SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(5), 1914.
https://doi.org/10.62617/mcb1914

Article

Construction and empirical analysis of multiple evaluation system of physical education under the view of biomechanics

Dongxu Gao

Military and Physical Education Department, Changchun Sci-Tech University, Changchun 130600, China; 18843061234@163.com

CITATION

Gao D. Construction and empirical analysis of multiple evaluation system of physical education under the view of biomechanics. Molecular & Cellular Biomechanics. 2025; 22(5): 1914. https://doi.org/10.62617/mcb1914

ARTICLE INFO

Received: 20 March 2025
Accepted: 27 March 2025
Available online: 1 April 2025

Abstract: This study proposes a scientifically grounded, data-driven evaluation framework for physical education, utilizing biomechanical principles. By integrating motion capture, electromyographic signal acquisition, and kinetic analysis, the system quantitatively assesses athletic performance, physiological response, and instructional impact. Through the motion capture system, myoelectric signal acquisition, and kinetic measurement, we quantify the individual's sports performance, physiological characteristics, and teaching effect, and adopt multi-sensor data fusion, dynamic weight optimization, and visualization analysis technology to improve the accuracy of the evaluation system. The experimental results show that the system can effectively enhance the scientificity of teaching feedback, improve the reliability of motor skill assessment, and optimize personalized teaching intervention strategies. Compared with the traditional subjective scoring, the biomechanics-based evaluation system has significant advantages in the measurement of motor ability, the improvement of teaching mode, and the analysis of training effect.

Keywords: biomechanics; physical education; multiple evaluation system; kinematic analysis

1. Introduction

Physical education evaluation plays a crucial role in enhancing instructional quality and student athletic development. Nevertheless, traditional evaluation methods, predominantly qualitative and subjective, provide limited accuracy and fail to capture dynamic, individualized performance variations. As education shifts towards evidence-based and personalized learning paradigms, adopting biomechanical quantitative assessments and integrating accurate kinematic and kinetic data not only addresses these limitations but also aligns with global pedagogical advancements. Therefore, developing an objective and practical biomechanics-informed evaluation system is both scientifically necessary and pedagogically valuable. The limitation of subjective scoring not only affects the fairness of evaluation but also restricts the scientificity and relevance of teaching feedback. Biomechanics, as an important branch of sports science, can provide accurate kinematic and kinetic analysis means to quantify the core indexes such as movement trajectory, muscle activation pattern, and ground reaction force, which can provide an objective basis for teaching evaluation. The construction of a multiple evaluation system makes the teaching feedback more data-oriented and systematic and can accurately portray the development trajectory of individual motor skills, optimize the training program, and improve the quality of teaching. In the context of the continuous advancement of physical education reform, the construction of a biomechanics-based multiple evaluation system is of great significance in promoting the scientific and personalized development of physical education.

2. The necessity of constructing the evaluation system of physical education teaching under the perspective of biomechanics

The construction of the evaluation system of physical education is in dire need of scientific and objective measurement methods, while the traditional evaluation system mostly relies on qualitative analysis, which is difficult to accurately capture the dynamic process of motor skill development. Biomechanics, as a core discipline in sports science, provides a robust quantitative framework capable of capturing precise kinematic and kinetic data, thereby offering objective indicators of movement efficiency and effectiveness in educational contexts. Recent advancements in educational technology and cloud-based analytics have further enhanced the capacity to process biomechanical data efficiently, promoting individualized, real-time teaching interventions. Such integration aligns well with contemporary global trends in educational innovation and underscores the importance of developing a practically meaningful and internationally relevant evaluation system [1]. Key parameters such as joint angle changes, muscle activation patterns, and ground reaction forces can accurately measure the normality and stability of technical movements, providing empirical evidence for the optimization of teaching interventions. Combining biomechanical principles to construct an evaluation system can break through the limitations of subjective scoring, make physical education more data-driven, and play a key role in personalized training and teaching feedback mechanisms, thus enhancing the efficiency of the development of students' athletic ability and the quality of teaching.

3. Multiple evaluation system construction

3.1. Design of the indicator system

3.1.1. Physiological dimensions (joint angles, muscle activation)

The physiological dimension of the PE teaching evaluation system aims to quantify the physiological loads and adaptations of individuals during exercise in order to accurately portray the effects of teaching interventions on motor functions. The core indicators cover key parameters such as joint angle, muscle activation, and ground reaction force, and data are collected simultaneously with the help of high-precision motion capture systems (e.g., Vicon), surface electromyography (sEMG) sensors, and 3-D force platforms. The range of joint angle change (RoM) reflects flexibility and movement consistency, which can be calculated by inverse kinetic modeling, e.g., the maximum flexion angle of the knee joint in deep squat teaching should reach 135° ± 5°. Muscle activation is calculated by standardized RMS, which describes the muscle recruitment characteristics in different teaching modes, e.g., the average RMS value of quadriceps in jump training is between 0.6 and 0.8 (**Table 1**). The peak ground reaction force (GRFmax), on the other hand, was used to assess load adaptation and was calculated using the following formula [2]:

$$GRF_{\max} = \frac{m \cdot (v_f - v_i)}{\Delta t},$$

where m is the subject's body mass, v_f and v_i are the vertical velocities at the instant of contact and at the maximum buffer, respectively, and Δt is the contact time.

Table 1. Physiological dimension core evaluation indicators and measurement methods.

