*Occupation and Professional Education Vol. 2 No. 6, 2025*

# Construction of a Quality Evaluation System for After-School Physical Education Services in Primary and Secondary Schools under the "Double Reduction" Policy

**Qingpeng Cai***
*Bai'an Primary School, Wanzhou District, Chongqing, China*
**Corresponding Author*

**Abstract: Under the "Double Reduction" policy, after-school physical education (PE) services in primary and secondary schools have become a key approach to improving students' physical health and promoting holistic development. However, the absence of a scientific and systematic quality evaluation system has hindered the standardization and improvement of such services. This study aims to develop a comprehensive quality evaluation system tailored to the "Double Reduction" policy context. The research adopts a literature analysis method to review relevant policies and studies on after-school PE services and quality evaluation. The Delphi method is used to consult experts in physical education and educational management to refine and finalize evaluation indicators, while the Analytic Hierarchy Process (AHP) is employed to determine indicator weights. Based on policy requirements, student needs, and service practices, the evaluation system includes dimensions such as service content, faculty quality, facility utilization, organizational management, and student feedback. The findings indicate that the proposed evaluation system is scientific and feasible, providing a comprehensive framework for objectively assessing service quality. This study contributes to improving after-school PE services, optimizing management mechanisms, and offering theoretical and practical guidance for enhancing their effectiveness.**

**Keywords: "Double Reduction" Policy; Primary and Secondary Schools; After-School Pe Services; Quality Evaluation System; Indicator Development**

## 1. Introduction

### 1.1 Research Background and Significance

Reforms in education have always aligned with societal development demands, with policy adjustments in basic education playing a crucial role in talent cultivation and improving national quality. The introduction of the "Double Reduction" policy marks a new phase of intrinsic development for China's basic education, aiming to reduce students' excessive academic workload and off-campus tutoring burdens. This policy seeks to create an environment conducive to students' well-rounded growth and health.

Physical education, a vital component of holistic education, has gained renewed emphasis under the "Double Reduction" policy. The policy highlights the importance of strengthening school-based physical activities, ensuring adequate student participation in physical education, and expanding the breadth and depth of PE services via after-school programs. Consequently, after-school PE services have become a critical vehicle for implementing the "Health First" educational philosophy.

While the scope and diversity of after-school PE services have expanded significantly, issues related to their quality have emerged. Problems such as homogeneous service content, insufficient teacher expertise, underutilized facilities, and inadequate safety management have hindered the achievement of policy goals. A primary cause of these challenges is the lack of a systematic quality evaluation framework, which leads to unclear standards for assessing performance, limits systemic improvement, and impedes effective oversight by management authorities.

Developing a comprehensive quality evaluation system for after-school PE services under the "Double Reduction" policy carries significant theoretical and practical implications. Theoretically, it enriches the application of educational evaluation theories in physical education, enhances the theoretical foundation

for school PE management, and offers conceptual and methodological references for future research. Practically, it provides tools for educational authorities to monitor service implementation, supports schools in identifying weaknesses and optimizing resource allocation, and enables teachers to improve service delivery, thus enhancing students' physical fitness and overall development.

### 1.2 Literature Review

Research on after-school PE services in China has gained momentum with the implementation of the "Double Reduction" policy, encompassing multiple dimensions:

Policy Implementation: Studies have explored how after-school PE services align with the "Double Reduction" policy, analyzing policy requirements and regional implementation disparities. While economically developed regions display greater diversity in service provision, underdeveloped areas struggle with resource limitations, affecting service quality.

Service Models: Research has highlighted various service models, including school-run sports clubs, collaborations with external coaches, and community-shared facilities. However, most studies rely on qualitative evaluations, lacking systematic and quantitative comparisons of model effectiveness.

Quality Evaluation: Existing studies often focus on specific dimensions, such as teacher qualifications or student satisfaction, without developing comprehensive evaluation frameworks. Additionally, the application of advanced evaluation methods, such as Delphi and AHP, remains limited.

