

The Evaluation System and Continuous Improvement of College Physical Education Teaching in Guangdong Province

Rongping Zhu *

Graduate School, Adamson University, CO 1000, Manila, Philippines

* Corresponding author Email: 18824700829@163.com

Abstract: The purpose of this study is to explore the construction of the evaluation scale of college physical education teaching in Guangdong Province, to ensure that the evaluation tool can provide accurate, reliable and effective evaluation results, thus supporting and guiding the improvement of the quality of college physical education teaching in Guangdong Province. Through the analysis of the validity of evaluation tools, the validity and reliability of the scale, the research found that most of the items were effective in evaluating the ability and performance of PE teachers. The result of factor analysis shows that the items of evaluation tools can be divided into six main factors: 1) Continuous Learning and Improvement; 2) Instructional Strategies and Assessment; 3) Collaborative Learning Environment; 4) Technology Integration; 5) Subject Knowledge and Curriculum Alignment; 6) Professional Ethics and Growth. In addition, it also verified the consistency of evaluation between teachers and managers and between different work units.

Keywords: Physical Education Teaching; Evaluation System; Performance Evaluation.

1. Introduction

College physical education plays an important role in cultivating students' physical fitness, teamwork, leadership skills and lifelong exercise habits. However, ensuring the quality and effectiveness of college physical education teaching is a challenge for educational institutions. Assessing the quality of university physical education teachers involves not only student learning outcomes, but also aspects such as teaching content, professionalism, digital literacy and pedagogical knowledge whether their physical education goals are achieved, help teachers improve their teaching methods, and also provide students with feedback to help them better understand their learning progress. However, establishing an effective university physical education teaching evaluation system is not a simple task because it requires taking into account multiple dimensions, different stakeholders, and multiple evaluation methods. Currently, there are problems in university physical education teaching in Guangdong Province such as lack of scientific guidance in teaching management, narrow adaptability of evaluation indicators, and incompatibility of evaluation results with reality.

Therefore, in order to solve the current problems, on the basis of domestic and foreign experience, combined with the actual situation of university physical education teaching in Guangdong Province, factors affecting teaching quality are identified, and research methods such as expert consultation and analytic hierarchy process are used to repeatedly verify and select the indicators to establish an evaluation model to provide suggestions for the sustainable development of university physical education science in Guangdong Province and promote the healthy, scientific and orderly rapid development of school physical education in my country.

To adapt to the new changes in the reform and development of school sports in the new era. General Secretary Xi Jinping (2017) pointed out: "Sports is an important symbol of social

development and human progress, and an important manifestation of comprehensive national strength and social civilization." The spirit of the report of the 19th National Congress of the Communist Party of China elaborates on the new requirements for the development of school sports in the new era, that is, "education must be in a primary position, deepen educational reform, accelerate educational modernization, and ensure that students can get more sense of gain from the development of school sports. Our country is carrying out drastic reforms in education, and its goals, missions, and methods are all After being reformed, physical education teaching, as a part of education, has also been reformed. With the progress of the reform, because there is no updated evaluation system that matches the teaching goals, missions, and methods, evaluation problems have become more and more prominent. In our country's foundation with the smooth implementation of education reform policies and guidelines, the purpose of physical education, the tasks and models of physical education have all undergone relatively profound reforms. Physical education is an essential component of school education that plays a vital role in promoting the healthy development of students' physical and mental health. It is also one of the main platforms for cultivating sports talents, and as such, it is a top priority to construct a Physical Education Teacher Evaluation Tool that identifies the knowledge, skills and behaviors needed to provide sound instruction in the physical education classroom.

The tool can be used to assess teacher's competency needed to perform a variety of physical activities, understanding of movement, concepts and principles strategies and tactics as they apply to the learning and performance of physical activities. It also emphasizes the importance of appropriate infrastructure, meaningful content defined by the curriculum, appropriate instructional practices, including classroom management, student and program by assessment and evaluation. It can also be used for self-assessment by teachers. The tool can be customized to target areas in professional

growth plan, provide in service programs to help teachers address point of emphasis or areas needed for improvement

The physical education teaching evaluation is a summary and analysis of the physical education teaching process. It has a good guiding role in the physical education teaching work. Through the teaching evaluation, the school physical education teaching can be taken to a higher level so that school administrators can more accurately grasp all the problems that arise in school physical education.

The physical education teaching evaluation in this study is the identification and diagnosis of the practical activities of teachers and students in teaching, aiming to better promote the high-quality development of physical education teaching in ordinary colleges and universities.

2. Statement of the Problem

This study will evaluate physical education teaching in some universities in Guangdong Province, discover the influencing factors of university physical education teaching, and then build an evaluation system for physical education teaching in universities in Guangdong Province. Specifically, it will answer the following questions:

- 1.What is the item validity of the proposed evaluation tool
- 2.What is the scale validity of the proposed evaluation tool
- 3.What is the reliability of the proposed evaluation tool
- 4.Based on the research results, what Performance Evaluation or Teachers Behavior Inventory Assessment tool can be developed

3. Scope and Delimitation of the Study

This study will be conducted at selected higher education institutions in Guangdong Province, China. It involved a random sample of universities from one hundred and sixty (160) universities in Guangdong Province, China, .221 faculty members or teachers and 36 administrators were selected as respondents

This study focuses on the factors that affect the evaluation of university physical education teaching.

4. METHODOLOGY

4.1. Research Design

Exploratory Analysis Research design will be utilized Exploratory research is often used when the issue you're studying is new or when the data collection process is challenging for some reason. You can use this type of research if you have a general idea or a specific question that you want to study but there is no preexisting knowledge or paradigm with which to study it³

4.2. Research Local

The study will be conducted in Guangdong Province and will involve (221) teachers and experts from (5) universities in Guangdong Province, (36) professors and experts were randomly selected from the twelve (12) universities to conduct a questionnaire survey. There are 36 experts in total, with the title of professor, with the title of associate professor, and lecturer. The following universities were selected because they have developed well in physical education and have sufficient development motivation, sufficient talent resource support, and experts and professors have sufficient physical education teaching experience.

