was not significant, Welch $t(67.1) = 1.17$, $p = .248$, suggesting consistent implementation across the two pilot classes. The means of the six dimensions all fell between 4.91 and 4.94 (Table 2), with a balanced distribution. The distribution of the 89 individual averages is shown in Figure 3.

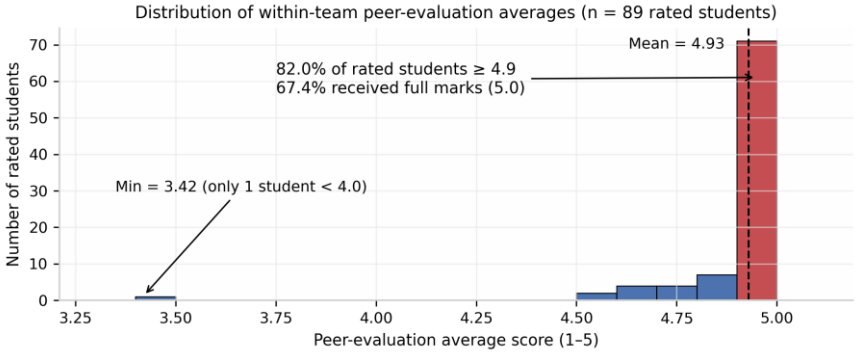


Figure 3. Distribution of Within-Team Peer-Evaluation Average Scores for Project 1 (N = 89 Rated Students), Showing a Marked Ceiling Effect.

Table 2. Dimension Means of Peer Evaluation for Project 1 (1–5 Scale, 526 Ratings Per Dimension)

Evaluation dimension	Mean	Evaluation dimension	Mean
Clear and reasonable division of work	4.92	Active participation in discussion	4.93
Timely completion of tasks	4.94	Contribution to deliverables	4.91
Substantive data contribution	4.92	Communication and collaboration attitude	4.94

Because peer ratings within a team are not independent and are subject to leniency effects [14, 15], these results are reported descriptively only and are triangulated with instructor rubric scores and process-portfolio records rather than subjected to inferential testing. Overall, the results indicate that teams functioned well and members were generally engaged in the first-round projects, but they also exhibit a clear ceiling effect: 82.0% of the rated students scored 4.9 or above, 67.4% received full marks, and the minimum was 3.42, leaving limited discrimination. This pattern is consistent with the leniency tendency documented in the peer-assessment literature [14, 15], and may have been amplified by the self-selected formation of teams, as teammates who choose one another are often acquaintances before rating; scoring anchors for peer evaluation will be further specified in subsequent projects. Consistent with this cautious positioning, team-level instructor rubric totals for Project 1 correlated only weakly and non-significantly with team peer-evaluation means (Spearman's $\rho = 0.17$, $p = .541$).

3.5. Inter-Group Peer Evaluation of Project Deliverables

After each project defense, teams rated the deliverables of every other team in their class on a 100-point scale; each team received seven ratings in Class 2 and six in Class 3 (98 ratings per project). For Project 1, team means averaged 90.02 (SD = 3.51; range 82.5–93.7); for Project 2, 90.80 (SD = 2.93; range 85.1–93.6), with no significant change across projects (paired $t(14) = 1.02$, $p = .324$; Table 4). Rater agreement within the same rated team was high (mean within-team SD of received ratings = 1.98 in Class 2, Project 1). Class 2 teams received higher mean ratings than Class 3 teams on both projects (Project 1: Welch $t(9.0) = 3.64$, $p = .005$; Project 2: Welch $t(6.7) = 3.28$, $p = .015$); because ratings were exchanged only within classes, this gap may partly reflect class-specific rating norms rather than deliverable quality alone. Compared with within-team peer evaluation, inter-group ratings discriminated substantially better among teams (team-mean range 82.5–93.7 versus 4.77–5.00 on the 5-point within-team scale; 53.3% of teams scored 90 or above in Project 1 versus 82.0% of students scoring 4.9 or above within teams), and team inter-group means were essentially uncorrelated with within-team peer means (Spearman's $\rho = -0.01$, $p = .974$), indicating that the two instruments capture complementary information about collaboration and product quality.