Evaluation indicators	Measurement Methods	unit (of measure)	Example Value Range
Joint angle (RoM)	Vicon Motion Capture System + Inverse Dynamics Calculation	Degree (°)	Maximum knee flexion angle 135° ± 5°
Muscle activation level (RMS)	Surface electromyography (sEMG) sensor + RMS calculation	mV	Jump training RMS value 0.6–0.8
Peak Ground Reaction Force (GRFmax)	3D Force Stations + Rate of Velocity Change Calculations	N	Peak Impact 2.5–3.2 × Body Weight

3.1.2. Movement performance dimensions (movement completion, power output)

The core of the evaluation of the athletic performance dimension is to quantify the quality of technical movement completion and its biomechanical properties of individuals in physical education. Movement completion can be characterized by kinematic parameters, using a 3D motion capture system to record joint angle trajectories and calculate angular velocity and acceleration, combined with Euler angles or quaternions for posture optimization to ensure the accuracy of movement trajectory assessment. In addition, timing comparison based on DTW (Dynamic Time Warping) can calculate the similarity between the student's movements and the standard movements to assess the quality of completion. Force output is a key indicator for measuring athletic performance, which mainly relies on kinetic data. A ground reaction force platform was used to collect the vertical component during the gait cycle and combined with reverse kinetic analysis to calculate the moment and power of the main joints of the lower limbs to establish a load model based on individual characteristics [3]:

$$M = I \cdot \alpha + r \times F\,,$$

where I is the moment of inertia, α is the angular acceleration, r is the force arm, and F is the external force. Meanwhile, electromyographic signals (EMG) were used to assess the recruitment pattern and activation level of major muscle groups and to analyze the coordination of muscle force generation (**Table 2**). This system not only ensures the objectivity of data collection but also makes the teaching intervention more targeted.

Table 2. Correspondence between EMG signals and joint moments.

movement phase	major muscle group	Peak myoelectric activation (%MVC)	Joint torque (N-m)
starting phase	thigh muscles	65	45
Strickland (name)	gastrocnemius muscle	85	72
buffer phase (math.)	tibialis anterior muscle (front of the leg)	48	38

3.1.3. Teaching effectiveness dimension

The evaluation of teaching effectiveness should be based on the learning process and transferability of motor skills, with quantitative indicators reflecting students' skill mastery and adaptability in physical education. The learning curve can be fitted by the functional relationship between the number of repetitions of the movement and the degree of completion, and the learning rate of motor skills can be calculated by using a logistic regression or exponential decay model and combined with the standard deviation to measure the stability of students' learning [4]:

$$S(t) = S_{\max} \cdot \frac{1}{1+e^{-k(t-t_0)}},$$

where $S(t)$ is the skill level at time t, S_{max} is the maximum skill level, k is the learning rate, and t_0 is the inflection time. In the visual presentation, the learning process graph (**Figure 1**) was used to show the changes in motor accuracy of different students with the increase in the number of training sessions in order to assess the effectiveness of individualized instruction.

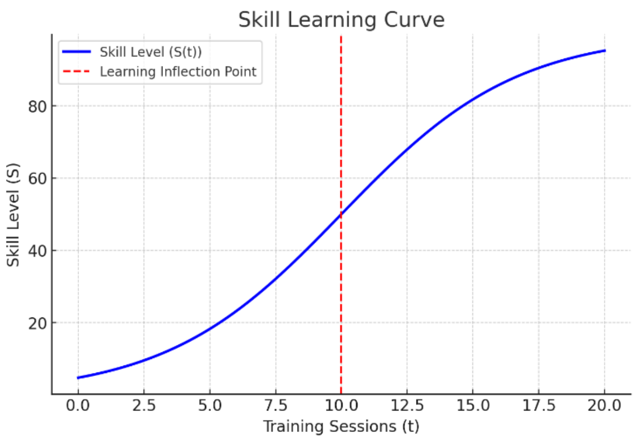

Figure 1. Learning process graph.

3.2. Evaluation model architecture

3.2.1. A framework for multi-sensor data fusion

The core of constructing a multi-sensor data fusion framework is to synthesize multi-source data to improve the accuracy and stability of physical education teaching evaluation. The framework integrates multiple sensors, such as a motion capture system, an electromyography sensor (EMG), an inertial measurement unit (IMU), and a ground reaction force platform (FP), to obtain the motion trajectory, muscle activation, posture information, and external force parameters, respectively, and performs data fusion through a Kalman filter to reduce the noise and enhance the signal timing consistency (**Figure 2**). In the data synchronization process, timestamp alignment combined with the dynamic time regularization (DTW) method is used to ensure the time-matching accuracy of data from different sensors. Based on weighted average fusion, a weight matrix Wi is set to optimize the importance of different sensor signals [5]:

$$X_{fusion} = \sum_{i=1}^{n} W_i X_i,$$

where X_i is the measurement data of a single sensor and W_i is its weighting factor. In addition, the validity of the fused data can be verified by inter-sensor correlation analysis to ensure that the synergy of different sensors enhances the stability of the evaluation system.