Table 1. List of experts

University	Title	Years of Teaching Experience(year)
Shantou University	Professor	16-20
Guangzhou University	Professor	16-20
Lingnan Normal University	Professor	16-20
South China Normal University	Associate Professor	6-10
Jinan University	Associate Professor	6-10
South China University of Technology	Associate Professor	11-15
South China Agricultural University	Associate Professor	16-20
Shenzhen University	Associate Professor	16-20
Guangzhou College of Commerce	Associate Professor	11-15
Nanfeng College Guangzhou	Lecturer	11-15
Guangzhou Software Institute	Lecturer	6-10
Guangdong University of Technology	Lecturer	6-10

4.3. Sampling Technique

Stratified sampling will be used to select the 221 teacher participants while total enumeration will be utilized to select the administrator participants.

Table 2. Teacher sampling

Name of University	Population PE teachers	n PE Teachers
Guangzhou University	101	45
Lingnan Normal University	111	50
South China Normal University	144	65
South China University of Technology	69	31
Guangdong University of Technology	68	30
Total	493	221

Table 3. Administrator sampling

Name of University	Administrators (n)
Guangzhou University	3
Lingnan Normal University	3
South China Normal University	3
Jinan University	3
South China University of Technology	3
South China Agricultural University	3
Guangdong University of Technology	3
Shenzhen University	3
Shantou University	3
Guangzhou College of Commerce	3
Nanfeng College Guangzhou	3
Guangzhou Software Institute	3
Total	36

4.4. Validation of Instrument

After suggestions and comments are provided to the consultant, a draft will be drawn up for approval so that team

members can verify the questionnaire.

5. Results and Analysis

Table 4. Content Validity Ratio

Items	CVR	Decision
1.Engage in continuous learning and improvement to enhance her professional knowledge skills and values	1.0	Retain
2.Demonstrates the skills or concepts that she /he wants students to learn this helps students know what is expected of them and how to apply this knowledge	1.0	Retain
3.Adapts her instructions and assessment to meet the diverse needs and abilities of students	1.0	Retain
4.Divides the class into small groups and assign them tasks or projects that require communication and collaboration	1.0	Retain
5.Helps students develop their social and interpersonal skills, creativity and leadership skills.	1.0	Retain
6.Use formative and summative assessment to monitor the students progress, identify gaps in their knowledge to be able to adjust their teaching accordingly	1.0	Retain
7.Provide students with information about their performance progress or understanding the content of Physical Education	1.0	Retain
8.Use various communication strategies to engage students in meaningful discussion or interaction.	1.0	Retain
9.Foster a positive and supportive learning environment that promotes collaboration and creativity among students	1.0	Retain
10.Establishes and maintain clear and consistent rules for student behavior in Physical Education	1.0	Retain
11. Creates a safe supportive and inclusive learning environment that fosters positive relationships among students and between students and teachers.	1.0	Retain
12.Uses positive reinforcement and corrective feedback to promote student engagement and regulation.	1.0	Retain
13.The teacher can explain the important concepts and skills in their subject area clearly and accurately	1.0	Retain
14The teacher is accurate in providing feedback or correcting errors in the assignments or test	1.0	Retain
15..Teachers knowledge of the subject is relevant to the learning goals.	1.0	Retain
16.Teacher connects the subject to other disciplines or real world situations.	1.0	Retain
17.Teacher is knowledgeable of the subject and its underlying principles and theories	1.0	Retain
18..Teacher is flexible in adapting the content or instructions to meet the diverse needs and abilities of the students	1.0	Retain
19. Develop and apply lesson based on effective course of study by integrating effective literacy instruction. throughout the curriculum and across content areas to enhance student learning	1.0	Retain
20. The teacher can design lesson and unit plans that reflect the curriculum standards and the pre-requirement relationships among the topics.	1.0	Retain
22. The teacher can grade the students test or assignments in a consistent fair manner, using appropriate criteria or rubric.	1.0	Retain
23.Use the cloud to allow teachers to create digital assignments, share resources with students, and provide feedback on student work.	1.0	Retain
24.Use the digital assessment tools to measure student learning and provide feedback.	1.0	Retain
25.Teachers use technology to teach the rules skills and tactics of a and use class time to coach students support or game	1.0	Retain
26. Act as a designer and create new learning experiences that use technology and to personalize instruction challenge the students to think more deeply and make learning more active engaging and authentic.	1.0	Retain
27 Students are placed in groups and given projects to complete. they use technology to track their physical activity and fitness level	1.0	Retain
28.Integrate technology across the curriculum not just as a separate subject or activity.	1.0	Retain
29.Use technology to differentiate instruction and meet the diverse needs of the learners .	1.0	Retain
30.Use technology to foster creativity innovation and problem solving skills in students	1.0	Retain
31.Know how to operate hardware software , trouble shoot problems or create digital content	1.0	Retain
32.Provide guidance feedback or opportunities for students to explore , collaborate or create with digital tools.	1.0	Retain
33.. Use technology to demonstrate analyze and correct the students movements and skills	1.0	Retain
34..Apply ethical principles and values in teaching practice such as honesty , fairness and responsibility	1.0	Retain
35. The teacher demonstrate a comprehensive understanding of the subject matter and the appropriate methods to teach Physical Education	1.0	Retain
36.The teacher designs and implement curriculum and instructions that are aligned to with the standards , objectives and outcomes of physical Education	1.0	Retain
37.The teacher establishes and maintain effective communication and collaboration with students colleagues, school administration and other stakeholders	1.0	Retain
38..The teacher embraces challenges innovations and opportunities for growth and improvement	1.0	Retain
39.The teacher possesses the necessary knowledge and skills and attitudes in Physical Education	1.0	Retain
40. The teacher contribute to the vitality of teaching profession	-0.6	Remove
41.The teacher engage in continuous learning to improve his or her own practice and contribute to the advancement of the profession	1	Retain
42The teacher. Provide students with assessments	0.6	Remove
43The teacher delivers effective instruction that meets the standards of quality teaching in Physical Education	1.0	Retain
44The teacher respects the rights and dignity of students and colleagues	1.0	Retain
45. The teacher respects and responds to the diverse needs and strengths interest and experiences of the learners.	-0.6	Remove
46.The teacher adapt learning experience in Physical Education	-0.6	Remove
47. The teacher communicate information about Physical Education to students	0.2	Remove
48. The teacher address the needs of all learners	0.6	Remove
49The teacher participate in local and global communication	-0.6	Remove
50.The teacher use learning content strategies	0.6	Remove

