Construction and Practice of a Digital Evaluation Index System for Education and Teaching in the Era of Artificial Intelligence

Linshu Zhang, Yuping Yang, Kaimei Li
Chongqing Polytechnic University of Electronic Technology, Chongqing, China

Abstract: Integrating the characteristics of vocational education courses and industry demands, this study refines the score proportions of three-dimensional course objectives and develops a task-based evaluation rubric that runs through the entire teaching process. By collecting data from students' prior courses and related task performance and conducting Pearson correlation coefficient analysis, a scientific and rational value-added evaluation model is constructed. Using this model, students' individual growth curves are generated for predictive analysis; through comparing predicted values with actual values, the effectiveness of value-added evaluation is defined. This approach effectively addresses problems common in many evaluation systems, such as vague competency indicators, difficulty in quantifying comprehensive qualities, limited participant roles, insufficiently holistic evaluation perspectives, lack of value-added levels, and unstable incentive mechanisms. The evaluation system and platform, applied across multiple institutions, demonstrate significant advantages in improving teaching efficiency and optimizing student learning experiences. They successfully facilitate innovation and upgrading in teaching models across various courses and effectively enhance the learning motivation of students with differing starting points.

Keywords: Regression Analysis; Residual Model; Value-Added Evaluation; Visualization

1. Introduction

Comprehensive quality evaluation, as an important topic in the field of education, has long received considerable attention and research. It is explicitly stated in The Overall Plan for Deepening the Reform of Education Evaluation in the New Era that we must "adhere to scientific and effective principles, improve result evaluation, strengthen process evaluation, explore value-added evaluation, and improve comprehensive evaluation; fully leverage information technology to enhance the scientific, professional, and objective nature of education evaluation." The plan aims to break away from the long-standing unscientific practice of labeling students solely with exam scores, to improve the mechanisms for fostering integrity and cultivating talents, and to correct distorted evaluation orientations.

Driven by reforms in the education evaluation system of the new era, many higher education institutions have begun gradually implementing reforms in both student evaluation systems and teacher evaluation systems. These reforms no longer rely solely on traditional credits and grades as the only standards for evaluating students, nor do they assess teachers solely based on academic output or conventional performance measures. According to Jiang Jing[1], in terms of evaluation agents, comprehensive quality evaluation is showing a trend toward diversification—not only involving student self-evaluation, teacher evaluation, and parent evaluation, but also incorporating participation from employers and society as external evaluators. Li Yongyi[2] argues that with societal development and educational transformation, the content of comprehensive quality evaluation continues to expand, becoming richer and more inclusive, covering cultural literacy, skills competence, humanistic literacy, and innovative qualities. Xie Longjian[3] notes that existing evaluation methods are relatively crude, relying mainly on "academic + skills" examinations, which are insufficient to achieve comprehensive evaluation in the domains of "morality, intelligence, physical fitness, aesthetics, and labor." Therefore, innovative evaluation methods must be explored. According to Zhou Guangkuo[4], comprehensive quality evaluation requires clearly defined assessment indicators to more accurately measure students' overall competence and to address the new challenges faced in student affairs within vocational colleges under new concepts, new models, new teaching approaches, and new technological conditions, ultimately improving students' core vocational abilities and professionalism.

Based on the above findings, current research on comprehensive quality evaluation predominantly focuses on diversification of evaluation systems while neglecting dynamic value-added evaluation of vocational students' growth. Examining existing evaluation systems reveals that "quality" evaluation mainly emphasizes inputs in teaching resources, school conditions, or one-time examination results. Although related to educational quality, these factors do not address the core issue of education quality—the growth and achievements of students. They measure only the attainment of knowledge objectives and fail to assess skills objectives or competency objectives, resulting in a single evaluation agent and a limited evaluation perspective.

As an innovative form of educational assessment, value-added evaluation holds substantial significance in vocational education. It emphasizes students' progress and improvement, breaking away from traditional result-oriented evaluation models and paying greater attention to students' growth and development throughout the learning process. Value-added evaluation focuses on students' developmental outcomes and treats educational quality as a process measure. It is represented by changes in student quality, measured as the difference between a student's "initial quality" at the beginning of the learning process and their "current quality," namely, the value-added in student quality[5]. By determining the value-added based on characteristics of students' initial and current states, educators can assess the degree of student growth and progress. To scientifically measure student growth increments, the model must first clarify the components of student quality, including—but not limited to—knowledge mastery, skill proficiency, innovation capability, and labor literacy. Next, data must be collected at different time points to reflect students' initial and current states. Such data may be obtained through exams, assignment completion, project performance, and other forms of assessment.