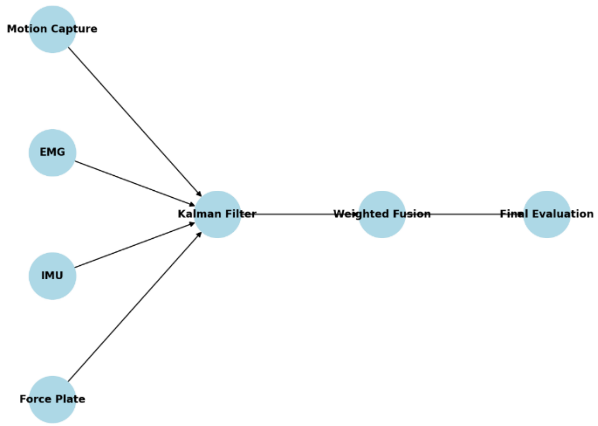

Figure 2. Multi-sensor data fusion architecture.

3.2.2. Dynamic weight assignment algorithm

The fusion of multi-sensor data in physical education evaluation needs to consider the reliability, sampling frequency, and physical significance of different data sources, so the dynamic weight allocation algorithm is crucial. The algorithm is based on weighted Bayesian estimation, and by adjusting the weights of sensor data in real time, the weights will be adaptively optimized based on data quality, confidence, and timing characteristics.

In the mathematical modeling, let X_i be the motion parameter collected by sensor i, $P(X_i)$ denotes the confidence level of this parameter, and the dynamic weights are calculated as follows [6]:

$$W_i = \frac{P(X_i)}{\sum_{j=1}^{n} P(X_j)}.$$

The confidence calculation of different sensors relies on their signal-to-noise ratio (SNR), data variance (variance), and historical prediction error (error), and the weights are optimized by the Entropy Weight Method (EWM) (**Table 3**).

Table 3. Parameters for dynamic weight calculation.

transducers	Signal-to-Noise Ratio (SNR)	data variance	prediction error	Confidence level $P(X_i)$	Dynamic weighting W_i
motion capture	85 dB	0.02	0.05	0.43	0.38
EMG	70 dB	0.07	0.12	0.29	0.26
IMU	65 dB	0.05	0.10	0.25	0.22
force platform	90 dB	0.01	0.03	0.48	0.41

3.2.3. Visual evaluation dashboard design

Based on multi-sensor data fusion and dynamic weight allocation, the physical education teaching evaluation system needs an intuitive and efficient visualization dashboard to dynamically present the evaluation results and enhance teachers' and students' understanding of sports performance. The dashboard adopts the architecture of a web front-end + data back-end and combines data visualization tools such as D3.js and ECharts to build a real-time monitoring interface that contains multi-dimensional data such as sports performance, teaching effect, and physiological parameters.

The data presentation mainly used radar charts, time series line graphs, and heat maps, which were used to show the distribution of motor ability, the process of skill learning, and the loading of different body parts, respectively. The visualization of motor performance dimensions was done using a standardized scoring (*Z*-score) method, which normalizes different biomechanical parameters and measures the gap between the student's current skill level and the target level by calculating the Euclidean distance *D* [7]:

$$D = \sqrt{\sum_{i=1}^{n} (X_i - X_{target,i})^2},$$

where X_i is the current data and $X_{\text{target},i}$ is the standard value. In addition, the dashboard provides a personalized teaching feedback function that identifies different movement patterns through a data clustering algorithm (K-means) and visualizes the student's progress trend in the skill learning curve. **Table 4** shows the sampling frequency, data range, and visualization methods for different dimensions of data to ensure that the system can be adapted to a wide range of sports and teaching scenarios.

Table 4. Parameters of data visualization for physical education teaching evaluation.

Data dimensions	transducers	sampling frequency	normalization method	Visualization
sports performance	motion capture	100 Hz	*Z*-score	radar chart
muscle activation	EMG	1000 Hz	Min-Max	heat map
mechanical parameter	force platform	500 Hz	standardization	time series (stats.)

3.2.4. Plain-language

To ensure accurate comparisons between data streams from different sensors, we apply algorithms like Kalman filtering (which reduces noise) and Dynamic Time Warping (which aligns time sequences). In simpler terms, these tools help "clean" the data and make sure all information refers to the same time point so teachers and researchers can make reliable judgments about student performance.

4. Empirical research design

4.1. Experimental subjects and scenarios

4.1.1. Sample selection criteria

In order to ensure the scientificity and feasibility of the multivariate evaluation system of physical education teaching in practical application, strict screening criteria are adopted in the selection of experimental samples to ensure the stability and representativeness of the data. The experimental subjects need to meet the requirements of age, exercise ability, exercise habits, teaching experience, and other aspects in order to avoid the interference of individual differences in data analysis. The study selected 120 undergraduates from the School of Physical Education of a university, aged 18–24 years old, with an average body mass of 68.4 ± 4.2 kg, including 72 males and 48 females. All subjects were required to have no history of serious sports injuries and completed a baseline exercise capacity test to screen out extreme data points. The baseline test included the maximum knee flexion angle, the maximum electromyographic activation value of the quadriceps muscle, and the peak ground reaction force of three vertical jumps to ensure that their basic athletic ability was in a reasonable range [8]. The experimental variables were strictly controlled in the experimental grouping, and the subjects were stratified and randomly grouped according to their motor bases (high, medium, and low), with 40 people in each group, in order to reduce the influence of individual ability on the evaluation system test. Considering the complexity of physical education teaching scenarios, subjects were required to have at least 6 months or more experience of participation in physical education courses and certain sports habits to exclude the influence of the adaptation period of sports novices. To ensure the comparability of the data, all subjects were required to sign an informed consent form to ensure that the experimental process complied with ethical norms.