The Cronbach Alpha reliability index will be used to further evaluate and validate the questionnaire to further examine and analyze its consistency and reliability. At least 20 experts and professors will receive a draft and will be asked to provide answers to determine whether the content of the questionnaire are clear, concise, accurate, reliable and understandable for content verification. The experts' suggestions will be incorporated to improve the instrument.

This section presents the gathered data represented in tabular form as well as its analysis and interpretation. The findings contained are based on the results of the statistical treatment applied using statistical software particularly IBM-SPSS v27.

Research Question 1: What is the item validity of the

proposed evaluation tool?

Content Validity Assessment

According to Morgado et al. (2018), content validity is the "adequacy with which a measure assesses the domain of interest" (Hinkin, 1995). If the items are to measure the things they are meant to measure, sufficient content is essential (Devellis, 2012). Additionally, content validity outlines content representations and relevance, meaning that the items accurately reflect the pertinent experiences of the target population (McPhail, 2007). The target population and experts' evaluations will be used to accomplish this. Each item is assessed by knowledgeable judges to determine if it fits within the domain of interest.

Table 5. Overall KMO and Bartlett's Test

KMO	0.980
Bartlett's Test of Sphericity	$\chi^2 = 8866.444$ df= 861 p = 0.000

Table 6. Communalities

Items	Initial	Extraction	Items	Initial	Extraction
1	.715	.641	22	.720	.637
2	.744	.633	23	.755	.650
3	.766	.667	24	.706	.645
4	.783	.698	25	.685	.584
5	.790	.722	26	.740	.677
6	.680	.616	27	.778	.702
7	.780	.633	28	.776	.740
8	.721	.641	29	.752	.661
9	.773	.647	30	.776	.659
10	.782	.718	31	.813	.747
11	.762	.658	32	.790	.703
12	.703	.619	33	.760	.664
13	.725	.597	34	.773	.690
14	.732	.678	35	.752	.671
15	.711	.655	36	.750	.683
16	.740	.638	37	.788	.702
17	.730	.648	38	.757	.645
18	.769	.673	39	.768	.662
19	.740	.675	40	.794	.744
20	.784	.717	41	.758	.633
21	.751	.670	42	.743	.668

CVR, which gauges an item's essentiality, was the second

kind of empirical analysis (Yamada et al., 2010). A higher

score denotes greater agreement among panel members, with CVR ranging from 1 to -1. The CVR formula is $CVR = (N_e - N/2)/(N/2)$, where N is the total number of panelists and N_e is the number of panelists who have indicated that a particular item is "essential" (Zamanzadeh et al., 2015). Formalized scaling and statistical procedures, such as the content validity index for measuring proportional agreement and the content validity ratio for quantifying consensus (Lawshe, 1995), will be used to quantify their evaluations (Lynn, 1986).

Based on the analysis of CVR values as rated by 6 experts, out of 50 items, 7 items will be removed since the computed CVR falls below the minimum threshold prescribed by Lawshe (1995), which is .99 for 6 experts. Specifically, items number 40, 42, and 45 to 50 will be removed and the remaining 42 items will be subjected to construct validity.

Research Question 2: What is the scale validity of the proposed evaluation tool?

Construct Validity Assessment

In estimating the sample size in performing Exploratory Factor Analysis, 200 respondents are considered fair according to Comrey and Lee (1992). There were 200

respondents involved in the pilot study.

Step 1: Extract Factors

To determine the factors influencing physical education instruction, an exploratory factor analysis was carried out. Bartlett's sphericity test and the Kaiser-Meyer-Olkin (KMO) test will be used to assess the suitability of the sample. KMO values between 0.7 and 0.8 and between 0.8 and 0.9, respectively, are considered appropriate for large sample sizes (Nia et al., 2014). The significance of the results ($\chi^2 = 8866.444$, $p = 0.000$) suggests that factor analysis is appropriate. The data's suitability for factor analysis is indicated by the Kaiser-Meyer-Olkin measure of sampling adequacy (MSA), which came in at 0.980. Data that have MSA values greater than 0.800 are deemed suitable for factor analysis in this sense.

Then, through principal component analysis and varimax rotation, latent constructs were extracted. Items with communalities lower than 0.4 will be eliminated from factor analysis, as the factor loading threshold will be set at 0.4 or higher (Samitsch, 2014). Analysis of the data revealed that all the communalities are over .40.