Against this background, the research team selected automotive manufacturing–related majors as samples. Based on the characteristics of the major and industry requirements, the team refined the evaluation indicators and constructed a value-added evaluation system for vocational college students based on a regression analysis model. By building a scientific and reasonable evaluation system, the model provides a more comprehensive and objective reflection of students' learning outcomes and progress.

2. Analysis and Selection of the Value-Added Evaluation Model

2.1 Model Analysis

Three commonly used value-added evaluation models were comprehensively studied and assessed, and their characteristics are shown in Figure 1.

The Student Growth Percentile (SGP) Model provides quantitative and descriptive measurements of student progress. By integrating students' historical data, it determines each student's improvement by comparing them with a group of peers with similar academic backgrounds (students of the same type). If a student's improvement exceeds most peers, it indicates good performance[6]; conversely, if progress falls below most peers, it indicates unsatisfactory improvement. This model is suitable for large-scale datasets. It is especially applicable when the data sample is abundant, dimensions are broad, and the correlation among collected indicators is low—conditions under which a multiple linear regression model becomes difficult to construct.

The Multiple Linear Regression Model can reveal relationships among multiple variables and is suitable for evaluating the impact of various influencing factors on student performance[7]. This method applies to relatively fixed general education courses and foundational professional courses, where assessment methods and indicators remain consistent. After data accumulation across multiple cohorts, a highly accurate regression model can be established, enabling growth prediction and value-added evaluation for future students taking the same course.

The Residual Model uses multiple prior test scores as predictors for regression analysis, allowing accurate identification of differences between students' actual and expected performance in specific subjects[8]. This provides educators with targeted intervention measures. Each of the three models has its strengths and can be flexibly applied according to course characteristics and evaluation needs.

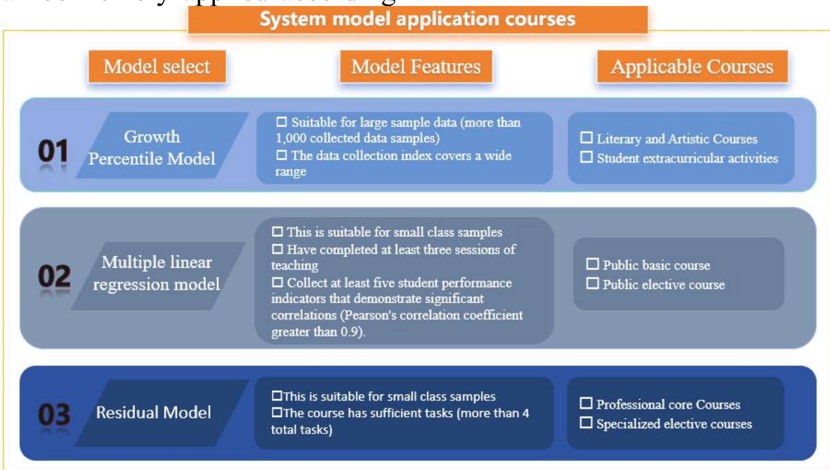

Figure 1. Model Selection and Applicable Course Types

2.2 Model Application

Based on the model characteristics, for small-scale practical courses in automotive manufacturing programs—where classes serve as the basic teaching unit and students enter with varied starting levels—the system backend adopts a residual model based on regression analysis to achieve scientific and balanced evaluation. The quantification process is illustrated in Figure 2. First, a course is divided into several module units, with each module containing 4 to 6 task items. After the completion of Module 1, multiple composite scores are generated. Two consecutive scores are defined as the entry score and the exit score, where the entry score corresponds to the x-axis value and the exit score corresponds to the y-axis value. Each score pair is plotted on the coordinate system, and a linear regression method is used to draw an individual student's growth curve.