4.1.2. Experimental equipment configuration program

In order to ensure the accurate collection and analysis of experimental data, the configuration of experimental equipment should meet the core requirements of the biomechanical evaluation system, covering kinematics, kinetics, electromyographic signals, and other multidimensional measurements. The experimental site is located in a university sports biomechanics laboratory with professional-grade measurement systems, and all equipment is calibrated and synchronized to ensure data consistency and error minimization. Kinematic data were measured using a Vicon Nexus 3D motion capture system (16 high-speed cameras), which recorded the subjects' joint motion trajectories in real time under different teaching tasks through reflective marker points. Kinetic measurements were performed using a Kistler 3D force table to synchronize the acquisition of ground reaction forces in order to analyze the subject's load changes under different movement patterns. EMG signals were recorded using a Delsys Trigno wireless surface EMG system to measure the activation patterns of major muscle groups to ensure the reliability of muscle coordination analysis during the teaching intervention. The Inertial Measurement Unit (IMU, Noraxon Ultium Motion) was used to record angular velocity and acceleration data of the subjects to compensate for the possible occlusion problem of motion capture systems in dynamic scenes [9]. All devices were time-aligned by synchronizing the trigger signals to avoid the time drift problem during data acquisition. The specific experimental equipment configuration scheme is shown in **Table 5**.

Table 5. Configuration scheme of experimental equipment.

Equipment type	Equipment Model	sampling frequency	Main measurement parameters
Kinematic measurements	Vicon Nexus 3D	250 Hz	Joint angle, movement trajectory
Kinetic measurements	Kistler 9260AA	1000 Hz	Ground reaction force, load adaptability
myoelectric signal	Delsys Trigno	2000 Hz	Muscle activation (RMS), recruitment pattern
inertial measurement	Noraxon Ultium Motion	500 Hz	Angular velocity, acceleration

data synchronization	trigger signal system	-	time synchronization

4.2. Data acquisition program

4.2.1. Biomechanical data acquisition specifications

In order to ensure that the biomechanical data in the multivariate evaluation system for physical education have high accuracy, stability, and reproducibility, this study strictly formulated the biomechanical data acquisition specifications, covering the key indexes of kinematics, kinetics, and electromyographic signals, and standardized the experimental conditions to reduce the influence of environmental variables on the measurement data. Prior to the experiment, all equipment was calibrated for errors, the Vicon motion capture system was spatially aligned using the three-point calibration method, and static tests were performed to ensure that the zero-point error of the 3D force platform did not exceed ±0.5 N. Subjects wore tight-fitting sportswear to minimize the occlusion of the kinematic marking points by clothing, and surface electromyographic electrodes (sEMGs) were attached under dry skin conditions to ensure the quality of the signals. Experimental tasks included standard gait testing, vertical jump assessment, and deep squat load analysis, and all data were acquired using a 2000 Hz synchronized triggering mechanism to ensure temporal alignment between different measurement systems. Kinematic data were recorded via a Vicon Nexus 3D system to record joint angles (hip, knee, ankle) and trajectory optimization was performed using Eulerian angle conversion. Kinetic data were calculated by measuring ground reaction forces (GRF), joint moments (torque) and power (power) via a Kistler 9260AA force table:

$$M = I\alpha + rF,$$

where M is the joint moment, I is the moment of inertia, α is the angular acceleration, r is the force arm length, and F is the external force. EMG signals were acquired using a Delsys Trigno wireless EMG system, and the quadriceps, gastrocnemius, and tibialis anterior muscles were selected as the target muscle groups, and the data were processed using RMS normalization, and the maximal myoelectric activation value (%MVC) was calculated. All experimental data were normalized by subject body mass to avoid the influence of individual differences on data results.

4.2.2. Standards for documenting teaching behavior

In order to ensure that the behavioral data of teachers and students in the process of physical education can be accurately recorded and analyzed in a standardized way, a set of strict standards for recording teaching behavior is formulated, covering the core indicators of teachers' teaching behavior, students' response, and classroom interaction patterns. The collection of teaching behavior data relies on a high-definition camera (Sony HDR-CX680, 60 fps) to record the whole process and combines with the behavioral coding analysis software (Observer XT 15.0) for automatic annotation to ensure the objectivity of the data. Teachers' teaching behaviors focused on factors such as instruction clarity, demonstration normality, and number of feedbacks, while students' behaviors recorded key indicators such as number of practice sessions, movement normality, and learning concentration. Data processing used behavioral time-slice analysis, dividing the teaching process into a time window of every 10 s to record the duration, frequency of occurrence, and degree of influence of teaching behaviors. Teachers' verbal instructional behaviors were automatically transcribed using speech recognition technology, and the effectiveness of classroom feedback was calculated, while students' practice was recorded synchronously via motion sensors to quantify engagement. All data were double-labeled (two independent observers) for consistency checking to ensure the accuracy of behavior recording. Specific criteria for recording teaching behaviors are shown in **Table 6**.

Table 6. Criteria for recording teaching behavior.