Table 7. Statistical indices of exploratory factor analysis using components analysis

	Rotated Component Matrix ^a					
	Component					
	1	2	3	4	5	6
1.	.392	.180	.438	.406	.364	.244
2.	.361	.472	.321	.417	.282	.119
3.	.521	.355	.167	.383	.168	.415
4.	.307	.369	.290	.581	.146	.299
5.	.334	.313	.513	.328	.347	.274
6.	.358	.291	.305	.466	.391	.177
7.	.330	.320	.243	.393	.293	.489
8.	.257	.306	.391	.591	.285	.157
9.	.348	.350	.337	.143	.591	.214
10.	.510	.369	.278	.327	.317	.300
11.	.377	.148	.277	.592	.261	.344
12.	.089	.269	.585	.436	.301	.296
13.	.654	.219	.181	.256	.392	.212
14.	.368	.363	.292	.475	.185	.274
15.	.307	.178	.420	.230	.272	.575
16.	.203	.447	.399	.332	.148	.466
17.	.649	.236	.289	.283	.242	.246
18.	.482	.328	.181	.392	.273	.358
19.	.340	.278	.330	.215	.289	.572
20.	.339	.364	.295	.307	.359	.375
21.	.406	.491	.358	.290	.145	.277
22.	.344	.428	.169	.415	.380	.274
23.	.267	.681	.320	.204	.237	.286
24.	.255	.285	.201	.326	.494	.468
25.	.538	.297	.426	.151	.226	.291
26.	.260	.504	.243	.266	.451	.356
27.	.426	.365	.475	.266	.190	.322
28.	.325	.413	.393	.377	.297	.305
29.	.360	.472	.358	.314	.298	.180
30.	.301	.571	.168	.420	.302	.245
31.	.310	.405	.325	.474	.185	.373
32.	.470	.351	.276	.431	.068	.421
33.	.646	.255	.269	.310	.253	.238
34.	.482	.412	.443	.231	.266	.215
35.	.308	.369	.533	.292	.375	.125
36.	.265	.427	.411	.244	.436	.296
37.	.504	.473	.415	.243	.202	.134
38.	.393	.234	.225	.471	.434	.179
39.	.426	.140	.311	.385	.546	.252
40.	.369	.376	.459	.257	.339	.328
41.	.335	.210	.696	.236	.147	.308
42.	.461	.345	.440	.314	.260	.182
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.						
a. Rotation converged in 20 iterations.						

The table provided contains the rotated component matrix resulting from a Principal Component Analysis (PCA) with Varimax rotation and Kaiser normalization. This analysis was conducted to identify underlying factors or components in the dataset consisting of 42 items related to various aspects of physical education teaching. The rotated component matrix

displays the loadings of each item on the extracted components, with values ranging from -1 to 1. Higher absolute values indicate stronger relationships between items and components. Each component represents a set of related items that tend to vary together.

Table 8. Rotated component matrix of each item using exploratory factor analysis with component analysis method

Items	Item Loading
1. Engage in continuous learning and improvement to enhance her professional knowledge skills and values	0.80
2. Demonstrates the skills or concepts that she /he wants students to learn this helps students know what is expected of them and how to apply this knowledge	0.80
3. Adapts her instructions and assessment to meet the diverse needs and abilities of students	0.82
4. Divides the class into small groups and assign them tasks or projects that require communication and collaboration	0.84
5. Helps students develop their social and interpersonal skills, creativity and leadership skills.	0.85
6. Use formative and summative assessment to monitor the students progress, identify gaps in their knowledge to be able to adjust their teaching accordingly	0.78
7. Provide students with information about their performance progress or understanding the content of Physical Education	0.81
8. Use various communication strategies to engage students in meaningful discussion or interaction.	0.80
9. Foster a positive and supportive learning environment that promotes collaboration and creativity among students	0.80
10. Establishes and maintain clear and consistent rules for student behavior in Physical Education	0.85
11. Creates a safe supportive and inclusive learning environment that fosters positive relationships among students and between students and teachers.	0.81
12. Uses positive reinforcement and corrective feedback to promote student engagement and regulation.	0.79
13. The teacher can explain the important concepts and skills in their subject area clearly and accurately	0.77
14. The teacher is accurate in providing feedback or correcting errors in the assignments or test	0.82
15. Teachers knowledge of the subject is relevant to the learning goals.	0.81
16. Teacher connects the subject to other disciplines or real-world situations.	0.80
17. Teacher is knowledgeable of the subject and its underlying principles and theories	0.81
18. Teacher is flexible in adapting the content or instructions to meet the diverse needs and abilities of the students	0.82
19. Develop and apply lesson based on effective course of study by integrating effective literacy instruction. throughout the curriculum and across content areas to enhance student learning	0.82
20. The teacher can design lesson and unit plans that reflect the curriculum standards and the pre-requirement relationships among the topics.	0.85
21. The teacher can grade the students test or assignments in a consistent fair manner, using appropriate criteria or rubric.	0.82
22. Use the cloud to allow teachers to create digital assignments, share resources with students, and provide feedback on student work.	0.80
23. Use the digital assessment tools to measure student learning and provide feedback.	0.81
24. Teachers use technology to teach the rules skills and tactics of a and use class time to coach students support or game	0.80
25. Act as a designer and create new learning experiences that use technology and to personalize instruction challenge the students to think more deeply and make learning more active engaging and authentic.	0.76
26. Students are placed in groups and given projects to complete. they use technology to track their physical activity and fitness level	0.82
27. Integrate technology across the curriculum not just as a separate subject or activity.	0.84
28. Use technology to differentiate instruction and meet the diverse needs of the learners.	0.86
29. Use technology to foster creativity innovation and problem solving skills in students	0.81
30. Know how to operate hardware software, trouble shoot problems or create digital content	0.81
31. Provide guidance feedback or opportunities for students to explore, collaborate or create with digital tools.	0.86
32. Use technology to demonstrate analyze and correct the students movements and skills	0.84
33. Apply ethical principles and values in teaching practice such as honesty, fairness and responsibility	0.81
34. The teacher demonstrates a comprehensive understanding of the subject matter and the appropriate methods to teach Physical Education	0.83
35. The teacher designs and implement curriculum and instructions that are aligned to with the standards, objectives and outcomes of physical Education	0.82
36. The teacher establishes and maintain effective communication and collaboration with students' colleagues, school administration and other stakeholders	0.83
37. The teacher embraces challenges innovations and opportunities for growth and improvement	0.84
38. The teacher possesses the necessary knowledge and skills and attitudes in Physical Education	0.80
39. The teacher engage in continuous learning to improve his or her own practice and contribute to the advancement of the profession	0.81
40. The teacher delivers effective instruction that meets the standards of quality teaching in Physical Education	0.86
41. Respects the rights and dignity of students and colleagues	0.80
42. The teacher respects and responds to the diverse needs and strengths interest and experiences of the learners.	0.82

Absolute value of .32 or higher was considered sufficiently high to assume a strong relationship between variable/item

and factor. Factor loadings less than .32 in absolute value were regarded as insignificant and the items containing such

loadings were removed from the scale (Costelo & Osborne, 2005).