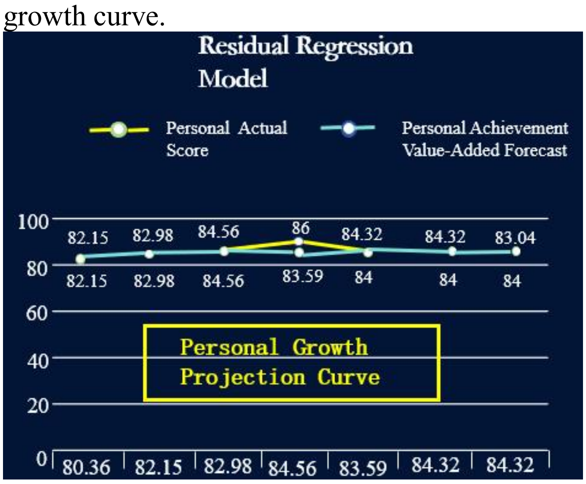

Figure 2. Residual Model Growth Prediction Diagram

The advantage of this value-added quantification scheme lies in its ability to effectively avoid the influence of differences in students' baseline abilities. By collecting data throughout the learning process, it provides continuous and intuitive descriptions of student growth. At the same time, since the residual model focuses solely on measuring each student's individual progress, it effectively reduces the impact of intra-class differences. Even students with weaker foundations can gain motivation from visualized incremental improvement, thereby enhancing their intrinsic learning drive.

3. Construction and Application of the Course Evaluation Index System

3.1 System Architecture Design

To achieve an effective value-added model design, this study focuses on diverse course characteristics and full-industry-chain job requirements. Centering on students as the primary subjects, the study formulates corresponding evaluation rubrics and collects evaluation data through multiple channels such as questionnaires, interviews, classroom observations, and platform-generated learning data. These multi-source data ensure comprehensive coverage of students' knowledge mastery, skill development, and overall quality improvement. For the collected data, Pearson correlation analysis is conducted, and based on the analytical results of student learning profiles, a multiple linear regression model or residual model is constructed. Utilizing the model, a dynamic tracking mechanism is established to visualize students' learning performance and growth trajectories at different stages[9], enabling more accurate assessment of their developmental outcomes. Evaluation results should be integrated with instructional feedback, offering teachers evidence to improve teaching methods and content, thereby forming a positive feedback loop. The system architecture design is shown in Figure 3.

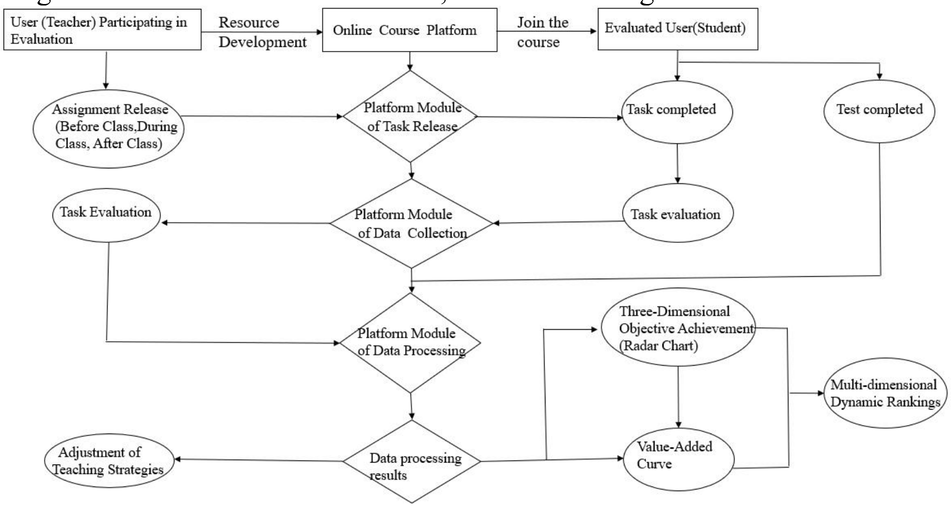

Figure 3. System Architecture Design

3.2 Development of Evaluation Rubrics

To achieve an effective value-added model design, the evaluation rubrics must first be developed in alignment with course characteristics and industry job requirements. Within the rubric, the weight of each indicator must be clearly defined to ensure its relevance to course content and real-world job skills[10]. The specific quantitative distribution of the indicators is shown in Figure 4.