Table 4. Inter-group Peer Evaluation of Team Deliverables by Class and Project (100-Point Scale, Team Means)

Class	Project 1 M (SD)	Project 2 M (SD)
Class 2 (8 teams)	92.32 (1.71)	92.64 (0.77)
Class 3 (7 teams)	87.40 (3.20)	88.69 (3.11)
Total (15 teams)	90.02 (3.51)	90.80 (2.93)

Note. Each team rated all other teams in its own class after the project defense. SDs are across teams.

4. Discussion

First, the design direction of the project-based reform matches students' demand structure. The pre-course questionnaire shows that students' expectations for authentic cases and software practice are significantly higher than their expectations for theoretical derivation, echoing the GAISE recommendations [8] and prior findings on project-based statistics instruction [9–12]; the implementation process and deliverable forms of the first-round projects verify the feasibility of the "tool orientation + task progression" pathway. Students' relatively low expectations for public defense and their concerns about difficulty should be translated into concrete instructional arrangements, such as presentation training before the defense and a gentler task gradient. At the same time, expectancy (Hawthorne-type) effects cannot be excluded because students were aware of their pilot status, and conclusions about demand structure will be revisited against post-course data.

Second, non-equivalent baseline groups are a problem that quasi-experimental designs must confront directly. The pilot group's prerequisite foundation was in fact slightly lower than that of the regular comparison group, which means that if the pilot group is observed to perform better after the course, the evidentiary strength of that observation is actually higher; nevertheless, the normative approach remains to conduct intergroup comparisons with pre-course levels as covariates and to report baseline differences truthfully. Although the transfer group's data cannot be used to attribute effects to the reform directly, they independently reveal a structural articulation gap between junior-college and undergraduate curricula rather than individual student deficiency: 61.6% of transfer students had no exposure to statistics or mathematics coursework during the junior-college stage. This pattern mirrors the preparation gaps documented for vertical transfer students in other Asian systems [13] and suggests that course articulation between the two stages-rather than remedial attribution to students-is the appropriate unit of intervention.

Third, the effectiveness of process-assessment instruments needs to be refined through iteration. The ceiling effect in the first-round peer evaluation confirms that anonymous peer evaluation is more suitable as a means of qualitative corroboration and anomaly detection than as a precise scale for differentiating student performance [14, 15]; low-score cases will be checked against classroom observation records and reflective journals, and more fine-grained scoring anchors will be piloted in the second-round implementation. The inter-group ratings offer a useful complement: they discriminate among teams far better than within-team ratings and are nearly uncorrelated with them, yet the significant between-class gap indicates that peer-based ratings are anchored to local class norms; cross-class comparisons should therefore rely on instructor rubric scores rather than peer ratings.

Beyond this single implementation, three transferable design principles emerge: (1) tool-first sequencing-removing software barriers before project work begins; (2) progressive withdrawal of scaffolding across seven project packages aligned with knowledge modules; and (3) four-component diversified assessment that separates individual knowledge attainment from team deliverables. These principles can be adapted to other quantitative courses in application-oriented programs, subject to local student profiles such as the articulation gap documented here.

5. Conclusion

The project-based reform of the Applied Statistics course completed a systemic restructuring from "knowledge logic" to "task logic." The first-round implementation verified the feasibility of the "1 + 7 + N" system and the diversified assessment mechanism in an authentic classroom setting, and accumulated three types of evidence: learning baselines, project deliverables, and process evaluation. The present analysis is an interim report of the first-round implementation; rigorous testing of the reform's effects awaits the collection of post-course questionnaires, post-test knowledge assessments, and final grade data, followed by within-group paired comparisons and intergroup analysis of covariance. Across the remaining three semesters of the two-year implementation, the project tasks, assessment instruments, and teaching work packages will be iteratively revised on this basis, and subsequent rounds will examine the feasibility of generalizing the reform experience to transfer-student classes and other economics and management courses.

Ethics and consent: The study used anonymized data collected in routine course teaching; students were informed of the research purpose and provided consent.

Data availability: De-identified data are available from the author upon reasonable request.

Conflicts of interest: The author declares no conflict of interest.