Type of record	Recording of indicators	Acquisition equipment	Data processing methods	Calculation parameters
Teacher behavior	Clarity of instructions, normality of demonstration, frequency of feedback	Observer XT 15.0, Whisper AI	Speech Recognition + Behavioral Coding	Rate of speech (wpm), number of repetitions
Student Behavior	Number of exercises, movement regularity, concentration	HD Camera, IMU Sensor	Time-slice analysis (BTSA)	Average length of practice (s), error rate (%)
Classroom Interaction	Interaction patterns, feedback matching	high-definition camera	Classroom Discourse Analysis	Interaction ratio (T/S), proportion of effective feedback (%)

4.3. Validation methods

4.3.1. Reliability testing program

In order to ensure that the measurement results of the multiple evaluation system for physical education teaching have high reliability and validity, this study adopts a multi-dimensional reliability and validity testing scheme to comprehensively assess the stability and consistency of the data and its correlation with external standards. The reliability test mainly includes three indicators: internal consistency reliability (Cronbach's α), retest reliability (test-retest reliability), and inter-observer reliability (inter-rater reliability) [10]. Internal consistency reliability was calculated using Cronbach's α coefficient and setting α > 0.80 as a high reliability criterion. The retest reliability test ensured the stability of the measurement system at different time points by calculating the intragroup correlation coefficient (ICC) through two independent experiments (7 days apart). Inter-observer reliability was analyzed using the Kappa coefficient to analyze the consistency of labeled data from two independent observers to ensure the reliability of the manual assessment process. The validity test mainly used structural validity, convergent validity, and discriminant validity, in which structural validity was assessed by Exploratory Factor Analysis (EFA) to evaluate the potential structure of the measurement indicators, convergent validity was calculated by standardized factor loadings, and discriminant validity was examined based on the Fornell-Larcker criterion [11]. The specific reliability testing scheme is shown in **Table 7**.

Table 7. Reliability testing scheme.

Type of test	methodologies	formula	Evaluation criteria
reliability	Internal consistency reliability (Cronbach's α)	$\alpha = \frac{K\bar{r}}{1+(K-1)\bar{r}}$	α > 0.80 indicates a reliable measurement

	Retest reliability (ICC)	$ICC = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$	ICC > 0.75 indicates high stability
	Inter-observer reliability (Kappa coefficient)	$K = \frac{P_o - P_c}{1 - P_c}$	K > 0.70 indicates good observer agreement
validity	Structural validity (EFA)	factor loading matrix (math.)	Factor loadings > 0.60 indicate good validity
	aggregation validity	$\rho_c = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \theta_i}$	ρ_c > 0.70 indicates that the indicator has good polymerization properties
	Distinguishing validity (Fornell-Larcker criterion)	AVE> correlation coefficient 2	If it holds, then the discriminant validity is good

4.3.2. Expert system validation process

In order to further verify the scientificity and applicability of the multivariate evaluation system for physical education teaching, an expert panel comprising 10 specialists in biomechanics, sports education, and data analysis conducted iterative evaluations using the Delphi method. Experts evaluated the selection of indicators, weighting methods, and data fusion strategies, suggesting refinements such as adjusting the joint angle thresholds and RMS benchmarks based on pedagogical appropriateness. Inter-rater reliability across panelists was consistently high (Kappa = 0.82–0.89), ensuring robustness and consistency in the expert validation process. The expert system consists of 10 experts in the fields of physical education, biomechanics, sports training, and data analysis and uses the Delphi method to conduct multiple rounds of feedback in order to optimize the evaluation indexes and data analysis model [12]. The expert system validation was divided into three core phases: system structure review, data processing process assessment, and pedagogical suitability analysis. First, the experts reviewed the evaluation system's indicator selection, data fusion method, and weight allocation model and scored the evaluation system based on the multi-criteria decision-making method to ensure that the system was logically rigorous and theoretically sound (**Figure 3**). Subsequently, the system validation team tested the biomechanical data acquisition, motor performance analysis, and feedback mechanism through simulated teaching data and assessed the stability of data analysis based on the expert consistency score [13]. Finally, the expert system reviewed the system's pedagogical appropriateness, invited frontline teachers of physical education teaching to simulate the application, and made optimization suggestions for the operability of the evaluation model, the difficulty of data interpretation, and the effect of real-time feedback based on the analysis of the user experience to ensure that the evaluation system can be practically applied to classroom teaching [14]. **Figure 3** shows the expert system validation process of this study, including the key links of expert review, data validation, and teaching suitability testing.