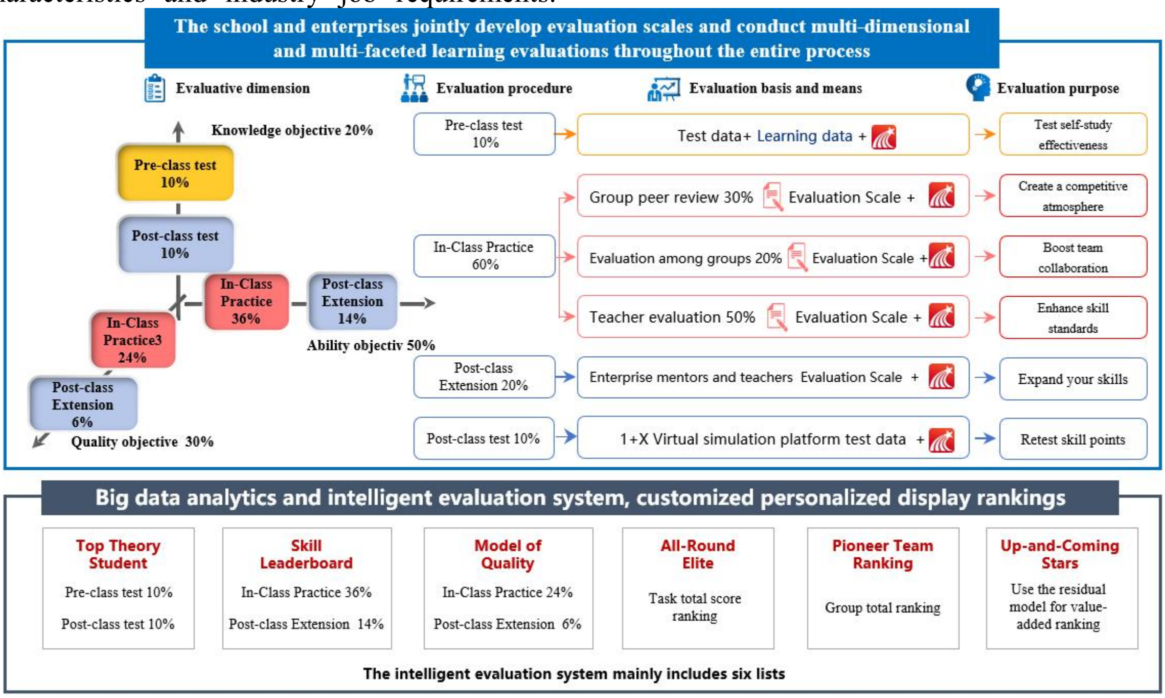

Figure 4. Indicator Quantification and Sample Evaluation Rubric

The evaluation rubric should contain quantifiable learning outcomes so that both teachers and students can clearly understand learning objectives and progress. Based on the different observation labels within the rubric, the system innovatively designs six different ranking lists—Theoretical Excellence, Skill Mastery, Quality Model, Pioneer Team, Rising Star, and All-Round Elite—to meet various learner needs and motivate students. Meanwhile, the design of the rubric should consider the characteristics of different disciplines to avoid constraining students with a single standard. Through periodic updates and optimization of the rubric, alignment with educational goals and industry trends can be ensured, providing students with more accurate and effective learning guidance.

3.3 System Construction and Practice

Using the core courses of the automotive manufacturing program at Chongqing Electronic Engineering Vocational University as an example, a school-enterprise collaborative task evaluation rubric was developed, refining the three-dimensional course objectives into 27 indicator observation points. Based on questionnaire and assessment results, the platform system analyzes the achievement of three-dimensional objectives at the overall, group, and individual levels in real time.

Before class, teachers adjust instructional focuses and strategies based on pre-task assessment results, designing interactive 3D game-based activities to vividly demonstrate internal principles, reduce conceptual difficulty, and increase student engagement. During class, teachers review task outcomes in real time, identify students who fall behind, track weak skill components, and—based on individual growth trends—help them reorganize task procedures and push micro-lesson videos as targeted support. After class, guided by the achievement levels of the three-dimensional objectives, group rankings, and leaderboard results, teachers assign three-tier extension tasks. For small-sample data at the class level, the system adopts the residual model algorithm to predict individual growth trends, producing a longitudinal value-added evaluation by comparing actual scores with predicted scores. This process generates six leaderboards such as the Rising Star list. Students who appear on these lists receive multiple forms of positive reinforcement, enhancing their sense of competition and achievement.

The system has been applied across multiple universities in Chongqing in both theoretical instruction and practical training platforms, covering core professional courses as well as general foundation courses. Several courses that adopted this system have won awards such as the First Prize in the Municipal Competition and the First Prize in the National Competition of the National Vocational College Teachers' Teaching Ability Contest. In practical teaching scenarios, the system has demonstrated significant advantages in improving teaching efficiency and optimizing student learning experiences.

4. Conclusion

Through an in-depth study of the student evaluation system in vocational colleges, this paper constructs a digital teaching evaluation index system based on an interactive value-added evaluation model for student groups, and applies it to the automotive manufacturing specialty as a practical sample. The system effectively addresses common issues in many evaluation systems, such as vague competency indicators, difficulty in quantifying comprehensive qualities, single-role participation, insufficient evaluation perspectives, absence of value-added tiers, and the inability to sustain incentive mechanisms. It provides strong support for improving the quality of education and teaching, and lays a solid foundation for further refinement of digital evaluation index systems. Future work will continue to enhance the construction of the evaluation system from the aspects of technological integration, course coverage, and system promotion and application.