International studies on after-school services, particularly in the U.S. and Japan, provide valuable insights. Programs like the U.S. "21st Century Community Learning Centers" emphasize safety, participation, and social impact, while Japan's "After-School Children's Classes" integrate sports with regional culture. However, these systems are shaped by distinct educational and policy contexts, limiting their direct applicability to China.

Although prior research offers foundational insights, the lack of a systematic and locally adapted quality evaluation system for after-school PE services remains a critical gap.

### 1.3 Research Methodology

This study adopts a systematic approach that integrates qualitative and quantitative methods:

Literature Analysis: A comprehensive review of "Double Reduction" policy documents, related studies on after-school PE services, and quality evaluation theories provides a theoretical basis for the study.

Delphi Method: Experts in physical education, educational management, and frontline teaching refine and finalize evaluation indicators through multiple rounds of consultation.

Analytic Hierarchy Process (AHP): This method is employed to determine the weight of each indicator by constructing judgment matrices and calculating eigenvalues.

Logical Analysis: This ensures the conceptual clarity, theoretical consistency, and practical relevance of the evaluation framework.

### 1.4 Innovation and Limitations

Innovation:

Policy Relevance: The study aligns the evaluation system with the "Double Reduction" policy, addressing the unique needs of this reform.

Comprehensive Framework: The evaluation system incorporates multiple dimensions, covering all aspects of after-school PE services, and balances qualitative and quantitative analysis.

Methodological Integration: By combining the Delphi method and AHP, the study ensures both expert-driven refinement and data-driven precision.

Limitations:

The empirical validation of the framework is limited in scope and may not fully account for regional and school-specific differences.

Indicator weights, though calculated scientifically, are influenced by expert subjectivity.

The dynamic nature of service quality necessitates ongoing updates to the framework, which this study cannot address in the short term.

## 2. Definition of Key Concepts and Theoretical Foundations

### 2.1 Core Connotations of the "Double Reduction" Policy

The "Double Reduction" policy represents a systematic reconstruction of the basic education ecosystem, aiming to achieve multiple objectives and return to the fundamental goals of education. While the policy directly targets reducing the

excessive academic workload and off-campus tutoring burdens on students, its deeper objectives include correcting the utilitarian trends in education, promoting equity, and building a holistic education system that integrates moral, intellectual, physical, aesthetic, and labor education.

In terms of regulating academic burdens, the policy sets explicit limits on homework quantity and time while emphasizing the quality of homework design. It discourages repetitive and mechanical tasks, allowing students sufficient time for rest and self-development. The policy does not intend to weaken academic education but to reform teaching approaches, enhancing classroom efficiency and achieving a “reduction in quantity with an increase in quality.” The governance of off-campus tutoring focuses on regulating the behavior of tutoring institutions and curbing excessive marketization in education by restricting academic-oriented training and promoting non-academic programs. These measures aim to re-establish the school as the primary site of education.

The policy places greater demands on schools to take responsibility for after-school services, offering enriched and high-quality programs to meet diverse student development needs. After-school physical education (PE) services, as a critical component, are assigned special responsibilities. Schools are required to ensure daily physical activity within school hours and provide diverse sports programs to cultivate students' interest in sports, enhance their physical health, and promote their holistic development. This elevates PE services from a peripheral "extracurricular supplement" to a "core educational activity," making it a vital vehicle for the policy's emphasis on balanced and comprehensive education.

### 2.2 Definition and Characteristics of After-School Physical Education Services

After-school physical education services refer to school-led PE programs provided to primary and secondary students after regular school hours. These programs rely on school resources or external partnerships to offer physical education services. They differ from regular PE classes, which follow curriculum standards with systematic and mandatory characteristics, and from extracurricular sports activities, which are usually self-organized and lack structured guidance. After-school PE services occupy a middle ground, characterized by a balance of educational structure and student autonomy. Guided by policy requirements and student interests, these services aim to expand the space for physical education and meet students' individual needs.