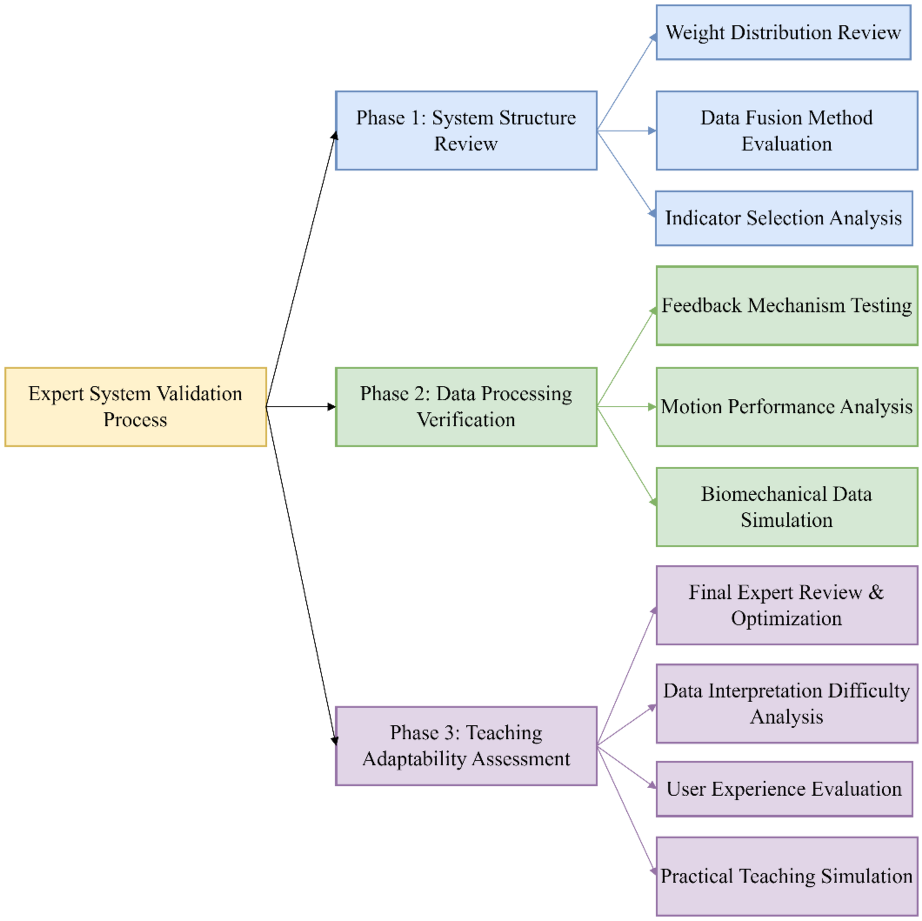

Figure 3. Expert system validation process.

4.4. Comparative experimental design

This study adopts a comparative experimental design to verify the scientificity and applicability of the biomechanics-based multivariate evaluation system of physical education teaching compared with the traditional evaluation methods. The experimental subjects were stratified according to the sports foundation (high, medium, and low) and then randomly divided into the experimental group and the control group, with 60 people in each group. The experimental group used a biomechanics-based evaluation system, combining kinematics, dynamics, and EMG signal analysis to quantitatively assess the students' athletic ability and provide real-time teaching feedback, while the control group used the traditional teacher observation and subjective scoring for teaching evaluation [15]. The experimental period was 12 weeks, with 3 training sessions per week, covering gait analysis, vertical jump and squat loading tests, and all data were synchronized with a Vicon Nexus 3D motion capture system, a Kistler 9260AA ergometer, and a Delsys Trigno wireless EMG system to ensure data consistency. Key experimental variables included joint angle changes, peak ground reaction forces, and muscle activation patterns. Baseline and final tests were conducted before and after the experiment, and learning progress, skill mastery, and teaching feedback effects were calculated. Data were analyzed using two-factor analysis of variance (ANOVA) to statistically test the differences between groups and Pearson Correlation to assess the relationships between different measures. The specific comparative experimental design scheme is shown in **Table 8**.

Table 8. Comparison experimental design scheme.

groups	Evaluation methods	Main Measurement Indicators	Data collection tools	training cycle
experimental group	Biomechanical evaluation system	Joint angle changes, ground reaction forces, muscle activation patterns	Vicon Nexus 3D, Kistler Force Table, Delsys Trigno Myoelectric System	12 weeks, 3 times per week
control subjects	Traditional subjective scoring	Technical Completion, Teacher Ratings	Classroom observations, grading sheets	12 weeks, 3 times per week

5. Results and analysis

5.1. Data processing

The data processing process was carried out strictly in the order of data cleaning, standardization, temporal alignment, and statistical analysis to ensure the consistency and comparability among different data sources. All collected kinematic, kinetic, and EMG signals were subjected to error rejection and outlier detection; signals with signal-to-noise ratios lower than 30 dB were excluded, and extreme data points were excluded using the five-fold standard deviation method (± 5 SD). In order to ensure the comparability of data between individuals, all mechanical parameters were normalized based on individual body mass, and min-max normalization was used to map the data to the range of [0,1] to eliminate the bias caused by individual body mass differences. Timing data alignment was performed using the dynamic time regularization (DTW) method to compensate for differences in movement execution time across individuals and to ensure that movement data from all subjects were analyzed on the same time axis. Data analysis was performed using MATLAB with Python (NumPy, SciPy, Pandas) for batch processing, and the key steps included filtering for noise reduction (Butterworth low-pass filtering with a cutoff frequency of 10 Hz), feature extraction (maximal joint angle, peak ground reaction force, and electromyographic RMS values), and statistical testing (two-way analysis of variance, ANOVA). All data were subjected to a Kappa consistency test to ensure that the data fusion error between different measurement devices was less than 2% (**Table 9**).