Based on the data from Tables 4 and 5, we can categorize these items into six main factors, namely Continuous Learning and Improvement, Instructional Strategies and Assessment, Collaborative Learning Environment, Technology Integration, Subject Knowledge and Curriculum Alignment, and Professional Ethics and Growth. These factors reflect the important characteristics and behaviors exhibited by teachers in physical education teaching. Here are the items included in each factor:

Component 1: Continuous Learning and Improvement. It reflects the importance of continuous learning and improvement for teachers. Including Engage in continuous learning and improvement to enhance her professional knowledge skills and values, The teacher can explain the important concepts and skills in their subject area clearly and accurately ect.

Component 2: Instructional Strategies and Assessment, emphasizes effective instructional strategies and assessment methods. Including Demonstrates the skills or concepts that she/he wants students to learn this helps students know what is expected of them and how to apply this knowledge, Use formative and summative assessment to monitor the students progress, identify gaps in their knowledge to be able to adjust their teaching accordingly ect.

Component 3: Collaborative Learning Environment, highlights the creation of a collaborative learning environment. Including Divides the class into small groups and assign them tasks or projects that require communication and collaboration, Helps students develop their social and

interpersonal skills, creativity and leadership skills, Foster a positive and supportive learning environment that promotes collaboration and creativity among students ect.

Component 4: Technology Integration, underscores the integration of technology in teaching. Including use various communication strategies to engage students in meaningful discussion or interaction, Use the cloud to allow teachers to create digital assignments, share resources with students, and provide feedback on student work ect.

Component 5: Subject Knowledge and Curriculum Alignment, pertains to subject knowledge and alignment with curriculum standards. Including Establishes and maintain clear and consistent rules for student behavior in Physical Education, Teachers knowledge of the subject is relevant to the learning goals, Teacher connects the subject to other disciplines or real-world situations ect.

Component 6: Professional Ethics and Growth, focuses on professional ethics and ongoing growth. Adapts her instructions and assessment to meet the diverse needs and abilities of students. Including Uses positive reinforcement and corrective feedback to promote student engagement and regulation, The teacher can design lesson and unit plans that reflect the curriculum standards and the pre-requirement relationships among the topics ect.

The PCA reveals underlying factors influencing physical education teaching practices. Teachers should prioritize continuous learning, employ effective instructional strategies, foster collaboration, integrate technology, align with curriculum standards, and uphold professional ethics. These findings can inform the development of training programs and improvement initiatives in physical education teaching.

Table 9. Differences in respondents' evaluation of physical education when respondent identity as a test factor

Indicators	Respondent	Mean	SD	Computed t-value	Sig	Decision on Ho	Interpretation
Continuous Learning and Improvement	Teacher	4.09	0.73	0.22	0.83	Accepted	Not Significant
	Administrator	4.12	0.75				
Instructional Strategies and Assessment	Teacher	4.10	0.69	0.74	0.46	Accepted	Not Significant
	Administrator	4.20	0.67				
Collaborative Learning Environment	Teacher	4.14	0.67	0.87	0.39	Accepted	Not Significant
	Administrator	4.25	0.70				
Technology Integration	Teacher	4.17	0.70	0.56	0.58	Accepted	Not Significant
	Administrator	4.25	0.70				
Subject Knowledge and Curriculum Alignment	Teacher	4.30	0.53	0.67	0.53	Accepted	Not Significant
	Administrator	4.23	0.62			Accepted	Not Significant
Professional Ethics and Growth	Teacher	4.21	0.60	0.62	0.43	Accepted	Not Significant
	Administrator	4.1	0.55			Accepted	Not Significant
Over-all	Teacher	4.16	0.65	0.61	0.53	Accepted	Not Significant
	Administrator	4.19	0.66				

Through the t-test, it was found that the average score for teachers was 4.16, with a standard deviation of 0.65, while for administrators, the average score was 4.19, with a standard deviation of 0.66. The calculated t-value was 0.61, with a significance of 0.53. Therefore, the null hypothesis is accepted, indicating that there is no significant difference in the overall evaluation of physical education between teachers and administrators.

Thus, there are almost no significant differences in the evaluation of various aspects of physical education between

teachers and administrators. This suggests that teachers and administrators generally hold similar views and opinions when evaluating physical education. Additionally, whether it is continuous learning and improvement, instructional strategies and assessment, collaborative learning environment, technology integration, or subject knowledge and curriculum alignment, both teachers and administrators have provided relatively consistent evaluations. This indicates that their evaluation of physical education considers multiple dimensions rather than being limited to a single aspect.