Following the implementation of the "Double Reduction" policy, after-school PE services have undergone significant changes:

The duration of services has increased, with schools offering programs several times a week, sometimes achieving full coverage.

The service providers have diversified, involving not only PE teachers but also retired athletes, social sports instructors, and university students majoring in sports.

The services are increasingly linked to comprehensive quality evaluations, with some regions incorporating students' performance in after-school PE services into their overall assessments, encouraging greater participation.

### 2.3 Theoretical Foundations for Quality Evaluation Systems

The construction of a quality evaluation system requires a robust theoretical foundation. This study integrates four key theories: service quality theory, total quality management theory, customer satisfaction theory, and educational evaluation theory.

Service Quality Theory

This theory provides a framework for designing evaluation dimensions. It defines service quality as the gap between perceived and expected services, encompassing five dimensions: tangibility, reliability, responsiveness, assurance, and empathy. Applied to after-school PE services:

Tangibility refers to physical facilities, equipment, and resources.

Reliability addresses the stability of service content and safety measures.

Responsiveness reflects how services meet student needs and adapt to feedback.

Assurance evaluates the professional qualifications of providers and adherence to standards.

Empathy considers respect for and accommodation of individual student needs.

These dimensions comprehensively capture the characteristics of service quality in after-school PE programs.

Total Quality Management Theory (TQM)

TQM emphasizes quality control across the

entire service process, guiding the systemic design of the evaluation system. It advocates for monitoring all aspects of service, including planning, implementation, and feedback, and involving all stakeholders. Applied to after-school PE, TQM suggests evaluating not only outcomes but also planning, resource allocation, and management processes, incorporating feedback from students, parents, and administrators to form a comprehensive evaluation network.

Customer Satisfaction Theory

This theory highlights customer satisfaction as a key indicator of service quality. In after-school PE services, students and parents are the primary "customers," and their satisfaction reflects actual service quality. Evaluation indicators should include students' satisfaction with content and organization, and parents' approval of safety measures and educational value. By integrating customer feedback, the system can enhance objectivity and relevance.

Educational Evaluation Theory

This theory ensures the system's alignment with educational goals and principles. It emphasizes that educational evaluations should reflect broader developmental objectives. For after-school PE services, indicators should measure improvements in students' physical fitness, interest in sports, and sports literacy, rather than focusing solely on athletic performance. This ensures that the evaluation remains educationally meaningful and developmentally oriented.

By integrating these theories, the evaluation system combines general service quality principles with the specific needs of the educational context. It balances hardware and process evaluation with the assessment of educational outcomes, providing a comprehensive theoretical foundation for the system's construction.

## 3. Principles and Framework of the Quality Evaluation System for After-School Physical Education Services

### 3.1 Principles for Constructing the Evaluation System

The evaluation system must adhere to key principles to ensure its scientific, reasonable, and practical applicability:

(1) Guidance Principle: The indicators must align with the "Double Reduction" policy and holistic education goals, prioritizing quality improvement over superficial service expansion.

(2) Scientific Principle: The system should be based on sound theories and student development principles, with clearly defined, operable indicators and scientifically determined weights.

(3) Systematic Principle: The system must cover all elements of after-school PE services, from planning and resource preparation to implementation and feedback, ensuring comprehensive evaluation.

(4) Practical Principle: The indicators should be actionable and measurable, avoiding overly abstract or complex metrics to facilitate straightforward application.

(5) Dynamic Principle: The system should remain flexible and adaptive to changes in student needs, policy environments, and service innovations, allowing for regular updates.

### 3.2 Framework Design of the Evaluation System

The evaluation system adopts a "three-dimension, multi-level" structure, encompassing three primary dimensions—service foundation, service process, and service outcomes—with multiple secondary indicators under each dimension.

Service Foundation Dimension: Focuses on prerequisites for implementing after-school PE services, covering:

Faculty Quality (teacher qualifications, training, experience).

Facilities and Equipment (availability, safety, and appropriateness).

Financial Support (stability, allocation, and efficiency).