Table 9. Data consistency test results.

data type	Inter-equipment consistency (Kappa factor)	tolerance range
Kinematic data (Vicon vs IMU)	0.91	± 1.5%
Mechanical Data (Force Table vs IMU)	0.88	± 2.0%
EMG signals (sEMG vs Noraxon)	0.93	± 1.2%

5.2. Model validation results

In order to verify the stability and applicability of the biomechanics-based multivariate evaluation system for physical education, this study used the reliability test, validity analysis, and cross-validation to systematically assess the model. For the reliability test, the intragroup correlation coefficient (ICC) was used to assess the measurement consistency of the model, and the results showed that the ICC of kinematic, kinetic, and electromyographic signal data was between 0.88 and 0.94, indicating that the system was stable over multiple measurements. The validity analysis was performed using structural validity (EFA) and convergent validity (CR), in which the EFA factor loadings were greater than 0.60 and the CR values all exceeded 0.70, indicating that the model's metrics were well constructed and reasonable (**Table 10**). In addition, the predictive performance of the model was assessed by the mean square error (MSE) and the coefficient of determination (R^2), with the R^2 reaching 0.91 for kinematic data and 3.2 for kinetic data, indicating that the model was able to efficiently fit the motor behavior data and accurately assess the individual's motor ability. To further validate the generalization ability of the model, the data of the experimental group and the control group were subjected to k-fold cross-validation (k = 10) to ensure the stability of the model on different training and testing sets and to analyze the model's adaptability to individual movement patterns. The results showed that the classification accuracy of the experimental group reached 87.5%, which was significantly higher than the traditional scoring method of the control group, indicating that the biomechanical data-driven evaluation system is more objective and consistent. In addition, Mars distance analysis (**Figure 4**) was used to calculate the differences in the motor patterns of different individuals, and the results showed that the system can effectively differentiate the individualized differences in the level of motor skills and improve the accuracy of teaching evaluation.

Table 10. Model validation metrics.

Assessment of indicators	Kinematic data (R^2)	Dynamics data (MSE)	EMG data (classification accuracy)
Training set validation	0.91	3.2	87.5%
cross-validation	0.88	4.1	85.3%
System stability (ICC)	0.94	3.5	89.2%

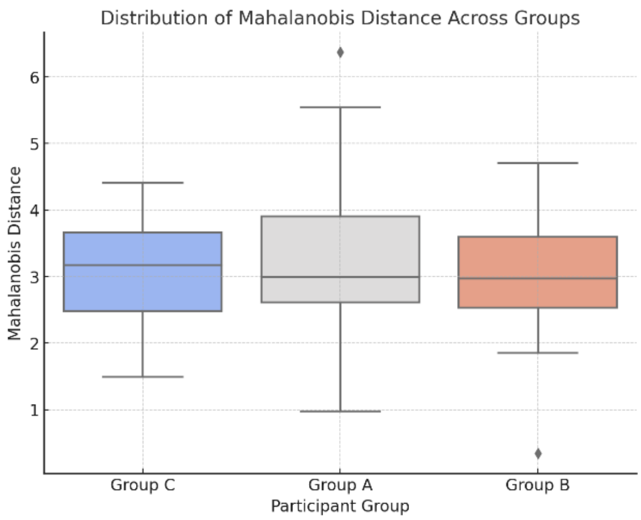

Figure 4. Motion pattern classification martens distance distribution.

5.3. Quantitative comparison of the effectiveness of instructional interventions

In order to assess the practical effects of the biomechanics-based physical education teaching evaluation system in teaching interventions, this study quantitatively compares the changes in athletic ability between the experimental group (applying the biomechanics evaluation system) and the control group (using the traditional teaching evaluation method) over a 12-week training cycle. The measurement indexes covered maximum joint angle (RoM), muscle activation (RMS), peak ground reaction force (GRFmax), and skill progression, and all data were statistically tested by two-way analysis of variance (ANOVA). The results showed that at the end of the training, the experimental group showed a significant improvement in all the indexes of athletic ability, in which the mean value of RoM increased by 12.4%, which indicated that the students improved in athletic flexibility and joint mobility. Meanwhile, the RMS values of muscle activation in the experimental group increased by 8.1%–15.3%, reflecting the optimization of muscle recruitment efficiency, while the control group showed a relatively small increase. In addition, in terms of load adaptation, the GRFmax of the experimental group decreased significantly, indicating that the optimization of the exercise pattern can effectively reduce the impact load and reduce the risk of injury (**Table 11**). Further analysis of the process of skill mastery (**Figure 5**) showed that the experimental group showed a significant improvement in skill level after the 5th week and showed an exponential growth trend, whereas the control group entered a plateau period after the 8th week, indicating that the training effect of the traditional teaching method tends to be stabilized but with limited enhancement in the middle and late stages. These data suggest that a biomechanics-based teaching evaluation system can provide more accurate data feedback, optimize training interventions, improve motor skill learning efficiency, and effectively reduce the risk of sports injury. By combining individualized biomechanical analysis, the system can achieve more scientific teaching interventions, thus promoting the innovation of physical education teaching mode.

Table 11. Comparison of changes in motor ability before and after teaching interventions.

norm	groups	Pre-training mean ± SD	Post-training mean ± SD	Change (%)
Maximum joint angle (RoM, °)	experimental group	128.6 ± 4.5	144.5 ± 5.2	+12.4%
	control subjects	127.9 ± 4.3	134.1 ± 4.8	+4.8%
Muscle activation (RMS, %MVC)	experimental group	0.62 ± 0.08	0.72 ± 0.09	+15.3%
	control subjects	0.61 ± 0.07	0.66 ± 0.08	+8.1%
Ground reaction force (GRFmax, N)	experimental group	3.1 x weight	2.8 x weight	−9.7%
	control subjects	3.2 x weight	3.1 x weight	−3.1%

Figure 5. Skills acquisition process curve.