Further exploration will focus on applying big data, artificial intelligence, and other technologies to the collection of raw assessment data. Algorithms will be used to conduct in-depth analysis of students' mastery of knowledge points, practical performance, and test results, thereby enabling more accurate baseline evaluation, needs prediction, and growth trend analysis.

Further exploration will focus on applying big data, artificial intelligence, and other technologies to the collection of raw assessment data. Algorithms will be used to conduct in-depth analysis of students' mastery of knowledge points, practical performance, and test results, thereby enabling more accurate baseline evaluation, needs prediction, and growth trend analysis.

Additionally, further research will be conducted on the replicability of the evaluation system. To facilitate wider application and dissemination of research outcomes, future work will focus on developing a more customizable upgraded system capable of automatically switching background settings, model selections, and visualization modes. Upon completion of the upgraded version, the system may be extended from institutional users such as schools to general users who require student learning data collection and processing.

Acknowledgements

This work was supported by the Chongqing Municipal Education Science Planning 2023 General Project "Research and Practice of Interactive Value-Added Evaluation System for Student Groups in Vocational Institutions" (Project Number: K23YG3090345), the Chongqing Municipal Education Science Planning 2022 General Project "Research on the Construction of an Integrated Model for Comprehensive Quality Evaluation of Higher Vocational Students in the New Era" (Project Number: K22YG309304), and the Key Project of Chongqing Vocational Education Teaching Reform Research "Research on the Construction and Application Model of Industry-Education Integration Textbooks in Vocational Undergraduate Education" (Project Number: Z2241064).

References

[1] Jiang Jing. Research on the Construction and Application Strategy of a Comprehensive Evaluation System for Students in Vocational Colleges. Jiangsu Science & Technology Information, 2024, 41(9):52–55. DOI:10.3969/j.issn.1004-7530.2024.09.015.

[2] Li Yongyi, Chen Mingxiang. Construction and Analysis of a Value-Added Evaluation Index System for Higher Vocational Students Under High-Quality Development. 2024(1):55–61.

[3] Xie Longjian, Tian Tengfei. Constructing a Quality Assurance and Evaluation System to Promote the Connotative Development of Vocational Education—Exploration of Quality Assurance and Evaluation Practices in Chongqing Vocational Education. Contemporary Vocational Education, 2018(5):6. DOI:10.16851/j.cnki.51-1728/g4.20180925.014.

[4] Zhou Guangkuo, Wu Gaochao. Research on a Comprehensive Quality Evaluation System for Higher Vocational Students Based on the Cultivation of Core Vocational Competencies. Journal of Yangling Vocational & Technical College, 2023, 22(01):33–38. DOI:10.19859/j.cnki.cn61-1403/G4.2023.01.009

[5] Hao Qilei, Liang Jinmiao, Li Guanghai. Review, Reflection, and Prospect of Research on Student Growth Value-Added Evaluation. Modern Education Management, 2024(7):63–73. DOI:10.16697/j.1674-5485.2024.07.007.

[6] Hu Qing, Qin Chao. Design and Implementation of a Blended Teaching Value-Added Evaluation System Based on Big Data//Proceedings of the 2023 Annual Conference of the China Tao Xingzhi Research Association (Vol. 5). 2023.

[7] Zhou Peng. Research on the Prediction of Minor Fault Extension Length Based on Multiple Linear Regression and Its Effect Evaluation. Inner Mongolia Coal Economy, 2024(22):1–4. DOI:10.13487/j.cnki.imce.026082.

[8] Chen Yao. Design and Application of an Interdisciplinary Thematic Learning Framework Based on Subject Knowledge Graphs. Southwest University, 2024. DOI:10.27684/d.cnki.gxndx.2024.000403.

[9] Li Yifan, Xiang Yongyong, Shi Luojie, et al. Efficient Reliability Analysis Method Based on Nonlinear Autoregressive Multi-Fidelity Surrogate Model and Active Learning (English). Journal of Zhejiang University–Science A (Applied Physics & Engineering), 2024, 25(11):922–938.

[10] Huang Weicheng, Wang Huiying. Application of a Value-Added Evaluation Model for Project-Based Learning Oriented Toward Core Literacy in Higher Vocational Education—A Case Study of “International Business Document Practice”. Journal of Heilongjiang Ecological Engineering Vocational College, 2025, 38(01):123–130.