Service Process Dimension: Evaluates organizational and operational quality, including:

Service Content (diversity, relevance, and educational value).

Organizational Management (efficiency, communication, and planning).

Safety Measures (protocols, inspections, and incident management).

Service Outcome Dimension: Measures the effectiveness of services, including:

Student Development (improvements in fitness, skills, and team spirit).

Parent Feedback (perceptions of quality and value).

Social Recognition (media reviews, external evaluations, and impact).

This framework aligns inputs (resources),

processes (implementation), and outputs (results), forming a comprehensive evaluation system that ensures practical applicability and systematic assessment.

## 4. Screening and Establishment of Quality Evaluation Indicators for After-School Physical Education Services in Primary and Secondary Schools

### 4.1 Establishment of the Preliminary Indicator Pool

The establishment of the initial indicator pool is a fundamental step in the construction of the evaluation system, and its quality directly affects the effectiveness of subsequent indicator screening and the scientific nature of the evaluation system. This study integrates information through multiple channels and dimensions to form a preliminary indicator pool covering all elements of services, ensuring the comprehensiveness and representativeness of the indicators.

The sources of the indicator pool mainly include three aspects. Firstly, theoretical indicators formed through literature review. Systematically review the research achievements in the fields of after-school physical education services, educational evaluation, and service quality, extract the evaluation indicators mentioned in relevant studies, such as the professional background of teachers, course diversity, and changes in students' physical fitness, and through concept integration and semantic standardization, form a set of indicators at the theoretical level. Secondly, the core indicators for the analysis and refinement of policy texts. Conduct a coding analysis of the requirements for after-school physical education services in the "Double Reduction" policy and its supporting documents, and transform policy expressions such as "ensuring activity time", "enriching activity content", and "guaranteeing activity safety" into operational evaluation indicators, such as "weekly service duration compliance rate", "number of sports events", and "completeness of safety plans", etc. Ensure the consistency between indicators and policy guidance. Thirdly, the realistic indicators obtained through practical research. Through semi-structured interviews with physical education teachers, researchers and parents in primary and secondary schools, the quality elements of concern in practice were collected, such as "students' participation enthusiasm", "timeliness of communication between home and school", "frequency of equipment update", etc., to supplement the practical details not covered by theoretical and policy indicators.

The construction of the indicator pool follows the "three-level indicator" design logic. Under the three first-level dimensions of service foundation, service process, and service effect, the second-level indicators are refined into measurable third-level indicators. In the dimension of service foundation, "teaching staff" is detailed into "the proportion of full-time physical education teachers", "the certification rate of part-time coaches", "the average annual training duration of teachers", etc. "Venue facilities" include "per-student sports area", "equipment integrity rate", "facility suitability for special groups", etc. "Funding guarantee" encompasses "per-student funding standards", "transparency of funding usage", "the number of resource crowdfunding channels", etc. In the dimension of the service process, "service content" is decomposed into "project update cycle", "proportion of stratified teaching implementation", "frequency of health education integration", etc. "Organizational management" includes "completeness of service plans", "frequency of communication between home and school", "student attendance rate", etc. "Safety assurance" is detailed into "safety inspection cycle", "emergency equipment provision rate", "accident handling satisfaction", etc. In the dimension of service effectiveness, "student development" includes "the increase in the pass rate of physical fitness tests", "the pass rate of sports skill levels", "the duration of sports interest", etc. "Parent Feedback" covers "service satisfaction", "safety recognition", "educational value recognition", etc. "Social recognition" includes "the number of regional demonstration cases", "the frequency of positive media coverage", "the assessment grade by the education department", etc.

To ensure the operability of the indicators, operational definitions are provided for each third-level indicator, clearly defining the measurement methods and data sources. For instance, the "proportion of full-time physical education teachers" is defined as "the ratio of the number of full-time physical education teachers participating in after-school services to the total number of service teachers", which is verified through the school's personnel file. The

"increase in the rate of passing the physical fitness test" refers to the "difference in the proportion of students who passed the physical fitness test in two consecutive times before and after participating in the service", which is obtained through the student physical health monitoring system. After the indicator pool was formed, three experts in the field of physical education were organized to conduct the first round of pre-review. Duplicate indicators were deleted, similar indicators were merged, and those with ambiguous expressions were revised. Eventually, a preliminary indicator pool containing 45 third-level indicators was formed, laying the foundation for subsequent screening.