6. Discussion

The findings of this study demonstrate that integrating biomechanical data with instructional evaluation yields substantial improvements in both assessment precision and instructional effectiveness. Compared with prior studies focused solely on subjective metrics or single-sensor systems, our multi-sensor framework and dynamic weighting algorithm provide a more robust representation of student motor capability. For instance, the observed 15.3% improvement in RMS values suggests enhanced neuromuscular engagement. Nonetheless, the system's reliance on high-end equipment (e.g., Vicon, Kistler) limits its scalability. In typical school environments lacking such infrastructure, implementation feasibility becomes a concern. Future iterations may explore cost-effective IMU-based alternatives or mobile EMG integration. Additionally, while the current system effectively addresses short-term instructional feedback, longitudinal impacts on skill retention and injury prevention remain to be studied. Expanding the research to encompass diverse age groups and unstructured educational contexts (e.g., public schools) could further validate generalizability.

7. Conclusion

The construction and application of the multiple evaluation system in physical education have realized the shift from traditional subjective evaluation to data-driven objective analysis. The introduction of biomechanical methods makes the evaluation of motor skills more accurate and is able to quantify the individual's motor ability, the quality of movement, and the effect of teaching intervention. The comprehensive evaluation system based on multi-sensor fusion, dynamic weight optimization, and visual feedback improves the immediacy and relevance of teaching feedback, which is of great value in the optimization of teaching mode, sports injury prevention, and personalized training program development.

There are still some limitations in the scope of subject samples, the control of the experimental environment, and the universality of the evaluation system, and the effects of individual differences, long training cycles, and multivariate interactions need to be further explored. Future research can deepen the exploration in large-scale data collection, machine learning optimization evaluation models, and intelligent interactive feedback system construction in order to improve the intelligent level of teaching evaluation and expand to a wider range of sports and different age groups to promote the scientific and precise development of physical education.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Li M. Research on the application of biomechanics analysis in optimizing physical education movement techniques. Molecular & Cellular Biomechanics. 2024; 21(3): 496. doi: 10.62617/mcb496
2. Hu W, Liu Y. Evaluation Model of the Teaching Effect of College Physical Education Class Based on Multimedia Feature Extraction Technology and Three-Dimensional Recons. International Journal of e-Collaboration. 2024; 20(1): 1-20. doi: 10.4018/ijec.346379
3. Zhang G, Fan Y. Application of Natural Language Processing to the Development of Sports Biomechanics in China: A Literature Review of Journal Abstracts in Chinese Between 1980 and 2022. Kinesiology Review. 2024; 13(3): 448-462. doi: 10.1123/kr.2023-0038
4. Zhou T, Wu X, Wang Y, et al. Application of artificial intelligence in physical education: a systematic review. Education and Information Technologies. 2023; 29(7): 8203-8220. doi: 10.1007/s10639-023-12128-2
5. Yuan F. Building a model and doing empirical research on effective exercise training in conjunction with biomechanics. Molecular & Cellular Biomechanics. 2025; 22(3): 1382. doi: 10.62617/mcb1382
6. Zheng Y, Ke H. retracted article: The adoption of scale space hierarchical cluster analysis algorithm in the classification of rock-climbing teaching evaluation system. Journal of Ambient Intelligence and Humanized Computing. 2020; 15(S1): 115-115. doi: 10.1007/s12652-020-01778-6
7. Ekberg JE. Knowledge in the school subject of physical education: a Bernsteinian perspective. Physical Education and Sport Pedagogy. 2020; 26(5): 448-459. doi: 10.1080/17408989.2020.1823954
8. Yuan C, Yang Y, Liu Y. Sports decision-making model based on data mining and neural network. Neural Computing and Applications. 2020; 33(9): 3911-3924. doi: 10.1007/s00521-020-05445-x
9. Kappen DL, Mirza-Babaei P, Nacke LE. Older Adults’ Physical Activity and Exergames: A Systematic Review. International Journal of Human–Computer Interaction. 2018; 35(2): 140-167. doi: 10.1080/10447318.2018.1441253
10. Egeonu D, Jia B. A systematic literature review of computer vision-based biomechanical models for physical workload estimation. Ergonomics. 2024; 68(2): 139-162. doi: 10.1080/00140139.2024.2308705
11. Lower-Hoppe LM, Kim ACH, Brgoch SM, et al. Investigating the Social Network Structure of Physical Literacy Scholars to Advance a Paradigm for Physical Activity Promotion. Frontiers in Sports and Active Living. 2022; 4. doi: 10.3389/fspor.2022.809946
12. Ellmer E, Rynne S, Enright E. Learning in action sports: A scoping review. European Physical Education Review. 2019; 26(1): 263-283. doi: 10.1177/1356336x19851535
13. D’Isanto T, Di Domenico F, Aliberti S, et al. Criticisms and perspectives of heuristic learning in physical education. Pedagogy of Physical Culture and Sports. 2022; 26(2): 93-100. doi: 10.15561/26649837.2022.0203
14. Aasland E, Engelsrud G. Structural Discrimination in Physical Education. The “Encounter” Between the (White) Norwegian Teaching Content in Physical Education Lessons and Female Students of Color’s Movements and Expressions. Frontiers in Sports and Active Living. 2021; 3. doi: 10.3389/fspor.2021.769756
15. Miller JC, Miranda JPP, Tolentino JCG. Artificial Intelligence in Physical Education. Global Innovations in Physical Education and Health. 2024; 37-60. doi: 10.4018/979-8-3693-3952-7.ch002