Table 10. Differences in respondents' evaluation of physical education when respondent identity as a test factor

Indicators	University	Mean	SD	Computed F-value	Sig	Decision on Ho	Interpretation
Continuous Learning and Improvement	1	4.19	0.83	1.61	0.20	Accepted	Not Significant
	2	4.10	0.68				
	3	3.50	0.66				
	4	3.98	1.18				
	5	4.01	0.40				
	6	4.11	0.72				
	7	4.08	0.59				
	8	4.30	0.62				
	9	3.92	1.05				
	10	4.14	0.39				
	11	4.27	0.90				
	12	4.08	0.47				
Instructional Strategies and Assessment	1	4.28	0.52	0.51	0.60	Accepted	Not Significant
	2	4.13	0.95				
	3	3.856	0.47				
	4	4.02	0.87				
	5	4.14	0.50				
	6	4.27	0.54				
	7	4.09	0.96				
	8	3.37	0.43				
	9	4.06	0.82				
	10	4.13	0.49				
	11	4.32	0.55				
	12	4.17	0.96				
Collaborative Learning Environment	1	4.21	0.37	0.67	0.52	Accepted	Not Significant
	2	4.19	0.79				
	3	3.80	0.78				
	4	4.30	0.48				
	5	3.99	0.51				
	6	4.11	0.78				
	7	4.24	0.94				
	8	4.10	0.53				
	9	4.31	0.69				
	10	4.17	0.96				
	11	3.81	0.98				
	12	3.91	0.66				
Technology Integration	1	4.17	0.95	1.15	0.32	Accepted	Not Significant
	2	3.81	0.96				
	3	4.49	0.62				
	4	4.21	0.94				
	5	3.86	1.05				
	6	4.57	0.63				
	7	4.23	0.93				
	8	3.91	0.81				
	9	3.40	0.55				
	10	4.19	0.79				
	11	4.24	0.65				
	12	3.70	0.48				
Subject Knowledge and Curriculum Alignment	1	4.20	0.76	0.98	0.38	Accepted	Not Significant
	2	4.16	0.60				
	3	3.71	0.64				
	4	4.16	0.83				
	5	4.13	0.68				
	6	4.23	0.66				
	7	4.27	0.79				
	8	4.03	0.62				
	9	4.06	0.75				
	10	4.18	0.76				
	11	3.96	0.63				
	12	4.03	0.78				
Professional Ethics and Growth	1	4.24	0.79	2.14	0.26	Accepted	Not Significant
	2	3.89	0.65				
	3	4.14	0.48				
	4	4.21	0.76				
	5	4.00	0.60				
	6	4.11	0.64				
	7	3.87	0.88				
	8	4.19	0.44				
	9	4.19	0.84				
	10	4.28	0.78				
	11	4.29	0.54				
	12	3.97	0.92				
Over-all	1	4.22	0.70	1.18	0.38	Accepted	Not Significant
	2	4.05	0.77				
	3	3.92	0.61				
	4	4.15	0.84				
	5	4.02	0.62				
	6	4.23	0.66				
	7	4.13	0.85				
	8	3.98	0.58				
	9	3.99	0.78				
	10	4.18	0.70				
	11	4.15	0.71				
	12	3.98	0.71				

The results for continuous learning and improvement, instructional strategies and assessment, collaborative learning environment, technology integration, subject knowledge and curriculum alignment, professional ethics and growth, and overall evaluation are similar. There are differences in the mean evaluations of physical education among respondents from different work units to some extent. However, the calculated F-values and p-values did not reach statistical significance, indicating that the differences are not significant. Therefore, we accept the null hypothesis, which states that the differences in the evaluation of physical education among

respondents from different work units, such as continuous learning and improvement, are not statistically significant.

Regarding the above results, although differences exist in some indicators, these differences are not sufficient to suggest significant disparities in the evaluation of physical education among respondents from different work units. This may indicate that regardless of the respondents' work units, their evaluations of physical education are generally consistent, and the differences in work units do not significantly affect the evaluation results.

Table 11. Fit Indices

Table 11. Fit Indices						
Test for Exact Fit						
χ^2			df		p	
1261.37			819		< .001	
Fit Measures						
			RMSEA 90% CI			
CFI	TLI	RMSEA	Lower	Upper	AIC	BIC
0.95	0.95	0.05	0.05	0.06	12361.95	12776.91

Table 12. Factor Loadings

Factor	Indicator	Estimate	SE	Z	p
Factor 1	1	0.64	0.05	13.65	< .001
	2	0.63	0.05	13.53	< .001
	3	0.71	0.05	14.10	< .001
	4	0.68	0.05	14.55	< .001
	5	0.76	0.05	14.95	< .001
	6	0.61	0.05	13.26	< .001
	7	0.66	0.05	13.99	< .001
	8	0.62	0.05	13.63	< .001
	9	0.66	0.05	13.72	< .001
	10	0.73	0.05	14.89	< .001
	11	0.69	0.05	13.94	< .001
	12	0.60	0.04	13.32	< .001
	13	0.58	0.04	12.94	< .001
	14	0.68	0.05	14.24	< .001
	15	0.67	0.05	13.89	< .001
	16	0.65	0.05	13.63	< .001
	17	0.63	0.05	13.76	< .001
	18	0.67	0.05	14.15	< .001
	19	0.67	0.05	14.22	< .001
	20	0.73	0.05	14.86	< .001
	21	0.69	0.05	14.14	< .001
	22	0.68	0.05	13.59	< .001
	23	0.67	0.05	13.82	< .001
	24	0.63	0.05	13.71	< .001
	25	0.59	0.05	12.76	< .001
	26	0.66	0.05	14.24	< .001
	27	0.70	0.05	14.62	< .001
	28	0.76	0.05	15.25	< .001
	29	0.68	0.05	13.98	< .001
	30	0.66	0.05	13.97	< .001
	31	0.77	0.05	15.37	< .001
	32	0.73	0.05	14.67	< .001
	33	0.68	0.05	14.01	< .001
	34	0.72	0.05	14.43	< .001
	35	0.68	0.05	14.13	< .001
	36	0.72	0.05	14.35	< .001
	37	0.70	0.05	14.62	< .001
	38	0.65	0.05	13.73	< .001
	39	0.64	0.05	13.95	< .001
	40	0.75	0.05	15.31	< .001
	41	0.62	0.05	13.57	< .001
	42	0.66	0.05	14.07	< .001

Research Question 3: What is the item validity of the proposed evaluation tool?