### 4.2 Index Screening and Revision Based on the Delphi Method

The Delphi method integrates expert opinions through multiple rounds of anonymous consultation and is suitable for the index screening process that lacks objective data support. This study uses this method to optimize the initial indicator pool, enhancing the scientificity and consensus of the indicators.

The selection of experts focuses on representativeness and authority, taking into account professional background, practical experience and regional distribution comprehensively. Twenty experts will be selected from sports colleges and universities, educational research institutions, primary and secondary schools, and educational administrative departments, among whom 30% are professors of physical education, 25% are special-grade physical education teachers in primary and secondary schools, 20% are teaching researchers, and 25% are educational management cadres. The average working experience of the experts exceeds 15 years. They are all familiar with the "double reduction" policy and the practice of after-school sports services, covering different regions in the eastern, central and western regions, ensuring the diversity of opinions.

The consultation process is carried out in three rounds. The first round of consultation will present the initial indicator pool to the experts via email, explaining the source and operational definition of the indicators. Experts will be invited to score them on a 5-point scale from three dimensions: "necessity", "feasibility", and "accuracy", and offer suggestions for revision. A total of 18 valid questionnaires were collected. The statistics showed that among the 45 indicators, 6 had a necessity score lower than 3.5, mainly involving indicators with relatively low attention in practice such as "the number of crowdfunding channels for funds" and "the frequency of positive media coverage". These were deleted in accordance with expert suggestions. At the same time, three new indicators, namely "Integration Degree of ideological and political Education in courses" and "Proportion of students' autonomous activities", were added based on the suggestions, forming the second round of indicator pool with 42 indicators.

The second round of consultation focused on the importance ranking of the adjusted indicator pool, requiring experts to rank the indicators from 1 to 42 in terms of importance and provide specific revision suggestions for the disputed indicators. A total of 19 valid questionnaires were retrieved. By calculating the Kendall's coefficient of Concordat ($W=0.72$, $P<0.01$), it was determined that the expert opinions were highly coordinated. A re-evaluation was conducted on the bottom 10% of the indicators (4 items). Based on expert feedback, "Equipment integrity rate" and "equipment update frequency" were merged into "equipment guarantee level", and "number of regional demonstration cases" was removed. Ultimately, 38 indicators were retained.

The third round of consultation presented the results of the second round of adjustments to the experts, confirming the rationality and completeness of the indicators. Among the 17 experts' feedback, 15 expressed their approval, and 2 suggested refining the operational definition of the "implementation ratio of stratified teaching". Based on this, the indicator was clearly defined as "the proportion of courses designed with differentiated content for students with different athletic abilities", forming the final 38 evaluation indicators. After three rounds of consultation, the expert opinion coordination coefficient increased to 0.81 ($P<0.01$), indicating that the index screening results have a high degree of consensus and provide a reliable index basis for the construction of the evaluation system.

## 5. Determination of the Weights of Evaluation Indicators for the Quality of After-School Physical Education Services in Primary and Secondary Schools

### 5.1 Application Principles of the Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) decomposes complex decision-making problems into ordered levels and determines the weights of indicators through pairwise comparisons. It is suitable for multi-dimensional and multi-indicator weight allocation scenarios. The core logic is to transform qualitative judgments into quantitative calculations, and on the premise of maintaining consistency in decision-making thinking, enhance the scientific nature of weight determination.