Funding: This research was supported by the 2026 Education and Teaching Reform Research Project of Geely University (Key Project), "Research and Practice on the Project-Based Reform of the Applied Statistics Course" (Grant No. 2026XJG2017).

References

1. Ministry of Education of the People's Republic of China, "Opinions on deepening undergraduate education and teaching reform and comprehensively improving the quality of talent cultivation (Jiao Gao [2019] No. 6)," 2019.
2. Ministry of Education, National Development and Reform Commission, and Ministry of Finance of the People's Republic of China, "Guidelines on steering some local regular undergraduate institutions toward application-oriented transformation (Jiao Fa [2015] No. 7)," 2015.
3. D. Kokotsaki, V. Menzies, and A. Wiggins, "Project-based learning: A review of the literature," *Improving Schools*, vol. 19, no. 3, pp. 267–277, 2016.
4. C. H. Chen and Y. C. Yang, "Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators," *Educational Research Review*, vol. 26, pp. 71–81, 2019.
5. L. Zhang and Y. Ma, "A study of the impact of project-based learning on student learning effects: A meta-analysis study," *Frontiers in Psychology*, vol. 14, p. 1202728, 2023.
6. P. Guo, N. Saab, L. S. Post, and W. Admiraal, "A review of project-based learning in higher education: Student outcomes and measures," *International Journal of Educational Research*, vol. 102, p. 101586, 2020.
7. M. Maros, M. Korenkova, M. Fila, M. Levicky, et al., "Project-based learning and its effectiveness: Evidence from Slovakia," *Interactive Learning Environments*, 2023, doi: 10.1080/10494820.2021.1954036.
8. R. Carver, M. Everson, J. Gabrosek, N. Horton, R. Lock, M. Mocko, et al., *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016*. Alexandria, VA: American Statistical Association, 2016. [Online]. Available: http://www.amstat.org/education/gaise
9. L. Dierker, J. R. Evia, K. Singer-Freeman, K. Woods, J. Zupkus, A. Arnholt, et al., "Project-based learning in introductory statistics: Comparing course experiences and predicting positive outcomes for students from diverse educational settings," *International Journal of Educational Technology and Learning*, vol. 3, no. 2, pp. 52–64, 2018.
10. F. Farrell and M. Carr, "The effect of using a project-based learning (PBL) approach to improve engineering students' understanding of statistics," *Teaching Mathematics and Its Applications*, vol. 38, no. 3, pp. 135–145, 2019.
11. A. Cujba and M. Pifarré, "Enhancing students' attitudes towards statistics through innovative technology-enhanced, collaborative, and data-driven project-based learning," *Humanities and Social Sciences Communications*, 2024, doi: 10.1057/s41599-024-03469-5.
12. L. Al Labadi and A. Ly, "Enhancing statistics education through project-based learning (PBL) and the emergence of ChatGPT," *Teaching Statistics*, 2025, doi: 10.1111/test.12405.
13. K. Cheung, J. Ng, H. Tsang, and K. W. A. Pang, "Factors predicting vertical transfer students' GPA and degree attainment in an Asian educational context," *Studies in Educational Evaluation*, vol. 70, p. 101041, 2021.
14. B. Sridharan, J. Tai, and D. Boud, "Does the use of summative peer assessment in collaborative group work inhibit good judgement?," *Higher Education*, vol. 77, no. 5, pp. 853–870, 2019.

15. F. Zhang, C. D. Schunn, W. Li, and M. Long, "Changes in the reliability and validity of peer assessment across the college years," *Assessment & Evaluation in Higher Education*, vol. 45, no. 8, pp. 1073–1087, 2020.
16. S. Wang, "China's experience in the development of applied universities," *Journal of Research in Vocational Education*, 2022.
17. Geely University, "The Educational Philosophy of Geely University (Ji Xiao Zi [2026] No. 17)," 2026.
18. G. Wiggins and J. McTighe, *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development, 1998.
19. W. G. Spady, *Outcome-Based Education: Critical Issues and Answers*. Arlington, VA: American Association of School Administrators, 1994.
20. H. S. Barrows, "A taxonomy of problem-based learning methods," *Medical Education*, vol. 20, no. 6, pp. 481–486, 1986.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.