There are several ways of assessing model fit. The first is a χ^2 -statistic that, if small, indicates that the model is a good fit to the data. These are the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI) and the Root Mean Square Error of Approximation (RMSEA) together with the 90% confidence interval for the RMSEA. Some useful rules of thumb are that a satisfactory fit is indicated by CFI > 0.9, TLI > 0.9, and RMSEA of about 0.05 to 0.08. A good fit is CFI > 0.95, TLI > 0.95, and RMSEA and upper CI for RMSEA < 0.05. Based on the computed values, all the items are good

items.

For convergent validity, the researcher used Factor Loadings to help the researcher determine if the items on the instrument are understood by the participants in the way that they were intended to by the creator of the test (Kock, 2017). The item loadings must be greater than 0.5, and the p-value of each item is less than 0.05, then the researcher can say that the instrument has a high convergent validity (Hair et al., 1987). Analysis of the data revealed that all the item loadings are greater than 0.50 and all the p-values are less than the 0.05 level of significance. Hence, it can be concluded that the scale satisfied the convergent validity.

Table 13. Scale Reliability Statistics

	Cronbach's α
Physical Education Teaching	0.98

Table 13 presents the reliability measurement, particularly the Cronbach's alpha of the unidimensional scale, the physical education teaching scale. Cronbach's alpha, abbreviated as CA, is a statistic used to assess construct reliability or internal consistency (Roldan & Sanchez-Franco, 2012; Kock, 2015). For Cronbach's alpha to be able to determine whether or not a measurement is reliable, the value of the measurement being analyzed must be equal to or greater than 0.70 (Fornell & Larcker, 1981; Nunnally, 1978). The analysis of the data showed that all the items satisfied the criterion for reliability and were considered to be good items with good internal consistency based on CA value of CA = 0.98.

6. Summary of Results

This study explored the construction of an evaluation scale for university physical education teaching in Guangdong Province, aiming to ensure that the assessment tools can provide accurate, reliable, and effective evaluation results, thereby supporting and guiding the improvement of the quality of university physical education teaching in Guangdong Province.

1. Item Validity of the Proposed Evaluation Tool

Through data analysis, it can be observed that the vast majority of the 50 items (42 items) were rated as having item validity, as their CVR values were all 1.0, indicating unanimous agreement among experts that these items are effective for assessing the capabilities and performance of physical education teachers, and are consistent with the purpose and intent of the evaluation tool.

However, 8 items were deemed to lack item validity, as their CVR values were below .99, and were thus decided to be removed. Most of these items (6 items) had negative CVR values, indicating that experts generally considered them unsuitable for assessing the capabilities of physical education teachers. These items may have ambiguities or lack relevance to the core objectives of physical education teaching.

2. Scale Validity of the Proposed Evaluation Tool

Exploratory Factor Analysis (EFA) was conducted to assess the scale validity of the proposed evaluation tool, based on the data analysis of Research Question 2. Firstly, the suitability of the sample was evaluated through Bartlett's sphericity test and Kaiser-Meyer-Olkin (KMO) test. The

results showed that Bartlett's sphericity test had a p-value of 0.000, and the KMO value was 0.980, indicating that the sample was suitable for factor analysis.

Subsequently, factor analysis was performed, and the table provided factor loadings for the 42 items, which are components of the latent factors extracted through EFA. It was observed that most items had high factor loadings, ranging from 0.7 to 0.9 in absolute value. This indicates a strong correlation between these items and the extracted factors, making them crucial for assessing key aspects of physical education teaching.

However, several items had lower factor loadings, ranging from 0.5 to 0.6 in absolute value. Although their loadings were still within an acceptable range, they might have lower contributions to the constituted factors. Therefore, further exploration may be needed in future studies to determine their actual importance in evaluating physical education teaching.

The consistency in the evaluation of physical education between teachers and administrators may reflect their shared pursuit and expectations of educational quality. It also indicates that teachers and administrators have similar values and goals when evaluating physical education, both aiming to provide high-quality educational services. Furthermore, analyzing the consistency of evaluation results between teachers and administrators helps understand their consensus and differences regarding physical education, providing references for educational management, teacher training, and teaching improvement.

Moreover, by comparing the evaluations of teachers and administrators, areas that require attention and improvement can be identified, thereby further enhancing the quality and effectiveness of physical education. Therefore, the comparative analysis results of respondent identities are of significant importance in understanding the evaluation of physical education by teachers and administrators, offering valuable insights and guidance for educational management and practices.

Respondents from different workplaces, that is, different universities, showed consistency in their evaluations of physical education. The evaluations of physical education by respondents from different workplaces were generally consistent, and the differences in workplaces did not significantly impact the evaluations. This is an important finding for physical education managers and decision-makers,

as it allows them to consider the overall educational needs more when formulating educational policies and allocating resources, rather than focusing on specific workplace evaluation differences.

3. Reliability of the Proposed Evaluation Tool

1) Item Validity

Through the analysis of fit indices and factor loadings, all items exhibited good validity. Regarding fit indices, the χ^2 statistic was small (1261.37), indicating a good fit between the model and the data. Both CFI and TLI were greater than 0.95, indicating good comparative fit indices for the model. RMSEA and its 90% confidence interval were within an acceptable range, indicating a small approximation error for the model and a good fit.

Factor loadings indicated that the estimated values for all items were greater than 0.50, with p-values less than 0.05, meeting the criteria for item validity. This suggests a significant association between all items and the extracted factors.

2) Convergent Validity

Through the analysis of factor loadings, all items met the criteria for convergent validity. The estimated factor loadings for all items were greater than 0.50, and all p-values were less than 0.05, indicating a high level of convergent validity for the evaluation tool, i.e., consistency between how respondents understand the items in the evaluation tool and the creators' intent.