The implementation steps of this method include four stages. First, construct a hierarchical structure, decomposing the target layer, criterion layer, and indicator layer step by step, and clearly defining the affiliation relationships among each level. In this study, the target layer is "Quality Evaluation of after-School Physical Education Services in Primary and Secondary Schools", the criterion layer corresponds to three first-level dimensions: service foundation, service process, and service effect, and the indicator layer includes 38 third-level indicators that have been screened. Secondly, a judgment matrix is constructed. Experts are invited to make pairwise comparisons of the importance of indicators at the same level relative to a certain element at the previous level, and values are assigned using the 1-9 scale method (1 indicates equal importance, and 9 indicates extremely important). The matrix element a_ij satisfies a_ij=1/a_ji, and the diagonal element is 1.

Subsequently, hierarchical single sorting and consistency checks are carried out. By calculating the maximum eigenvalue λ_max of the judgment matrix and its corresponding eigenvector, the relative weights of the indicators at the same level are obtained. Consistency checking is achieved by calculating the consistency index CI= (λ_max-n)/(n-1) (where n is the order of the matrix) and the consistency ratio CR=CI/RI (where RI is the mean random consistency index). When CR<0.1, it is judged that the matrix meets the consistency requirements; otherwise, the matrix elements need to be adjusted. Finally, perform a total hierarchical ranking, calculate the combined weight of the lowest-level indicators relative to the target level, and verify the consistency of the total ranking to ensure the rationality of the overall weight distribution.

The applicability of the Analytic Hierarchy Process (AHP) in this study is reflected in three aspects. Firstly, it conforms to the hierarchical structure of the evaluation indicators and can reflect the affiliation relationship between different dimensions and indicators. Secondly, by combining the subjective judgment of experts with quantitative calculation, the subjectivity and objectivity of the judgment of the importance of indicators have been balanced. Thirdly, the consistency check mechanism can effectively avoid logical contradictions in expert judgments and enhance the reliability of the weight results. The application of this method assigns reasonable weight values to the evaluation indicators, making the evaluation system not only include "what to evaluate", but also clarify "how to weight", thereby enhancing the discrimination and persuasiveness of the evaluation results.

### 5.2 Calculation of Index Weights Based on the Analytic Hierarchy Process

The calculation of index weights is a key link in applying the Analytic Hierarchy Process (AHP) to the evaluation system. The process must strictly follow the method norms to ensure the rationality and consistency of weight distribution. First, construct the judgment matrix. Ten core experts (including representatives from various fields) who participated in the Delphi method consultation were invited to conduct pairwise comparisons and score the indicators. For the criterion layer, experts need to assess the significance of the service foundation, service process, and service effect relative to the overall objective. Statistics show that the service process (with an average score of 5.2) is considered more important than the service foundation (3.8), followed by the service effect (4.5). Based on this, the criterion layer judgment matrix is constructed. The indicator layer judgment matrix is constructed respectively according to the affiliation relationship. For example, under the service foundation dimension, the relative importance of the teaching staff, venue facilities, and financial guarantee is independently evaluated by experts, forming three third-order matrices. The service process and service effect dimensions respectively form three and three judgment matrices (corresponding to their respective secondary indicators).

The weight calculation adopts a combination of the sum method and the root method. The sum

method normalizes the columns of the judgment matrix, sums them row by row, and then normalizes them to obtain the weight vector. The root method calculates the NTH root of the product of elements in each row of the matrix and normalizes it to obtain the weight. The average of the results of the two methods is taken as the final weight to reduce the error of a single method. Taking the criterion layer as an example, the calculated service base weight is 0.25, the service process weight is 0.40, and the service effect weight is 0.35, indicating that experts believe that the service implementation process has the greatest impact on quality.

The consistency test results show that the CR values of all judgment matrices are less than 0.1, meeting the consistency requirements. Among them, the CR of the criterion layer was 0.07, the CR of the "Safety Assurance" index matrix under the service process dimension was 0.05, and the CR of the "student Development" index matrix under the service effect dimension was 0.09, all of which passed the inspection. Among the weights of the indicator layer, indicators such as "the completeness of the safety management system" (0.12), "the increase in the pass rate of physical fitness tests" (0.10), and "the proportion of full-time physical education teachers" (0.08) have relatively high weights, reflecting experts' emphasis on service safety, student physical fitness improvement, and the professionalism of teachers. The weights of indicators such as "Transparency of fund usage" (0.03) and "Frequency of positive media coverage" (0.02) are relatively low, which is consistent with the implicit characteristics of resource input and the indirectness of social evaluation in practice.

The calculation results of the combined weights show that the weight values of the 38 indicators range from 0.01 to 0.12, presenting a gradient distribution. The cumulative weight of the service process dimension (0.40) is higher than that of other dimensions, among which "diversity of service content", "security guarantee measures" and "organizational management efficiency" contribute the most. The cumulative weight of the "student development" related indicators in the service effect dimension reaches 0.28, highlighting the student-centered evaluation orientation. The results of weight distribution not only reflect the policy's requirements for the standardization of the service process, but also demonstrate the educational attribute of after-school physical education services, which takes students' development as the fundamental goal, providing a quantitative basis for the practical application of the evaluation system.

## 6. Empirical Validation of the Quality Evaluation System for After-School Physical Education Services in Primary and Secondary Schools

### 6.1 Empirical Subjects and Data Collection

The empirical validation aims to assess the applicability and effectiveness of the evaluation system using real-world data, providing a practical foundation for its refinement. A stratified sampling approach was adopted, selecting 20 schools from different regions (eastern, central, and western) and different types of areas (urban, suburban, and rural), ensuring diversity and representativeness. The sample includes 12 primary schools and 8 secondary schools, all of which have been conducting after-school physical education services for over a year. Various service modes, such as self-managed, school-enterprise partnerships, and community collaboration, were included to reflect current practices comprehensively.

Data collection employed multiple methods to ensure comprehensiveness and accuracy.

Questionnaires: Student questionnaires (38 indicators, 5-point Likert scale) were distributed to 1,200 students (20% randomly sampled per school), with 1,086 valid responses (90.5% response rate). Parent questionnaires (focused on safety and educational value) were distributed to 600 individuals, with 528 valid responses (88.0% response rate).

Interviews: Semi-structured interviews with 1 PE department leader and 2 frontline teachers per school covered topics like indicator clarity, data accessibility, and evaluation results, generating over 400,000 words of transcripts.

Observations: Structured observation forms were used to document service processes, teacher-student interactions, and facility usage, with over 300 hours of observation conducted on-site.

Document Analysis: Data such as teacher resources, funding reports, and fitness test results were collected to cross-verify questionnaire and interview findings.

To control for errors, standardized investigator training, strict data collection procedures, and

double-entry data processing were employed. All data were analyzed using SPSS for reliability and validity tests.

### 6.2 Reliability and Validity Testing of the Evaluation System

Reliability and validity are critical for ensuring the quality of the evaluation system.

Reliability Testing: Internal consistency was assessed using Cronbach's α. The total questionnaire's α was 0.92, indicating high reliability. The three primary dimensions—service foundation (α=0.86), service process (α=0.91), and service outcomes (α=0.89)—all exceeded the acceptable threshold of 0.8. Secondary indicators ranged between 0.78 and 0.87, with "Safety Assurance" (0.87) and "Student Development" (0.86) scoring highest, while "Funding Assurance" (0.78) was relatively lower, possibly due to parents' limited understanding of financial details. Split-half reliability yielded a corrected Spearman-Brown coefficient of 0.92, further confirming consistency.

Validity Testing:

Content Validity: Assessed via expert review and a pilot study, with a Content Validity Index (CVI) of 0.91. Feedback from 8 experts and 50 participants led to adjustments of 3 unclear indicators.

Construct Validity: Exploratory factor analysis (EFA) showed a KMO value of 0.90 and a significant Bartlett's test ($P<0.001$), supporting factor analysis suitability. Using principal component analysis, three factors were extracted, explaining 68.7% of the variance, consistent with the predefined dimensions. All indicators had loadings >0.6 on their respective factors and cross-loadings <0.3, confirming clear structures and reasonable factor divisions.