3) Reliability

The internal consistency of the evaluation tool was measured by Cronbach's α coefficient. In this study, Cronbach's α coefficient was 0.98, well above the standard value of 0.70. The Cronbach's α coefficient values for all items met the reliability standards, indicating high internal consistency of the evaluation tool, i.e., consistent performance of items in measuring the same concept or construct.

In conclusion, despite slight differences in certain evaluation indicators, overall, respondents from different workplaces showed similar trends in their evaluations of physical education. This suggests that respondents may have been influenced by similar factors when assessing physical education, whether from teaching experience, professional backgrounds, or educational policies. The data indicate that factors such as continuous learning and improvement, instructional strategies and assessment, and collaborative learning environments received high average scores from respondents, with minimal differences across different workplaces. This indicates that these factors are perceived as significant contributors to the quality of physical education by respondents. However, this evaluation consistency may be influenced by various factors, such as consistency in educational policies, uniformity in teaching standards, and commonalities in teacher professionalism. These factors may contribute to the convergence of evaluations of physical education among respondents from different workplaces. Based on the finding of evaluation consistency, managers and decision-makers can consider overall educational needs more comprehensively, rather than focusing excessively on the evaluation results of specific workplaces. When formulating educational policies and allocating resources, a more holistic approach can be adopted to meet the educational needs of a broad audience.

7. Conclusion

The results derived from this study lead the researchers to draw the following conclusions:

1. Most of the proposed items in the evaluation tool were deemed effective and suitable for assessing the performance of physical education teachers. However, there are still some items that require further modification or removal to enhance the accuracy and effectiveness of the evaluation tool. This finding holds significant implications for improving the design and implementation of physical education evaluation tools, ensuring that they accurately reflect the actual performance of physical education teachers and provide support and guidance for enhancing teaching quality.

2. The factor loading table provides crucial information regarding the scale validity of the proposed evaluation tool, indicating a strong correlation between most items and the extracted factors, thus supporting the tool's validity. The proposed evaluation tool demonstrates good scale validity as it captures key factors in physical education teaching and explains the variability in most of the data. Therefore, this tool is effective in assessing the teaching abilities and performance of physical education teachers.

3. Through the analysis of these data, it can be concluded that the evaluation tool exhibits excellent results in terms of item validity, scale validity, and reliability. These findings suggest that the evaluation tool is a reliable and effective instrument for assessing the teaching evaluation of university physical education.

4. In conclusion, the study revealed the effectiveness of physical education assessment tools and the consistency among teachers, administrators, and respondents from different workplaces in evaluating physical education. These findings provide valuable references for educational management, teacher training, and instructional improvement, contributing to the enhancement of the quality and effectiveness of physical education.

8. Recommendations

1. Continuous Professional Development Workshops: Organize workshops focusing on areas such as continuous learning and improvement, effective communication strategies, and integration of technology in teaching. These workshops can help instructors enhance their skills in these areas, thereby improving their performance in key aspects of physical education teaching.

2. Diversified Instructional Strategies Training: Offer training sessions on adapting instructions and assessments to meet diverse student needs and abilities. Providing instructors with strategies for creating inclusive learning environments and fostering collaboration among students can contribute to more effective teaching practices.

3. Technology Integration Training: Develop training programs to help instructors effectively integrate technology into their teaching, including using digital assessment tools, creating digital assignments, and leveraging technology to personalize instruction. This can improve instructors' technological competence and enable them to utilize digital tools to enhance student engagement and learning outcomes.

4. Feedback and Assessment Workshops: Conduct workshops on providing constructive feedback, using formative and summative assessments to monitor student progress, and grading assignments consistently and fairly. Equipping instructors with assessment best practices can lead

to more accurate evaluation of student performance and a more supportive learning environment.

5.Cultural Competency Training: Offer cultural competency training to help instructors respect and respond to the diverse needs, strengths, interests, and experiences of students. Enhancing instructors' cultural awareness can promote inclusivity and equity in physical education classrooms, leading to better student engagement and success.

6.Regular Communication and Collaboration Platforms: Establish regular communication and collaboration platforms for instructors, students, colleagues, and administrators to exchange ideas, share resources, and discuss teaching practices. This can facilitate knowledge sharing, foster a supportive professional community, and promote continuous learning and improvement among instructors.

7.Utilization of Student Feedback Mechanisms: Implement student feedback mechanisms to gather input on teaching effectiveness, course content, and learning experiences. Analyzing student feedback can provide valuable insights for instructors to identify areas for improvement and make necessary adjustments to their teaching approaches.

8.Ongoing Evaluation and Reflection: Encourage instructors to engage in ongoing self-reflection and evaluation of their teaching practices against established standards and objectives. Providing opportunities for instructors to review their performance, set goals for professional development, and seek feedback from peers can support continuous

improvement in physical education teaching.

References

- [1] Xing Wei & Zhang Yingqiu. (2023). The value implication, goal orientation and practical path of physical education teaching reform in colleges and universities. *Educational Theory and Practice* (06), 58-61.
- [2] Jiang Yong & Chen Chuang. (2022). Factors influencing the quality of physical education teaching in colleges and universities and optimization paths from the perspective of learning paradigm. *Anhui Sports Science and Technology* (01), 73-77.
- [3] Sun Juan & Wang Yue. (2014). Misunderstandings and resolutions of physical education teaching in colleges and universities from the perspective of the modern university system. *Journal of Nanjing Institute of Physical Education (Social Science Edition)* (05), 6-11. doi:10. 15877/j. cnki . nsic. 2014.05.002.
- [4] Song Guoliang. (2012). Discussion on the current situation and development of physical education teaching in colleges and universities. *China Science and Education Innovation Guide* (05), 209.
- [5] Wang Ling. (2012). Comprehensive thinking on optimizing the physical education teaching model in colleges and universities. *Education and Occupation* (02), 112-114. doi:10. 13615/j.cnki. 1004-3985.2012.02.088.