Criterion-Related Validity: Correlation analysis between system scores and annual education department assessments showed a significant positive correlation ($r=0.76$, $P<0.01$), indicating the system effectively reflects service quality. Differentiation analysis showed urban schools scored higher than rural schools ($t=3.24$, $P<0.01$), while secondary schools scored slightly higher than primary schools ($t=1.89$, $P>0.05$), aligning with regional resource disparities and confirming the system's discriminant validity.

## 7. Conclusions

This study develops a quality evaluation system for after-school physical education services in primary and secondary schools under the "Double Reduction" policy, yielding the following conclusions:

Integration of Policy and Practice: The evaluation system, based on three primary dimensions (service foundation, process, and outcomes) and 38 tertiary indicators, integrates "Double Reduction" policy requirements with practical needs. Indicators cover resource allocation, implementation, and outcomes, ensuring both specificity and comprehensiveness.

Scientific Methods Enhance System Reliability: The Delphi method and Analytic Hierarchy Process (AHP) enhanced scientific rigor. Expert consensus informed indicator selection, while AHP determined indicator weights, prioritizing service processes (weight: 0.40) and key indicators like "Safety Assurance" and "Physical Fitness Improvement," reflecting a "health first" philosophy.

Empirical Validation Confirms System Robustness: The system demonstrated high reliability (Cronbach's α=0.92) and validity. EFA confirmed the three-dimensional structure, and significant correlations with external benchmarks verified effectiveness. Differentiation analysis validated its ability to identify disparities across regions and school types.

## References

[1] Yang Qingyu, Zhang Juan. Study on Educational Needs of Ethnic Traditional Sports Courses in Southwest Frontier Universities [J]. Contemporary Sports Technology, 2023, 13(27): 1-6. DOI:10.12677/ae.2024.146916.

[2] Liu Chang, Wang Lei. Strategies for Improving the Quality of After-School PE Services in Primary and Secondary Schools under the "Double Reduction" Policy [J]. Journal of Physical Education, 2022, 29(4): 89-94. DOI:10.12677/ae.2024.146916.

[3] Li Hua, Chen Chen. Application of Delphi Method in Screening Educational Quality Evaluation Indicators [J]. China Journal of Education, 2021(7): 68-73. DOI:10.12677/ae.2024.146916.

[4] Zhang Wei, Zhao Li. Empirical Study on the Application of AHP in PE Education Evaluation [J]. Journal of Beijing Sport University, 2020, 43(8): 102-108. DOI:10.12677/ae.2024.146916.

[5] Wang Ming, Liu Yang. Construction of an Evaluation System for After-School PE Services from the Perspective of Service Quality Theory [J]. Journal of Sports Culture Guide, 2021(5): 112-117. DOI:10.12677/ae.2024.146916.

[6] Guo Qiang, Wu Min. Application of TQM Theory in PE Education Management [J]. Journal of Wuhan Sports University, 2020, 54(6): 85-90. DOI:10.12677/ae.2024.146916.

[7] Chen Hong, Zhou Yu. Research on After-School PE Service Quality Evaluation from the Perspective of Customer Satisfaction [J]. Journal of Chengdu Sport University, 2022, 48(3): 105-110. DOI: 10.12677/ae.2024.146916.

[8] Zhao Gang, Sun Li. Application of Educational Evaluation Theory in Assessing the Effectiveness of After-School PE Services [J]. Journal of Shanghai University of Sport, 2021, 45(2): 78-83. DOI:10.12677/ae.2024.146916.

[9] Liu Min, Zhang Tao. Collaborative Governance Mechanism for After-School PE Services under the "Double Reduction" Policy [J]. Sports Research, 2024, 38(1): 65-71. DOI:10.12677/ae.2024.146916.

[10] Wang Dan, Li Yang. Construction of Quality Evaluation Indicator System for After-School PE Services in Primary and Secondary Schools [J]. Sports Science, 2023, 43(9): 56-63. DOI:10.12677/ae.2024.146916.